TOWARDS DEMYSTIFICATION OF MODEL-FREE TECHNICAL ANALYSIS: THE POWER OF FEEDBACK CONTROL

By

Shirzad Malekpour

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(ELECTRICAL AND COMPUTER ENGINEERING)

at the

UNIVERSITY OF WISCONSIN–MADISON

2016

Date of final oral examination: 12/21/2015

The dissertation is approved by the following members of the Final Oral Committee:

B. Ross Barmish, Professor, Electrical and Computer Engineering

John A. Gubner, Professor, Electrical and Computer Engineering

Nigel Boston, Professor, Electrical and Computer Engineering & Mathematics

David P. Brown, Professor, School of Business

James A. Primbs, Associate Professor, Finance, California State University, Fullerton.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Professor Bob Barmish, who has been a great teacher, mentor and inspiration. The first lesson I learned from him was the importance of having a "clear mind" in research. Professor Barmish patiently taught me basic research skills and showed me the power of creativity, especially in developing a new line of research. His integrity and high standard for work ethics makes him my ideal role model.

I am grateful to Professors Nigel Boston, David Brown, John Gubner and James Primbs for serving on my dissertation committee and providing great feedback. Each of them had an important and pivotal role in my academic life at UW-Madison. I took my first graduate-level courses with Professor Boston and Professor Gubner. Their amazing teaching skills and passion for research inspired me. My regular chats about Football (soccer) with Nigel will always be fondly remembered. I am thankful to Professor Brown who taught me the first principles of investment theory and patiently answered my "naive" questions. My first collaborative research project was carried out with Professor Primbs. His brilliance and excellent communication skills taught me the importance of these qualities in collaborative research.

I am so grateful to have many people who supported me in the course of pursuing my PhD. I was so blessed to be raised by my parents, Mina and Parviz. Their major contributions to society through their careers in academia encouraged me towards higher education. I am also thankful to my sisters, Shirin and Sheida, along with their families who made this journey easier while I was away from my wife. During the last months, their hospitality made me feel at home which helped me focus on my thesis.

The long hours spent at my desk were interrupted by daily coffee breaks with brilliant friends. We chatted, joked and brainstormed. My friends Amin Farmahini-Farahani and Laura Balzano were regular coffee break companions. They listened to my concerns and made me feel energized throughout the day. I am blessed to have such great friends.

Finally and foremost, I would like to express my gratitude to my wife, Anaram, who has constantly and unconditionally supported me throughout these years. Her kind words, reminding me of the hope for a bright future together, helped me go through tough times. I could not have accomplished what I have if it had not been for her love and support.

TABLE OF CONTENTS

Page

Page

Abstract

The take-off point for this dissertation is the body of literature in finance and control theory which involves stock-trading strategies based on technical analysis. A salient feature of the line of research which we pursue is that neither a parameterized model for stock prices nor a behavioral model involving "agents" is used. This class of trading methods is said to be *model free*. Many papers in finance documenting the "efficacy" of such model-free trading methods are based on backtesting using historical price data. This reliance on data in lieu of a formal theory explaining successes and failures is one of the main reasons that many in the finance community have criticized this method of trading. In addition, many of these strategies are heuristic in nature which makes them difficult to carry out mathematical analysis. In direct contrast to the approaches above, the main objective of this dissertation is to further the development of a relatively new line of research emerging from control community: using simple ideas involving robust and adaptive control to provide a theory explaining successes and failures of various classes of technically-based trading strategies. In a sense, we seek to "demystify" model-free technical analysis.

Analysis here is carried out under the assumption of an "idealized market" which is similar to the well-known concept of a "frictionless market" in finance. In our setting, the feedback controller which determines the investment level is parameterized by a gain denoted by K and its gain-loss performance is benchmarked using Geometric Brownian Motion (GBM), the most famous price-generating process in the finance literature. In this GBM setting, one of our first results is a new formula for the skewness of the probability distribution of the trading gains which is seen to be an increasing function of the feedback gain K. The second result in the thesis is motivated by the fact that a highly-skewed distribution often leads to a significant probability of loss and large "drawdown," a well-known measure of risk. With this is mind, we derive a formula for the drawdown in the wealth when the strategy is again based on a linear feedback trading strategy.

Following the initial results above, we study a so-called Proportional-Integral (PI) controller, a generalization of our linear feedback controller to exploit memory. One of the main results is that in an idealized market with stock price governed by a non-trivial GBM, a combination of two PI controllers leads to a positive expected value for the trading gain. Since this holds independently of the parameters underlying the GBM process, it is called the Robust Positive Expectation Property. The theory is then extended to accommodate a variation of this PI controller which involves an exponentially-weighting scheme to put more emphasize on recent performance and reduce the impact of “old information” on the investment level.

While the results described above are for feedback-based trading in continuous time, we also consider the discrete-time case which is apropos for a “low-frequency” trader such as a typical small investor. In this setting, we consider a trading rule which involves a *controller with delay* and, analogous to PI controller, is motivated by a desire to include weighting of recent performance to obtain the investment level. Once introduced, it is proven that the Robust Positive Expectation Property holds for this delay system.

The final results in the thesis are motivated by skewness considerations and are considered to be an “off-shoot” of our research. In the presence of skew, classical mean-variance based analysis can provide a distorted picture of the prospect for success. This can become even more crucial in “mission-critical” applications with the possibility of “model distrust.” With these motivations in place, we introduce a new “conservative” reward-risk pair which not only discounts long tail of distribution but is also independent of individual’s risk-preference; i.e., utility function. To this end, the *Conservative Expected Value* (CEV) and *Conservative Semi-Variance* (CSV) are formally defined. They are calculated for some of famous probability distribution and some of their most important properties are established.

Chapter 1

Background and Overview of this Dissertation

The take-off point for this dissertation is a new line of research aimed at studying *model-free* stock trading strategies based on concepts from control theory; e.g., see [1–14]. Motivation for our study of this problem area is derived in part from the limited ability of existing models to predict prices. The failure of these models is epitomized in the turbulent markets experienced in 2000 and 2008-2009. For example, the model used by Markowitz [15] relies on covariance considerations in order to bring diversification to the portfolio. However, this model failed dramatically in volatile days of 2008 and 2009 as price movements became excessively correlated. Given this context, the so-called model-free trading strategies which we consider in this dissertation and their associated buy and sell signals are based on gain-loss performance rather than any type of parameterized stock-price model or agent-based model. These trading rules fall under the umbrella of "technical analysis," for example, see [16–21].

In addition to the class of *model-free* strategies defining the scope of this thesis, there are other "flavors" of technical analysis in the literature. For example, one method of technical analysis includes consideration of agents, rational expectations, equilibria and information content in the stock prices; e.g., see [22] and [23]. Yet a third type of technical analysis, some of which is inspired by control theory, focuses on trading strategies aimed at maximization of the trader's "utility function" or using a known model for dynamics; e.g., see [24–52]. The fact that model parameters are assumed to be known is what differentiates this work from ours. Finally, we note that many of these cited references are only tangentially related to the research described in this thesis in that they deal with issues in financial markets other than the narrow focus here: trading a stock.

In the finance literature dedicated to the analysis of technically-based trading strategies, the use of backtesting with historical price data has been the method of choice. Reference [53] provides an excellent survey and bibliography covering this literature; see also [21, 54–59] for more detail. While technical analysis is widely used by practitioners, many of the statistically-based performance analyses in finance literature claiming "excess" profits have been challenged. Critics claim that these strategies have not demonstrated a significant "edge" over market benchmarks; e.g., see [58, 60–62]. Furthermore, many academics tend to be skeptical about the efficacy of technical analysis based on the belief that stock prices are unpredictable. As pointed out in [21], some authors have gone so far as to refer to technical analysis as "voodoo finance." In [63], technical analysis is referred to as "anathema." One of main reasons underlying such skepticism is the "Efficient Market Hypothesis," originally described in [35, 64–66], which indicates that there is no "pattern" in asset prices which can be exploited in order to make excess profit.

In response to the type of criticism above, there is a large body of literature suggesting that technical analysis can be quite successful; e.g., see [19, 20, 54, 55, 67–74]. For example, in [54], the case for the efficacy of technical analysis is made via a number of significant empirical studies involving the use of statistics and historical data under a number of market conditions; e.g., see [21] and [56]. In addition, it should also be noted that the use of technical analysis is quite popular in many quarters of Wall Street. For example, nowadays, hedge funds as well as individual investors often document their performance; e.g., see [75] and [76]. Finally, it should be mentioned that some of this literature on the empirical performance of technically-based strategies has been challenged on the grounds that "data snooping" is suspected. For example, it is argued that the use of a small data set exaggerates the performance of a strategy; e.g., see [77–80].

It is pointed out in [18] that the heuristic nature of many technically-based strategies makes it difficult to carry out mathematical analysis. In this regard, in a number of publications such as [18] and [24], the possibility of a formal theoretical framework to analyze the performance of such strategies has been raised. This dissertation concentrates on developing such a theoretical framework. Beginning with the initial results from the control community in [1–5], we pursue theoretical explanations for the performance of various *model-free* technically-based trading algorithms. That is, instead of studying the efficacy of a trading strategy via statistical methods or use of a price model, we develop a formal model-free theory based on feedback control considerations.

Given the context above, the main goal of this dissertation is to develop a proof-based theoretical framework addressing some special cases of technical analysis — the hope being to open the door to demystification of more general situations. The new strategies which we consider involve linear feedback and belong to a class of trend-following strategies; e.g., see [26, 81–83]. In this setting, our controllers adapt the investment level in response to gains and losses as they evolve over time; e.g., see [1–5]. It should be noted that the main point in this dissertation is not to develop new "market-beating" algorithms. It is to develop a feedback-based theory which explains both successes and failures of new and existing strategies.

In the remainder of this chapter, we provide an overview of the class of linear feedback strategies being considered and some of the main dissertation results are previewed. The analysis is carried out under the assumption of an "idealized market," described in Section 1.2. Such a market is closely related to the so-called "frictionless market" described in [84]. In our framework, for the continuous-time case, the performance is benchmarked by driving the trading system using prices obtained as sample paths of a Geometric Brownian Motion (GBM) process with drift μ and volatility σ which are unknown to the trader. This classical process is the most famous *benchmark* in the finance literature and is used to test different strategies; e.g., see [85–89]. For the case of discrete-time benchmarking, it is assumed that the returns $\rho(k)$ are independent and have common mean $\mathbb{E}[\rho(k)] = \mu$. It is important to note that the model-free controllers which we consider use no apriori information about the underlying parameters of the benchmark price process. Most notably, we seek performance certification theorems which are robust with respect to sign of the μ; i.e., performance is guaranteed even when the trend of the price is unknown. That is, we expect the controller to perform well in both upward and downward trending markets.

1.1 Introduction to Trading via Model-Free Feedback Control

As stated above, in this work, no model for the price or the trader's preference is used and no parameter estimation is involved. Instead, the stock price $p(t)$ is treated as an external uncontrolled input and robust performance is sought. The stock-trading strategies which we consider in this category use "patterns" in prices or the resulting gain-loss function $g(t)$ to generate buy and sell signals. The block diagram in Figure 1.1.1 shows a controller which uses the gain-loss performance to modify the investment level.

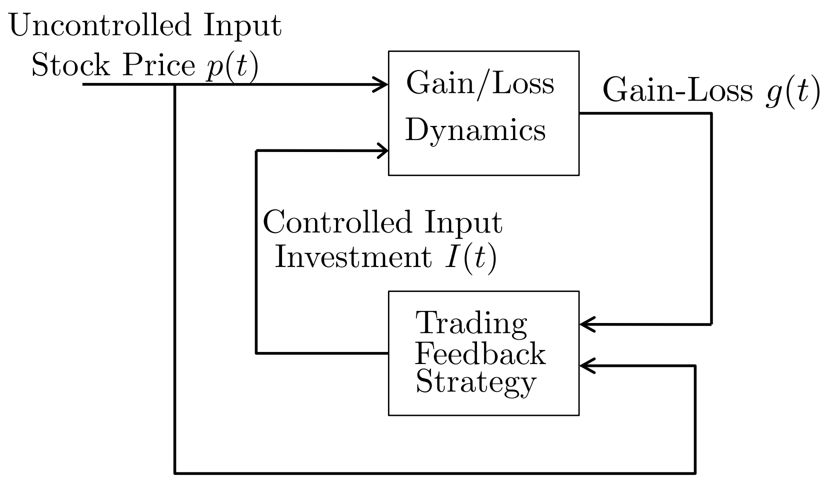

Figure 1.1.1: Block Diagram of Model-Free Trading via Feedback Control

1.1.1 Linear Feedback: To provide a first example of a model-free strategy, we consider a trader who modifies the amount invested via a linear feedback rule. More specifically, the instantaneous investment $I(t)$ is modulated in continuous time based on the cumulative trading gain or loss $g(t)$ over $[0, t]$, where $g(t) < 0$ is understood to be loss. This continuous-time modulation of the investment is consistent with the first assumption regarding "idealized markets" to be covered in the next section. The classical linear time-invariant feedback trading rule we consider is given by

$$I(t) = I_0 + Kg(t)$$

where I_0 is the initial investment and K is the feedback gain. In the formulation above, with $K > 0$, the variation of the investment $I(t)$ is proportional to the variation of the trading gain or loss $g(t)$; that is, the investment follows the "trend" associated with the $g(t)$; i.e., it is a trend-following strategy of sorts.

When $I(t) > 0$, the investment is said to be "long" and when $I(t) < 0$, it is called "short." A trader with long investment owns shares hoping for the price to rise in order to obtain a profit. In case of a short investment, the trader borrows shares from the broker and sells them in the open market in the hope of a price drop. At any time, the trader can buy back the shares and return them to the broker with either a realized gain or loss. The block diagram of the resulting linear feedback system is shown in Figure 1.1.2.

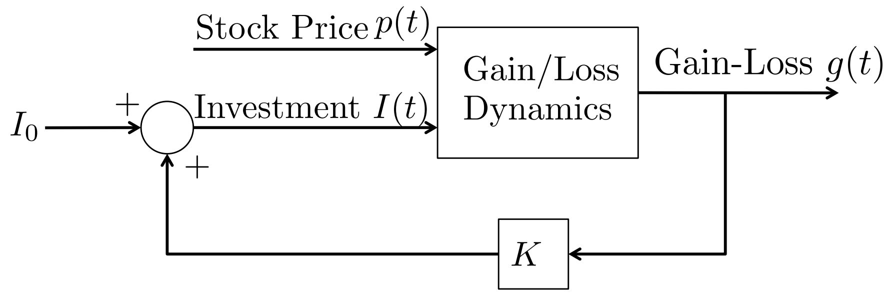

Figure 1.1.2: Block Diagram of Trading via Linear Feedback Control

1.2 The Notion of an Idealized Market

An *idealized market* is defined via a set of assumptions about the trading process. We describe these assumptions below and indicate how they relate to a real market. Our view in this dissertation is that proof-based analysis of a trading strategy in this idealized market accompanied by significant backtesting against historical prices provides an indication of the prospect of success or failure. As previously mentioned, the idealized market is similar to the so-called "frictionless market" in the finance literature; e.g. see [84]. The notion of an idealized market is described as follows:

- *Continuous Trading:* It is assumed that the trader can react instantaneously to price changes and continuously adjust the investment. This also allows the trader to hold a fractional number of shares. The assumption of continuous trading is consistent with high-frequency trading scenarios involving many transactions over short time intervals. This assumption is classical in finance; e.g., see [34, 35] and is also used in the well-known Black-Scholes model; see [89].

- *Perfect Liquidity and Price-Taker Assumption:* It is assumed that the trader can transact as many shares as desired, at the instantaneous price $p(t)$. In practical terms, this means that the trader faces no "gap" between the bid and ask prices. It also means that the trader is a price taker in the sense that the stock price remains constant at $p(t)$ during the course of the transaction. For example, this would be the case if the trader, at time t, is not buying or selling "sufficiently large" blocks of stock so as to have an influence on the price. Note that this assumption would be faulty in the case of a "large" hedge or mutual fund buying or selling millions of shares per day. For example, if this large trader is a buyer, the price typically increases during the course of the transaction. As the demand-supply gap grows, the "last" of these purchased shares becomes more costly to acquire than the earlier shares.

- *Costless Trading:* It is assumed that the trader incurs no transaction costs such as brokerage commissions, margin costs or exchange fees. In practice, these costs are often negligible for high-volume traders such as investment companies or hedge funds. Even for the smaller trader, transaction costs have droppd dramatically in recent years.

- *Adequate Resources:* It is assumed that the trader has sufficient resources so that no transaction is "stopped" or no "liquidation" occurs due to failure to satisfy the broker's collateral requirement. For example, this assumption is satisfied if the trader's account has a large cash reserve or if the securities in the account, not bought on margin, provide adequate collateral.

1.3 Price Benchmarks

When it comes to evaluating a trading strategy, one obvious way to start is to work with theoretical benchmarks and historical data. In this regard, as previously mentioned, one of the most well-known price processes, which is widely used in the finance literature is Geometric Brownian Motion (GBM); e.g., see [88]. The GBM process is described by the stochastic equation

$$\frac{dp}{p} = \mu dt + \sigma dZ,$$

where dp/p is the percentage change in price over the time increment $[t, t+dt]$, μ is the *drift*, $\sigma \geq 0$ is the *volatility* of the process and $Z(t)$ is a standard Wiener process. The increment dZ can be viewed as a zero-mean normal random variable with variance dt; that is, $dZ \sim \mathcal{N}(0, dt)$.

1.3.1 Discrete-Time Benchmark: In this thesis, we also consider feedback-based trading rules in a discrete-time setting. The classes of stock-price process $p(k)$ which we consider for these strategies are described by their returns

$$\rho(k) \doteq \frac{p(k+1) - p(k)}{p(k)}; \quad k = 0, 1, 2, \cdots$$

which are assumed to be independent with common mean

$$\mathbb{E}[\rho(k)] = \mu.$$

1.4 Existing Results Motivating Dissertation Research

In this section, we describe some existing results which provide motivation for this dissertation. First, we consider the simple linear feedback trading strategy which was introduced in Subsection 1.1.1 and give formulae for the probability density function, the mean and the variance of the associated gain-loss $g(t)$, with the price driven by a GBM process. We similarly review results in the literature for a modification of the linear feedback strategy which is called *Simultaneous Long-Short* (SLS). This strategy, first introduced in [1], is aimed at achieving robustness with respect to unknown stock price process parameters. In this setting, the model-free trader is agnostic about the direction of stock price movement and volatility. This controller is introduced in Subsection 1.4.3. Subsequently, in Theorem 1.4.4, we give formulae for the mean and variance of the gain-loss function $g(t)$ resulting from the SLS feedback control when the price is driven GBM process.

The description of existing results given in this section is rather detailed. There are two main reasons for this *detailed* exposition: First, to illustrate the type of theoretical proof-based framework we want to develop; i.e., to demonstrate what we mean by "performance certification theorems" for feedback-based trading rules. The second reason is that these results were the take-off point of this thesis leading to some of our initial research. More specifically, Theorem 1.4.2 was the starting point of the research reported in Chapter 2. Also, the existing results described in Subsection 1.4.3 and Theorem 1.4.4 led to a more generalized theory we developed in Chapter 4.

1.4.1 Simple Linear Feedback Case: In this subsection, the formulae for the expected value, the variance, and the probability density function of the trading gain-loss function $g(t)$ are given for the case when the linear feedback is employed. These results are derived under the assumption of an idealized market with prices driven by a GBM process. It is important to note that the model-free framework of this thesis dictates that the trader does not have apriori knowledge of the GBM parameters μ and σ. That is, the feedback gain $K \geq 0$ is not a function of μ and σ. The following theorem is a consequence of the analysis in [4].

1.4.2 Theorem: *Consider an idealized market with prices driven by a Geometric Brownian Motion process with drift μ and volatility σ. With the long linear feedback investment controller $I(t) \doteq I_0 + Kg(t)$, with $I_0 > 0$ and $K \geq 0$, used to determine the investment level, at time $t > 0$, the resulting gain-loss function $g(t)$ has the lognormal probability density function*

$$f_g(x,t) = \frac{1}{\sigma\sqrt{2\pi t}(I_0 + Kx)} \times e^{-\frac{\left(\log(1+\frac{Kx}{I_0})+0.5K^2\sigma^2 t-\mu K t\right)^2}{2K^2\sigma^2 t}}$$

for $x > -I_0/K$. Furthermore, the mean and variance of $g(t)$ are given by

$$\mathbb{E}\big(g(t)\big) = \frac{I_0}{K}\left[e^{\mu K t} - 1\right]; \quad \text{var}\big(g(t)\big) = \frac{I_0^2}{K^2}\, e^{2\mu K t}\left[e^{K^2\sigma^2 t} - 1\right].$$

1.4.3 Simultaneous Long-Short Feedback Control Case: The ideas which we describe below are generalized in Chapter 4 to handle PI controller and considered in the discrete-time in the context of a controller with delay in Chapter 5. Indeed, the trader with long investment, following the linear feedback strategy as described in the theorem above, typically makes profit when the price rises. However, if the price falls, a loss is likely to occur. Motivated by a desire to hedge against market declines, another strategy based on linear feedback control, introduced in [1], is called *Simultaneous Long-Short* (SLS). The objective for a trader with no model for the price is to benefit from price trends in either direction. This is accomplished by combining two linear feedback controllers, one dedicated to a long trade and the other dedicated to a short trade. That is, with $I_0 > 0$ and $K > 0$, the amount to invested in the long trade is given by

$$I_L(t) \doteq I_0 + Kg_L(t)$$

where $g_L(t)$ is the gain-loss function corresponding to this long trade. Similarly, we use the subscript "S" to denote the *short* position; i.e., for $I_0 > 0$ and $K > 0$, the short trade is given by

$$I_S(t) \doteq -I_0 - Kg_S(t)$$

where $g_S(t)$ is the gain-loss function associated with this short trade. As the gain $g_S(t)$ goes up, the short investment $I_S(t)$ increases in magnitude; i.e., $I_S(t)$ becomes more negative. The trader following the SLS strategy has the investment $I_L(t)$ long and $I_S(t)$ short with the overall investment given by

$$I(t) = I_L(t) + I_S(t) = K\big(g_L(t) - g_S(t)\big),$$

and the overall gain-loss function is

$$g(t) = g_L(t) + g_S(t).$$

In practice, a broker typically implements the Simultaneous Long-Short trade by "netting out" the two investments $I_L(t)$ and $I_S(t)$ rather than carrying two separate positions. At time $t = 0$, the trade is "flat" since $g_L(0) = g_S(0) = 0$ implies that $I(0) = 0$. Then for $t > 0$, the controller adapts to gains and losses. For example, if the long trading gain $g_L(t)$ is trending upward, the feedback rules amplify $I_L(t)$ and attenuate $I_S(t)$. The exact opposite will occur in a declining market. That is, the short side investment $I_S(t)$ will increase in magnitude becoming more negative and the long side $I_L(t)$ decreases. Some of the main results in the literature for this strategy are provided below.

The theorem to follow, established in [4], provides formulae for the expected value, the variance of gain-loss function $g(t)$ under the assumption of an idealized market with the price driven by a Geometric Brownian Motion.

1.4.4 Theorem: (Robust Positive Expectation) *Consider an idealized market with price driven by the Geometric Brownian Motion with the drift μ and volatility σ. Then, for $t \geq 0$, the expectation and variance of the gain-loss function $g(t)$ resulting from the SLS feedback control are given by*

$$\begin{aligned} \mathbb{E}[g(t)] &= \frac{I_0}{K}[e^{\mu K t} + e^{-\mu K t} - 2]; \\ \mathrm{var}[g(t)] &= \frac{I_0^2}{K^2}\left(e^{\sigma^2 K^2 t} - 1\right)\left(e^{2\mu K t} + e^{-2\mu K t} + e^{\sigma^2 K^2 t}\right). \end{aligned}$$

Furthermore, for the non-trivial case $\mu \neq 0$, we have

$$\mathbb{E}[g(t)] > 0.$$

1.4.5 Remarks: The fact that $\mathbb{E}[g(t)]$ is positive and independent of the sign of μ justifies the name Robust Positive Expectation Property above. Establishing this property for the feedback-based controllers in this thesis will be one of the objectives in the chapters to follow. To conclude this survey, we provide a sample of a more technical result demonstrating the depth of technical analysis which is possible using the methods in this thesis. The notation and to follow can be skipped by the less technically-minded reader without loss of continuity in the reading of this dissertation. The theorem to follow, established in [1] and [4], provides a closed-form for the probability density function of the gain-loss function $g(t)$ resulting from the use of SLS control driven by Geometric Brownian Motion as described in the preceding subsections. Indeed, for the SLS controller with initial investment I_0 and feedback gain K, using the GBM drift μ and volatility σ, at time $t > 0$, we introduce the notation

$$X_{\pm}(x,t) \doteq \frac{1}{2}e^{\frac{1}{2}\sigma^2(K-K^2)t}\left[\Big(\frac{K}{I_0}x+2\Big) \pm \sqrt{\Big(\frac{K}{I_0}x+2\Big)^2 - 4e^{-\sigma^2K^2t}}\right];$$

$$\nu \doteq \mu - \frac{\sigma^2}{2};\;\; Z_{\pm}(x,t) \doteq \frac{\log X_{\pm}^{\frac{1}{K}}(x,t) - \nu t}{\sigma\sqrt{t}};$$

$$A(x,t) \doteq \frac{1}{\sigma I_0\sqrt{2\pi t}\sqrt{(\frac{K}{I_0}x+2)^2 - 4e^{-\sigma^2K^2t}}}.$$

Using the notation above, the probability density function for $g(t)$ is provided in the theorem below.

1.4.6 Theorem: *Consider an idealized market with price driven by Geometric Brownian Motion with the drift μ and volatility σ. Then for $t > 0$, the probability density function $f_g(x,t)$ for the gain-loss function $g(t)$ when the Simultaneous Long-Short strategy is used, is given by*

$$f_g(x,t) = A(x,t)\big(e^{-\frac{1}{2}Z_+^2(x,t)} + e^{-\frac{1}{2}Z_-^2(x,t)}\big)$$

for $x \geq g^*(t)$ where

$$g^*(t) \doteq \frac{2I_0}{K}\big[e^{-\frac{1}{2}\sigma^2K^2t} - 1\big].$$

1.4.7 Example: The plot of the probability density function for the gain-loss function $g(t)$ of SLS for various values of the drift μ is provided in Figure 1.4.1. The other parameters used are $\sigma = 0.2$, $K = 4$, $t = 0.5$ and $I_0 = 1$.

1.5 Dissertation Results at a Glance

In this section, a brief overview of the dissertation results is provided with the details relegated to later chapters. In Subsection 1.5.1, the formula for the skewness of the probability density function of trading gain-loss function $g(t)$ given in Theorem 1.4.2 is obtained when a linear feedback strategy is employed. As a consequence, we see how a highly-skewed distribution can result in a large

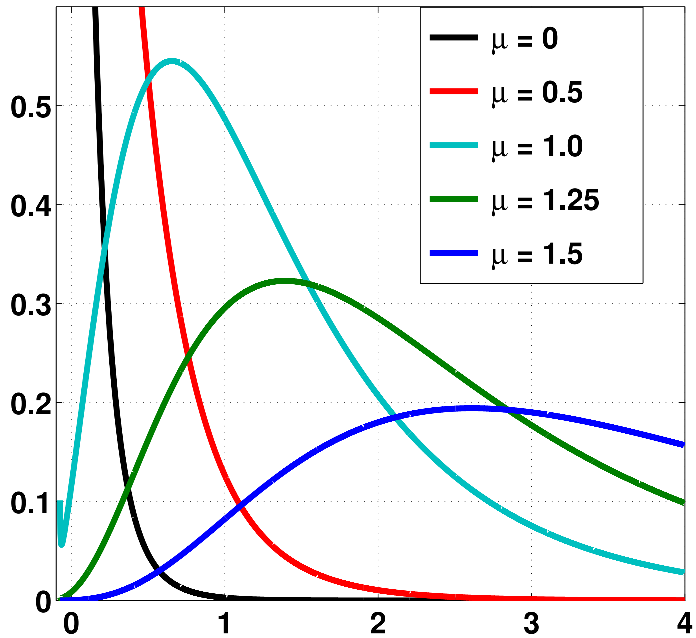

Figure 1.4.1: Probability Density Function of $g(0.5)$ for Various Values of μ

drawdown in the trader's wealth. A formula for drawdown, a well-known and widely-used measure of risk, is given in Subsection 1.5.2. In Subsection 1.5.3, a modification of SLS controller is proposed to include *memory* in the investment rule. This new controller, based on a *Proportional-Integral (PI)* controller, is shown in Chapter 4 to have the Robust Positive Expectation Property. This sort of result serves as an illustrative case for the type of analysis which is possible in our control-theoretic setting.

In contrast to the results given for feedback-based trading in continuous time, Subsection 1.5.4 considers the discrete-time case. For a "low-frequency" trader such as a typical small investor, discrete-time results are more realistic. In this regard, to demonstrate the style of analysis in this setting, a new linear feedback strategy is introduced. This trading rule involves a *controller with delay* and is motivated by a desire to include weighting of recent performance to obtain the investment level. Once introduced, in Chapter 5, it is proven that the Robust Positive Expectation Property holds for this delay system.

Finally, Subsection 1.5.5 overviews an interesting "offshoot" of our research involving a new risk and reward pair, the *Conservative Expected Value* (CEV) and *Conservative Semi-Variance* (CSV) respectively. Much of the motivation for the development of this pair comes from the fact that use of possibly-large feedback gains K often leads to a fat-tailed, highly-skewed probability distribution for the gain-loss function $g(t)$. In this case, the analysis of a trade based on classical expected value $\mathbb{E}[g(t)]$ can become unduly optimistic. Furthermore, in such cases, classical risk analysis involving variance may further distort one's overview of the risk at hand. That is, the variance "penalizes" large positive profits which are desirable. In this regard, in finance, it is standard to use a *semi-variance measure*, for example, see [90] and [91], and this motivates the definition of CSV. As illustrated by examples in Chapter 6, we envision the (CEV,CSV) theory to be applicable in different areas. Accordingly, the corresponding exposition in Chapter 6 is more general than the "finance-flavored" focus of Chapters 1-5. These metrics are shown to have certain "promising" properties and they are calculated for some of important, widely-used probability distributions.

1.5.1 Formula for Skewness and Efficiency Considerations: The first thesis contribution, given in Chapter 2, is a formula for the skewness of the distribution of the gain-loss function $g(t)$, when the linear feedback rule

$$I(t) = I_0 + Kg(t)$$

is used and prices are driven by Geometric Brownian Motion. At any pre-specified time t, we prove that skewness is given by the formula

$$S(K) = \left(e^{K^2\sigma^2 t} + 2\right)\sqrt{e^{K^2\sigma^2 t} - 1}.$$

Further, a similar approach is used to derive the formula of the skewness for the trading gain when the Simultaneous Long-Short trading strategy, described in Section 1.4.3, is employed. It is also shown that $S(K)$ is a monotonically increasing function of the feedback gain K. This provides a "warning" regarding the use of classical mean-variance analysis. That is, mean-variance based measures of performance may be entirely inappropriate when a feedback control law is used.

1.5.2 Analysis of Drawdown: Depending on the size of feedback gain $K > 0$, the feedback-induced probability distributions can have large right-sided skewness. This means that even with a large positive expected gain, $\mathbb{E}[g(t)]$, the probability of loss can become significantly large and a large drawdown may result. Attention to drawdown in wealth is one of the most important aspects of risk management; e.g., see [56, 92–94]. For such situations and along the lines described earlier in Subsection 1.5.1, a classical mean-variance analysis as in the celebrated work of Markowitz [15] often does not suffice since higher order moments are in play. In this regard, measures of drawdown have received considerable attention over recent years. The important results on stock price drawdown in [92] and [95] strongly motivate our analysis in Chapter 3.

More formally drawdown is defined both in absolute and percentage terms. More specifically, suppose $V(t)$ represents a trader's wealth (account value) for $0 \leq t \leq T$. Then the *maximum absolute drawdown* is defined as

$$D_{max}(V) \doteq \max_{0 \leq s \leq t \leq T} V(s) - V(t)$$

and *maximum percentage drawdown* is defined as

$$d_{max}(V) \doteq \max_{0 \leq s \leq t \leq T} \frac{V(s) - V(t)}{V(s)}.$$

In the Figure 1.5.1 these two quantities are shown for $T = 10$.

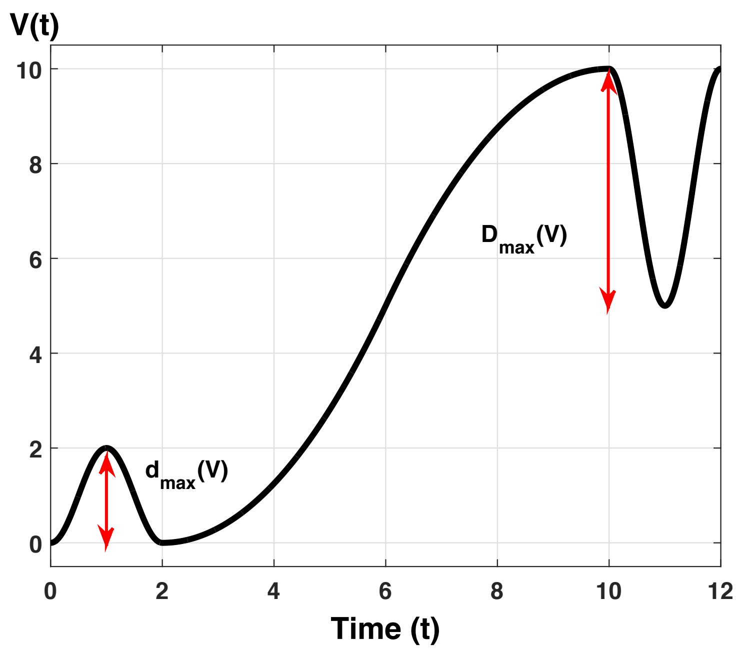

Figure 1.5.1: Maximum Absolute and Percentage Drawdown

With the price driven by Geometric Brownian Motion, we studied the random variable corresponding to the maximum percentage drawdown in wealth. In Chapter 3, we provide an upper bound on the expected value of d_{max} when linear feedback is used. Depending on the size of the feedback gain K, it is seen that this upper bound has three different possibilities for its asymptotic behavior as T gets large. In this regard, when the prices are driven by Geometric Brownian Motion, the signal-to-noise ratio μ/σ is seen to play an important role.

1.5.3 Generalization to Dynamic Controller: In Chapter 4, results on the Robust Positive Expectation Property are generalized to include a dynamic controller. To this end, we consider a classical Proportional-Integral (PI) controller

$$I(t) = I_0 + K_P g(t) + K_I \int_0^t g(\tau) d\tau.$$

The inclusion of the integral above means that "memory" of past performance is used to determine the instantaneous investment; see Figure 1.5.2 below.

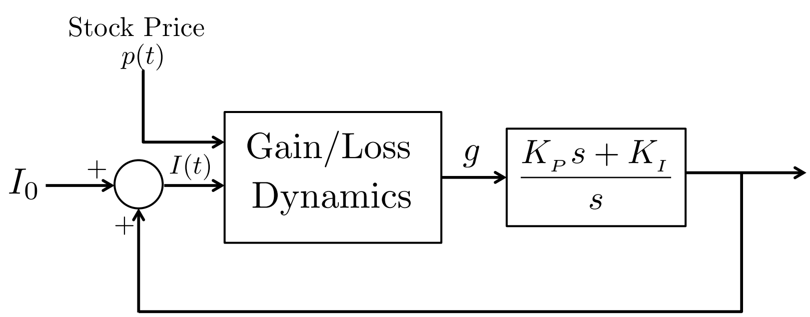

Figure 1.5.2: Block Diagram of Trading via PI Controller

In this setting, we define a long-short version of the PI controller above. Indeed, with subscripts "L" and "S" denoting the Long and Short components respectively, the investment components are defined by

$$I_L(t) \doteq I_0 + K_P g_L(t) + K_I \int_0^t g_L(\tau) d\tau;$$
$$I_S(t) \doteq -I_0 - K_P g_S(t) - K_I \int_0^t g_S(\tau) d\tau$$

and we consider the case when a Geometric Brownian Motion (GBM) drives the stock price. More specifically, for the non-trivial case of $(K_P, K_I) \neq (0,0)$ and the price drift $\mu \neq 0$, at any time $t > 0$, the main result in Chapter 4 is that

$$\mathbb{E}[g(t)] > 0.$$

The chapter also considers some practical considerations related to collateral requirements of the broker. The results are generalized to an SLS version of the *exponentially weighted moving average* controller

$$I(t) = I_0 + K_P g(t) + K_I \int_0^t e^{-\gamma(t-\tau)} g(\tau) d\tau$$

where $\gamma \geq 0$ is chosen by the trader.

1.5.4 Discrete-Time Controller With Delay: In Chapters 2-4, the trader modifies the investment level continuously in time. Despite the development of high-frequency trading platforms in recent years, the assumption of trading continuously in time can be impractical, especially for a typical small trader. Accordingly, Chapter 5 addresses discrete time while noting that the results also apply to high-frequency trading as the discretization interval Δt becomes very small. In this setting, we demonstrate the type of analysis which is possible in our control-theoretic framework by considering an SLS linear feedback *controller with delay* and prove that a discrete-time version of the Robust Positive Expectation Property holds.

At stage k, this new controller focuses on recent performance via inclusion of the term $g(k-m)$ where m is some pre-specified look-back period. More specifically, at stage k, the long and short components of the investment level are given by

$$\begin{aligned} I_L(k) &\doteq I_0 + K\big(g_L(k) - g_L(k-m)\big); \\ I_S(k) &\doteq -I_0 - K\big(g_S(k) - g_S(k-m)\big), \end{aligned}$$

with the initial investment $I_0 > 0$ and $K \geq 0$ being the feedback gain. Notice that the investment levels $I_L(k)$ and $I_S(k)$ are attenuated as the "winning power" goes away and gets rekindled if the trade reverts to profitability.

In this setting, we consider the discrete-time price benchmark described in Subsection 1.3.1. More specifically, we show that except for the trivial break-even case when either $K\mu = 0$ or $N = 1$, the Robust Positive Expectation Property holds; that is, at any step $N > 1$,

$$\mathbb{E}\big[g(N)\big] > 0.$$

1.5.5 New Conservative Reward-Risk Pair: Recalling the discussion about highly-skewed probability distributions in Section 1.5.1, the takeoff point for our CEV-CSV theory is a random variable X for which larger values are preferred and the corresponding probability distribution is highly skewed. For example, in the thesis framework, X represents the trading gain-loss function $g(t)$ resulting from the use of linear feedback in which the possibility of a long fat-tailed distribution can lead to an expected value, $\mu = \mathbb{E}[X]$ which is unduly optimistic. This motivates the definition of the *Conservative Expected Value* (CEV) and the associated *Conservative Semi-Variance* (CSV).

To provide a quick overview of the CEV, we take the leftmost support-point of the cumulative distribution function $F_X(x)$ for random variable X, denoted by α_X, to be known and finite. This assumption on the "worst-case" outcome is realistic in many applications: For example, the worst-case loss of a linear feedback trading rule when the price is driven by Geometric Brownian Motion is given in Theorem 1.4.6; that is $\alpha_X = g^*$. Additional examples involve the price of a stock or the lifetime of a component in a system which are both non-negative random variables; that is $\alpha_X = 0$.

To motivate the definition below, we imagine a mission-critical decision being made by a risk-averse individual who is rewarded or penalized based on X. If this person has a "minimal acceptable target" for X, denoted by γ, a conservative approach to the analysis of this gamble X would be to shift the probability mass associated with $X \leq \gamma$ to α_X, and to shift the probability mass associated with $X > \gamma$ to γ; see Chapter 6 for elaboration. This procedure maps the original random variable to a Bernoulli random variable X_γ described by

$$X_\gamma \doteq \begin{cases} \alpha_X & \text{with probability} \quad P(X \leq \gamma) = F_X(\gamma); \\ \gamma & \text{with probability} \quad P(X > \gamma) = 1 - F_X(\gamma). \end{cases}$$

After this conservative procedure is carried out, one is left with a simple Bernoulli random variable X_γ with expected value

$$\mathbb{E}[X_\gamma] = \alpha_X F_X(\gamma) + \gamma\big(1 - F_X(\gamma)\big)$$

and the *Conservative Expected Value (CEV)* is defined to be

$$\mathbb{CEV}(X) \doteq \sup_\gamma \mathbb{E}[X_\gamma].$$

Consequently, motivated by downside risk considerations, the associated *Conservative Semi-Variance (CSV)* is also defined and we proceed to demonstrate the mathematical breath and applicability of our new theory. That is, for many classical distributions, the (CEV,CSV) pair is compared with their (μ, σ^2) counterparts and many mathematical properties are established. We also provide some numeric examples using real-world data to show the potential use in a number of applications.

1.5.6 Concluding Remarks: In Chapter 7, we provide concluding remarks and directions for future research. In particular, we introduce a new discrete-time feedback-based trading rule which is *triggered by moving average crossing,* discuss the problem of choosing the controller parameters, describe three new strategies and outline a research path to extend the analysis in this dissertation to trading a portfolio of stocks.

Chapter 2

On Skewness of the Trading Gain

In this chapter, we derive a formula for the skewness of gain-loss function $g(t)$, when linear feedback is used to determine the level of investment.[1] In classical finance, when a stochastic investment outcome is characterized in terms of its mean and variance, it is often implicitly taken for granted that the underlying probability distribution is not heavily skewed. For example, in the "perfect" case when outcomes are normally distributed, mean-variance considerations tell the entire story. The main point of this chapter is that mean-variance based measures of performance may be entirely inappropriate when a feedback control law, as described in Subsection 1.5.1, is used instead of buy-and-hold to modulate one's stock position as a function of time. For example, with Geometric Brownian Motion generating prices, when using a feedback gain K to increment or decrement one's stock position, we see that the resulting skewness measure $S(K)$ for the trading gains or losses can easily become dangerously large.

In view of the above, we argue in this chapter that the selection of this gain K based on a classical mean-variance based utility function can lead to a distorted picture of the prospects for success. To this end, our analysis begins in an idealized market with prices generated by Geometric Brownian Motion; see subsections 1.2 and 1.3 respectively. In addition to the "red flag" associated with skewness, a controller efficiency analysis is also brought to bear. While all feedback gains K lead to efficient (non-dominated, Pareto optimal) controllers, we show that the same does not hold true when we use a return-risk pair which incorporates more information about the probability distribution for profits and losses. To study the efficiency issue in an application context, the chapter also includes a simulation Pepsico Inc. using the last five years of historical data.

Motivation for our work in this chapter is derived by first considering the classical buy-and-hold strategy in lieu of linear feedback above. In this case, the trading gain or loss at the terminal

[1] The results reported in this chapter have been published in [7].

time $g(T)$ is governed entirely by the log-normally distributed stock price with skewness given by

$$S = (e^{\sigma^2 T} + 2)\sqrt{e^{\sigma^2 T} - 1}.$$

In practice, in financial markets, it is often the case that this leads to S-values which are sufficiently small to justify reliance on mean-variance performance evaluation; for example, see [96]. Even when a stock trades with annualized volatility as high as fifty percent, that is, $\sigma = 0.5$, the resulting skewness after one year, let $T = 1$, is $S \approx 1.75$, which is rather modest.

2.0.1 Classical Considerations: Many of the well-known tools in investment theory are based on the statistics of the return such as the mean and variance; e.g., see [22,97,98]. Once such statistics are available, various strategy options can be assessed using a number of different measures. More generally, it is well-known that mean-variance considerations can be inadequate for various classes of problems in finance; e.g., see [99] and more recent work such as [100] accounting for higher order moments. Further to illustrate, the main point of this chapter is that mean-variance based measures of performance may be entirely inappropriate when a feedback control law is used instead of buy-and-hold to modulate one's stock position as a function of time. For example, when using a feedback gain K to increment or decrement one's stock position, we see that the resulting skewness measure $S(K)$ for the trading gains or losses can easily become dangerously large. In fact, we see in Section 2.1 that the skewness formula above, when modified by linear feedback, becomes

$$S(K) = (e^{K^2\sigma^2 T} + 2)\sqrt{e^{K^2\sigma^2 T} - 1}.$$

This formula clearly shows that when K increases, the skew can get very large. For example, with modest annualized volatility of 5% and feedback gain $K = 10$, with $\sigma = \sqrt{0.05}$ and $T = 0.4$ above, we obtain $S(10) \approx 23.73$. In other words, it is arguable that the information contained in the higher order moments cannot be neglected. To reiterate our motivation underlying this type of demonstration, our hypothesis is the following: Given the high-frequency adjustment of a stock position, the possibility of highly skewed probability density function for $g(T)$ can result from linear feedback. While this type of hypothesis is self-evident to some degree, it is shown here by examples that this skewing effect is actually realizable rather than just a mathematical possibility.

2.0.2 Efficiency Considerations: The second issue which we consider, related to the first, is the issue of "efficiency" which is central to modern finance; e.g., see [15] and [101]. First, we argue that no matter what linear feedback gain K is used, the resulting risk-return pair $(\text{var}(g(T)), \mathbb{E}[g(T)])$ is always efficient. In terms of the language used in the control literature, this can be understood to mean that this pair is non-dominated or Pareto optimal. Our point of view is that all feedback gains being deemed efficient is a distortion resulting from the inappropriate use of mean and variance with a skewed distribution. Motivated by the efficiency and skewness considerations above, we consider the possibility of alternative risk-return coordinates which implicitly include information about the skewness. To this end, we propose an alternative risk-return pair and show via an example that inefficiency occurs for a certain range of the feedback K. That is, a feedback K which is efficient from the classical mean-variance point of view may no longer be efficient with the new risk-return measure.

The remainder of the chapter is organized as follows: Section 2.1, devoted to the issue of skew and includes a motivating example which illustrates the deleterious effects of feedback in this context. Section 2.2 is also devoted to the study of skew using a different linear feedback scheme from the literature, the so-called Simultaneous Long-Short (SLS) controller described in Section 1.4 and [3–5]. Section 2.3 concentrates on the efficiency and finally, in Section 2.4, a brief conclusion is provided.

2.1 The Skewing Effect of Feedback

As mentioned above, if a linear feedback control with gain $K \geq 0$ is used, the resulting skewness $S(K)$ of the probability distribution for $g(T)$ can be so large as to render mean-variance information of questionable worth. Consider the case in which the amount invested is

$$I(t) = I_0 + Kg(t).$$

For the sake of self-containment, we provide the reader with the definition of skewness which is used. For the random variable $g = g(T)$, as shorthand, we denote its mean by M_g and its standard deviation by σ_g. Then, associated with the feedback gain K is the skewness

$$S(K) = \frac{\mathbb{E}[(g - M_g)^3]}{\sigma_g^3}.$$

Now, to demonstrate the skewing effect of feedback, we provide the following result.

2.1.1 Theorem: *Consider the idealized GBM Market with drift μ, volatility σ and linear feedback controller with gain K. Then, at time $T > 0$, the resulting trading gain or loss $g(T)$ has a probability density function with skewness which is independent of μ and given by*

$$S(K) = (e^{K^2\sigma^2 T} + 2)\sqrt{e^{K^2\sigma^2 T} - 1}.$$

Sketch of Proof: From the GBM model, we begin by noting that for any sample path $p(\cdot)$ for the price, it is easy to show that the resulting trading gain at time T is given by

$$g(T) = \frac{I_0}{K}\left[a\left(\frac{p(T)}{p(0)}\right)^K - 1\right]$$

where $a \doteq e^{\frac{1}{2}\sigma^2(K-K^2)T}$. It follows that the desired quantity $S(K)$ is precisely equal to the skewness of the log-normally random variable

$$X = \left(\frac{p(T)}{p(0)}\right)^K$$

whose scale parameter is given by $\sigma_X = K\sigma\sqrt{T}$. Now using the skew formula for the log-normal random variable given in the beginning of the chapter, we obtain

$$S(K) = (e^{K^2\sigma^2 T} + 2)\sqrt{e^{K^2\sigma^2 T} - 1}.$$

This completes the proof of the theorem. □

2.1.2 Illustrative Example: To illustrate how feedback affects the skewness of the probability distribution for $g(T)$, we consider Geometric Brownian Motion for the stock price with $\sigma^2 = 0.05$ which represents about a 22.36% annualized volatility. Taking $T = 0.4$ to represent a little over 100 trading days, in Figure 2.1.1 below, we provide a plot for the feedback-induced skew which results when the probability distribution for the trading gain $g(T)$ is considered. The key point to note is that for practical typical feedback gain values, say $0 \leq K \leq 10$, the skew becomes unreasonably large; e.g., as indicated in the introduction, $S(10) \approx 23.73$ which is an order of magnitude larger than the skew $S(1) \approx 0.4813$ associated with buy and hold at time $T = 0.4$.

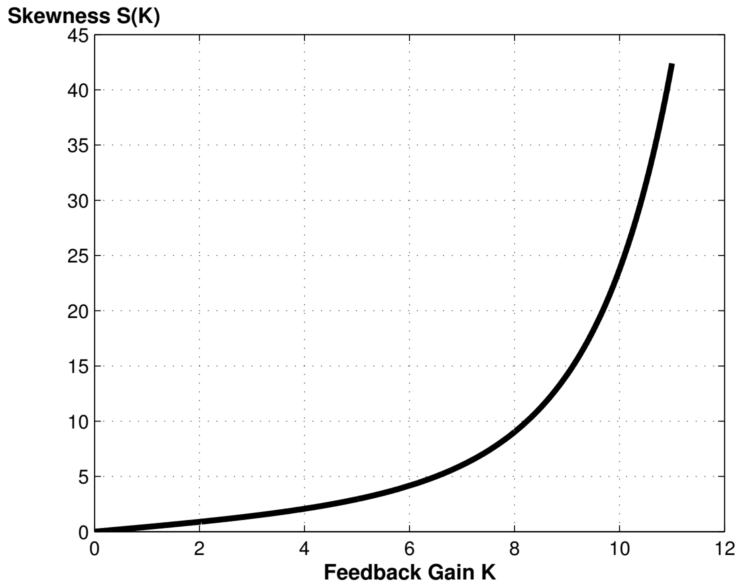

Figure 2.1.1: Plot of Skew for Linear Feedback in GBM Market

2.1.3 Example Continued (Mean-Variance Optimization): To further drive home the point that reliance on mean-variance with feedback may be inappropriate, we continue with the running example above and include the drift parameter $\mu = 0.25$ representing an annualized return of 25%. To carry out our analysis, we make use of two formulae for mean and variance of the trading gain-loss, which are easily derived from the results in [4] and were given earlier in Section 1.4.1.

Using these formulae, we now consider a classical mean-variance optimization, for example, see [97], to find a so-called "optimal" feedback gain K. More specifically, to work on a per-dollar basis we first set $I_0 = 1$ so that $g(T)$ corresponds to the rate of return. Now, with terminal time $T = 0.4$ and *risk aversion coefficient* $A \geq 0$, we construct a classical quadratic objective function as in the literature for our running example. Namely, we take

$$\begin{aligned} J(K) &= \mathbb{E}[g(T)] - 0.5A \operatorname{var}\big(g(T)\big) \\ &= \frac{1}{K}\left(e^{\frac{K}{10}} - 1\right) - \frac{0.5A}{K^2} e^{\frac{K}{5}} \left(e^{\frac{K^2}{50}} - 1\right) \end{aligned}$$

to be maximized. For this case involving drift $\mu = 0.25$, a trader seeing this upward trending bull market might declare "risk on" and reasonably set $A = 0.1$ to capture as much of the stock gain as possible without completely ignoring downside risk. For this case, from the resulting plot of $J(K)$

below in Figure 2.1.2 we obtain optimal feedback $K = K^*$, optimal cost $J = J^*$ and resulting skewness $S^* = S(K^*)$ given by $K^* \approx 9.81; \quad J^* \approx 0.15; \quad S^* \approx 21.42.$

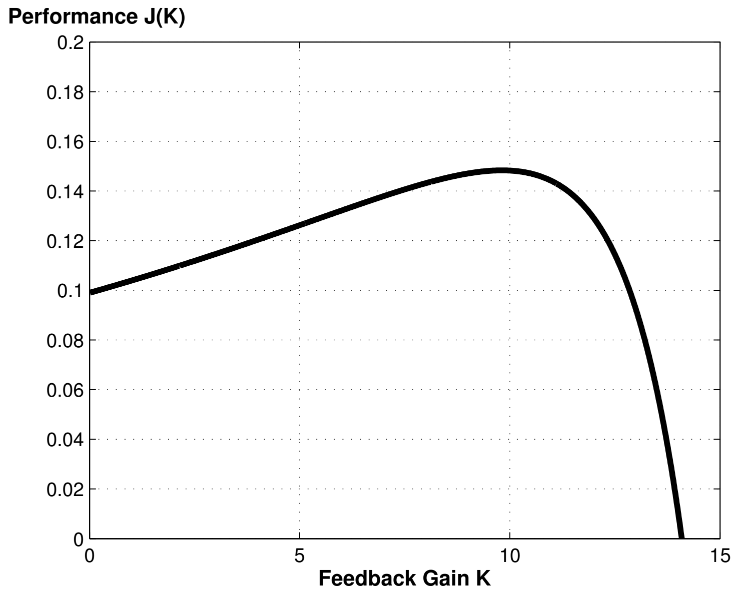

Figure 2.1.2: Mean-Variance Based Performance in GBM Market

To demonstrate how misleading this mean-variance optimization can be, we construct the *true cumulative distribution function* associated with the linear feedback controller. That is, via a lengthy but straightforward calculation,

$$P(g(T) \leq \gamma) = \phi\left(\frac{\log(I_0 + K\gamma) - \log(aI_0) - K\mu_Y}{K\sigma_Y}\right),$$

where $\phi(\cdot)$ is the cumulative distribution of the standard normal random variable $\mathcal{N}(0,1)$ and

$$\mu_Y \doteq (\mu - \frac{1}{2}\sigma^2)T; \quad \sigma_Y^2 \doteq \sigma^2 T.$$

Now, computing the mean and variance at the optimum feedback gain $K^* \approx 9.81$, we obtain

$$M_g^* \doteq \mathbb{E}[g(T)] \approx 0.1699; \ \sigma_g^* \doteq \text{var}\big(g(T)\big) \approx 0.4327.$$

Using the normal distribution $\mathcal{N}(M_g^*, \sigma_g^*)$, we can calculate the *implied probability* $P(g(T) \geq \gamma)$, which we call it $P_N(\gamma)$. We compare this quantity with the true probability which we call $P(\gamma)$ and consider rates of return $\gamma \geq 0$. These two, cumulative distributions are provided in Figure 2.1.3

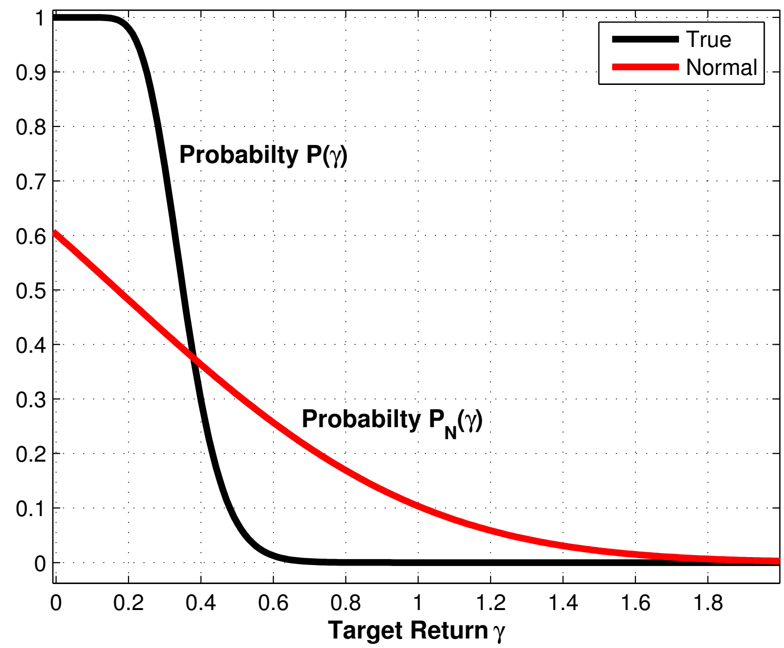

Figure 2.1.3: Probability of Reaching a Target Return γ

below. In the figure, the large differences between the two distribution functions is immediately evident. For example, when a target return of 20% is sought, the normal distribution understates the probability of success by nearly 50%. On the other hand, when a high target return such as 60% is sought, the normal distribution overstates the probability of success by about 25%. To conclude, mean-variance optimization provides a distorted view of a trader's prospect for success.

2.2 Skewness Formula for Simultaneous Long-Short Feedback Controller

In this section, we provide an extension of the skew formula $S(K)$ above which applies to another type of linear feedback controller called *Simultaneous Long-Short* introduced in [1] and [4] and pursued further in [3] and [5]. The description of strategy and the analysis of the probability density function and its first two moments are summarized in Section 1.4.3. For this type of trading, a formula for the resulting skew $S(K)$ in the overall trading gain $g(T)$ is now provided. Similar to the earlier analysis in Section 2.1, $S(K)$ can become unreasonably large.

2.2.1 Theorem: *Consider the idealized GBM Market with drift μ, volatility σ and SLS linear feedback controller with gain K. Then, at time $T > 0$, the resulting trading gain or loss $g(T)$ has probability density function with skewness given by*

$$S(K) = C(K,\mu,\sigma,T)[A(K,\mu,\sigma,T) - B(K,\mu,\sigma,T)]$$

where

$$A(K,\mu,\sigma,T) \doteq [e^{K\mu T} + e^{-K\mu T}]^2(e^{K^2\sigma^2 T} + 2);$$

$$B(K,\mu,\sigma,T) \doteq 3(e^{\frac{K^2\sigma^2 T}{2}} + e^{\frac{-K^2\sigma^2 T}{2}})^2;$$

$$C(K,\mu,\sigma,T) \doteq \frac{\sqrt{e^{K^2\sigma^2 T} - 1}(e^{K\mu T} + e^{-K\mu T})}{(e^{2\mu K T} + e^{-2\mu K T} + e^{-\sigma^2 K^2 T}\big]^{\frac{3}{2}}}.$$

Sketch of Proof: To simplify notation in the calculations to follow, we work with the scaled random variable $X \doteq p(T)/p(0)$ and define $M_p \doteq \mathbb{E}(X^K)$; $M_n \doteq \mathbb{E}(X^{-K})$; $a \doteq e^{\frac{1}{2}\sigma^2(K-K^2)T}$ and $c \doteq e^{\frac{-1}{2}\sigma^2(K+K^2)T}$. Now using the fact that the stochastic differential equation for g is integrable, for sample path $p(\cdot)$ we obtain

$$g(T) = \frac{I_0}{K}(aX^K + cX^{-K} - 2).$$

Next, we use the fact that the GBM process leads to X being log-normal. More precisely, we have $\log X \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ with K-th moment

$$\mathbb{E}[X^K] = e^{K\mu_Y + \frac{1}{2}K^2\sigma_Y^2}.$$

Now, a lengthy but straightforward computation leads to

$$\begin{aligned} M_g &= \frac{I_0}{K}\left(a\mathbb{E}(X^K) + c\mathbb{E}(X^{-K}) - 2\right) \\ &= \frac{I_0}{K}\left(e^{\mu K T} + e^{\mu K T} - 2\right) \end{aligned}$$

and denominator in the skewness given by

$$\sigma_g^3 = \frac{I_0^3}{K^3}\big[(e^{K^2\sigma^2 T} - 1)(e^{2\mu K T} + e^{-2\mu K T} + e^{-\sigma^2 K^2 T})\big]^{\frac{3}{2}}.$$

Next, to obtain the numerator in the skewness formula, we expand $(g - M_g)^3$ and use the fact that $M_p M_n = (ac)^{-1}$. A lengthy calculation yields

$$\begin{aligned}
\mathbb{E}[(g - M_g)^3] &= \frac{I_0^3}{K^3} \times \Big\{ a^3 \Big[\mathbb{E}(X^{3K}) - 3M_p \mathbb{E}(X^{2K}) + 2M_p^3 \Big] \\
&\qquad + 3a \Big[\frac{-1}{M_n} + 2M_p - \frac{1}{M_n} \mathbb{E}(X^{2K}) \Big] \\
&\qquad + 3c \Big[\frac{-1}{M_p} + 2M_n - \frac{1}{M_p} \mathbb{E}(X^{-2K}) \Big] \\
&\qquad + c^3 \Big[\mathbb{E}(X^{-3K}) - 3M_n \mathbb{E}(X^{-2K}) + 2M_n^3 \Big] \Big\}.
\end{aligned}$$

Substituting for $\mathbb{E}(X^{2K}), \mathbb{E}(X^{-2K}), \mathbb{E}(X^{3K}), \mathbb{E}(X^{-3K})$ and in last step for μ_Y and σ_Y leads to

$$\begin{aligned}
\mathbb{E}[(g - M_g)^3] &= -3(e^{\frac{1}{2}K^2\sigma^2 T} - e^{-\frac{1}{2}K^2\sigma^2 T})^2 (e^{K\mu T} + e^{-K\mu T}) \\
&\quad + (e^{3K^2\sigma^2 T} - 3e^{K^2\sigma^2 T} + 2)(e^{3K\mu T} + e^{-3K\mu T}).
\end{aligned}$$

Then, dividing by the expression for σ_g^3, and further simplifying we arrive at

$$S(K) = C(K, \mu, \sigma, T)[A(K, \mu, \sigma, T) - B(K, \mu, \sigma, T)].$$

This completes the proof of the theorem. □

2.3 Controller Efficiency Considerations

In accordance with the discussion in the beginning of the chapter, we now look at efficiency issues. The starting point for our analysis is that both the expected value and the variance of the trading gain are increasing with respect to the feedback K. This is true for both of the linear feedback trading schemes discussed earlier in this chapter. As a consequence of this monotonicity, when the expected value is plotted against the variance as a function of K, no pair $(\mathrm{var}\big(g(T, K)\big), \mathbb{E}[g(T, K)])$ dominates any other from a two-coordinate risk-return point of view. That is, all mean-variance pairs are efficient.

Our hypothesis is that this "all-K efficient result" gives an erroneous impression about efficiency because skew is neglected. When a different return-risk pair is used which incorporates more information about the distribution of $g(T)$ beyond the second moment, our hypothesis is that feedback gains in certain ranges can be ruled out based on efficiency considerations. That is, even the trader with a utility function reflecting very low risk aversion will limit the selection of the feedback K to those values in the efficiency regime. Said another way, on an equal return basis, the classical point of view is that all traders, independent of their utility functions, prefer less risk. Similarly, on an equal risk basis, all traders, independent of their utility functions, prefer a higher return.

2.3.1 Preamble on Efficiency Basics: We quickly review the notion of efficiency which is standard in the finance literature, for example, see [15]. Efficiency considerations also arise occasionally in the control literature in the context of Pareto optimality analysis. Indeed, we consider a process with possible outcomes $X \in \mathcal{X} \subseteq \mathbf{R}^2$ with components X_1 and X_2 representing some measure of risk and return respectively. In finance and engineering, when dealing with an investment with gain or loss g, the most classical measure of risk is the variance, var(g) and the most classical measure of return is the mean $\mathbb{E}[g]$.

Now, given a possible outcome $X \in \mathcal{X}$, it is said to be *inefficient* if there exists some $X' \in \mathcal{X}$ with either $X_1' \leq X_1$ and $X_2' > X_2$ or alternatively, $X_1' < X_1$ and $X_2' \geq X_2$. In other words, X' either has a higher return than X with no additional risk or it has a lower risk than X with at least as much return. We observe that $\mathcal{X}$ can be partitioned into a union of two disjoint sets: the inefficient set and its complement, the efficient set. In the case where X is inefficient as demonstrated by X', under some basic assumptions about utility functions, it can be argued that both investors will discard X in favor of X'.

2.3.2 Alternative Risk-Return Pair: As previously stated, the monotonicity of the mean and variance of $g(T)$ with respect to feedback gain K implies that all feedback controllers are efficient. The question we address in this section is the following: Might there be alternative risk-return measures which lead to a different conclusion in the presence of skewness? The example which we give below enables us to answer this question with a qualified "yes." That is, for the idealized GBM market and the alternative risk-return pair which we describe, we show via an example that the linear feedback controller $I(t) = I_0 + Kg(t)$ leads to risk-return combinations which are inefficient for low values of K. That is, there exists some feedback gain $K^* < \infty$ with the property that feedback gain $K < K^*$ has associated risk-return pair which is inefficient.

2.3.3 Construction of Proposed Risk-Return Pair: We begin by fixing a target return $\gamma \geq 0$ and take off from the fact that the idealized GBM market leads to a finite lower bound on the support of the trading gain $g(T)$. We let $W(K)$ denote the worst-case loss produced by the model and let $P_\gamma(K) = P(g(T) \geq \gamma)$; i.e., the probability of a successful trade. That is, instead of variance and mean of $g(T)$, we envision a trader whose underlying utility function depends on these new variables which are functions of the feedback gain K. Since γ is a parameter, one can vary this parameter and work with a family of curves which serve as a "menu" corresponding to differing returns.

2.3.4 Example Demonstrating Realization of Inefficiency: To show how the efficiency issue arises, we continue with the SLS trading scheme described in Section 2.2 using the parameters given in Section 2.1. We consider the target return 10% described by $\gamma = 0.1$. Via a lengthy calculation we obtain the worst-case loss and the corresponding probability of a successful trade

$$\begin{aligned} W(K) &= \frac{2I_0}{K}\left[1 - e^{-\frac{1}{2}\sigma^2 K^2 t}\right]; \\ P_\gamma(K) &= 1 - \phi\left(\frac{y_+ - \mu_Y}{\sigma_Y}\right) + \phi\left(\frac{y_- - \mu_Y}{\sigma_Y}\right) \end{aligned}$$

where $\phi(\cdot)$ denotes the cumulative distribution for the standard $\mathcal{N}(0,1)$ random variable, μ_Y, σ_Y and a are given in Section 2.1 and

$$y_\pm \doteq \log\left[\frac{1}{2a}(2 + \frac{K\gamma}{I_0} \pm \sqrt{(2 + \frac{K\gamma}{I_0})^2 - 4ac}\right]^{\frac{1}{K}}.$$

Examining the $(W(K), P_\gamma(K))$ plot in Figure 2.3.1, we note that the point where the worst-case trading loss $W(K)$ is maximized can readily be characterized. That is, by setting the derivative of $W(K)$ to zero, it is straightforward to show that this point of maximality is characterized by $K^* \approx 8.91$ and $W(K^*) \approx 0.13$.

Noting that the right side of the figure corresponds to low gain K, we see that there are many $(W(K), P_\gamma(K))$ pairs which are inefficient. That is, the risk-return pair associated with feedback gain K is inefficient because the same probability of success is guaranteed with some other gain K' with a lower level of risk.

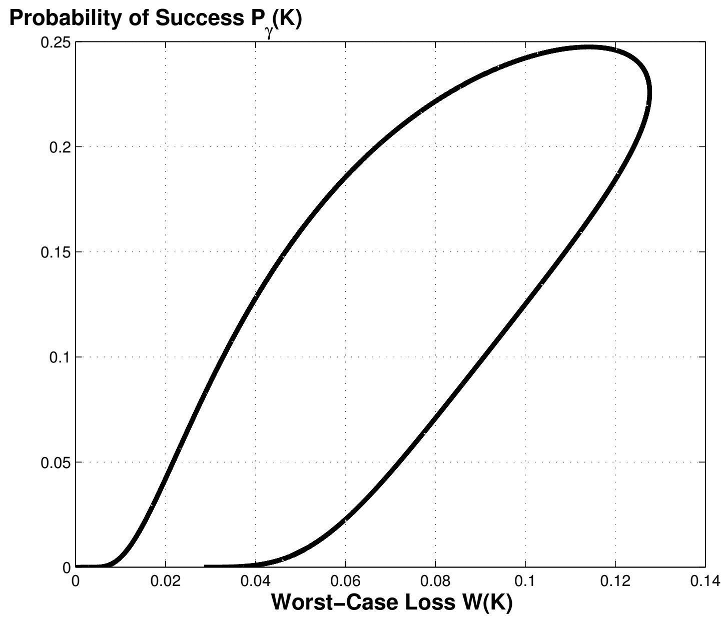

Figure 2.3.1: Demonstration of Inefficiency in SLS Trading

2.3.5 Numerical Example: The historical price for Pepsico Inc. (ticker PEP) over the 5 year period February, 12, 2007 to February, 10, 2012 was selected to generate an empirical plot using the new risk-return pair $(W(K), P_\gamma(K))$. We used a time interval $T = 1$ which was represented by 252 trading days and carried out a number of back-tests using the SLS trading strategy. To generate about 1000 one-year long sample paths, we used each of the days in the first four years as a starting point. Using $I_0 = 1$ and $\gamma = 0.1$, we generated empirical estimates of $P_\gamma(K)$ and $W(K)$ and plot this pair as we increased K over the interval $[0, 10]$. As seen in Figure 2.3.2, a regime of inefficiency was detected for $K \geq 5.2$ which corresponds to the maximum of $W(K)$.

2.4 Conclusion and Further Research

In this chapter, the focal point was the skewing effects of feedback controller gains on the probability distribution for the trading gain or loss $g(T)$. It is demonstrated that there are pitfalls associated with reliance on mean-variance based measures of performance. That is, when the feedback leads to a level of skewness $S(K)$ which is large, performance metrics based on mean and variance provide a distorted picture of the prospects for success.

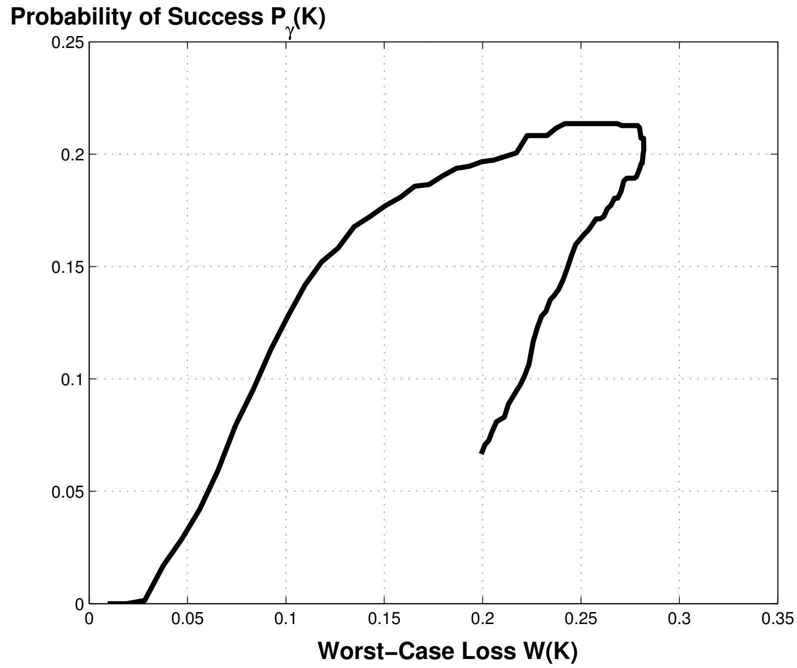

Figure 2.3.2: Efficiency Plot for Pepsico 2007-2012

Based on the work presented herein, one obvious question presents itself for consideration in future research: What type of objective function $J(K)$ should be used in carrying out an optimization problem for selection of the feedback gain K? Based on our arguments that a classical mean-variance utility function is inappropriate, one possibility would be to use a different risk-return pair which includes information about higher order moments of $g(T)$. Utility functions which capture higher moments especially the cubic utility function, have been discussed in portfolio optimization such as [100], [102], and [103] can be extended to our framework with appropriate risk aversion coefficients. Another possible example is when the issue of efficiency was considered, we introduced the pair $(W(K), P_\gamma(K))$. For such pair, the natural alternative to the quadratic objective would be

$$J(K) = P_\gamma(K) - AW(K)$$

where $A \geq 0$ is the risk aversion coefficient. Extending the skew analysis to a portfolio of stocks when linear feedback is employed would be another possible path to pursue; see Chapter 7 for further discussion along these lines.

Chapter 3

Drawdown Analysis for Stock Trading Via Linear Feedback

In this chapter, we provide the analysis of drawdown for a stock-trading strategy which is based on linear feedback control.[1] Significant motivation for our analysis is derived from the fact that *drawdown* is very important to stock traders and fund managers. That is, one typically monitors "drops in wealth" over time from highs to subsequent lows and investors often shy away from funds with a past history of large drawdowns. That is, when tracking a time-varying portfolio value $V(t)$, risk aversion dictates that drops from peaks to later valleys should not be too large. Careful monitoring of drawdown in $V(t)$ is one of the most important aspects of risk management; e.g., see [56,92–94]. A large drawdown is associated with significant "tail risk" and large skewness of the probability distribution for wealth. For such situations, a classical mean-variance analysis as in the celebrated work of Markowitz [15] does not suffice.

To provide further context for this chapter, we note that over the last two decades, a significant body of research has been involved in the development of new risk measures for financial markets. Until 1999, perhaps the most widely used of these measures was the well-known "Value at Risk" which is known as VaR; e.g., see [104]. In a seminal paper [105], the notion a of "coherent risk measure" is introduced. To complete this extremely brief overview of risk measures, we mention [104, 106–108] which provide examples of other risk measures commonly used. Suffice it to say, the *drawdown measure* is one of these and has received considerable attention over the past ten years. In this regard, the important results on stock price drawdown in [92] strongly motivate the research reported in this chapter.

With this motivation in mind, this chapter addresses the analysis of drawdown when a feedback control is employed in stock trading in an idealized market with prices governed by Geometric Brownian Motion. We begin with a result in the applied probability literature which is applicable to cases involving buy-and-hold. Subsequently, after modifying this result via an Ito correction to

[1] The results reported in this chapter have been published in [8].

account for geometric compounding of the daily stock price, we consider the effect on drawdown when a simple pure-gain feedback control is used to vary the investment $I(t)$ over time. That is, letting $V(t)$ denoting the trader's account value at time $t \geq 0$, when a feedback control $I = KV$ is used to modify the amount invested, the buy-and-hold result no longer applies. Our first result is a formula for the expected value for the maximum drawdown in logarithmic wealth $\log(V(t))$. This formula is given in terms of the feedback gain K, the price drift μ, the price volatility σ and terminal time T.

Subsequently, using a fundamental relationship between logarithmic and percentage drawdowns, we obtain an estimate for the expected value of the maximum percentage drawdown of $V(t)$. This chapter also includes an analysis of the asymptotic behavior of this drawdown estimate as $T \to \infty$ and Monte Carlo simulations aimed at validation of our estimates. These simulations are also used to illustrate how drawdown behavior can be studied for other classes of controllers. To this end, we also consider the Simultaneous Long-Short feedback trading strategy and give a brief discussion of its expected maximum percentage drawdown.

A distinction between the stock price drawdown results cited above and our research involving trading is important to make: When we use feedback to continuously modify the amount invested $I(t)$, drawdown in the account value $V(t)$ can be dramatically different from drawdown in the stock price $p(t)$. They are only equivalent when $I(t)$ corresponds to buy-and-hold with full investment $I(t) \equiv V(t)$. In our case, when feedback is used, the result can be an account value with a highly skewed probability distribution; this effect is quantified in Chapter 2 and [7]. Finally, context for this chapter is provided by the emerging line of research on the use of feedback in trading; e.g., see [2–8, 13, 26–28, 31, 36, 37, 40–44]. This chapter addresses the issue of drawdown in this context. Although an investment may end up with a large gain at the terminal time, if there is a large drawdown along the way, the performance may be deemed to be unsatisfactory.

3.0.1 The Starting Point: The starting point in this chapter is an idealized market, fully described in Section 1.2, with stock price p, governed by the Geometric Brownian Motion (GBM), as described in Section 1.3; that is,

$$\frac{dp}{p} = \mu dt + \sigma dZ$$

where μ is the drift, σ is the volatility and $Z(t)$ is a standard Wiener process. When a trader is brought into the picture with investment $I(t)$ for $0 \leq t \leq T$, the resulting stochastic increment for the account value becomes

$$dV = \frac{dp}{p}I + r(V - I)dt$$

where $r \geq 0$ is the risk-free rate of return and for simplicity, we also take r as the margin rate. In this setting, the issue of drawdown, formally defined in Section 3.1, arises.

Our analysis of drawdown begins by modifying the results in [92] to address dp/p instead of dp. Subsequently, we generalize the existing results to analyze the drawdown in the account value $V(t)$ when the *proportional-to-wealth* investment scheme introduced earlier is used; that is, $I = KV$ with $K > 0$. This sort of investment scheme is rather classical in financial markets; e.g., see [109]. To this end, we obtain a formula for drawdown in $V(t)$ which specializes to the existing result in the literature for the special case of buy-and-hold which is obtained when $K = 1$.

3.0.2 Collateral Considerations: The results in this chapter apply to an "idealized market" as introduced in Section 1.2. One of the assumptions in an idealized market is that the trader has adequate collateral to make a trade when $I(t) > V(t)$. In practice, when such trades are made, margin interest is involved. Hence, in the sequel, we take $r \geq 0$ to be the margin interest rate which will apply when $K > 1$; i.e., with $I = KV$, the trader pays interest on $I - V = (K - 1)V$. For the case when $K \leq 1$, we assume the same r to be the risk-free rate of return; more specifically, it denotes interest accrued on $V - I = (1 - K)V$. In practice, the margin interest rate and risk-free rate of return are not equal, and this assumption is used to simplify the analysis.

3.0.3 The Governing Stochastic Equations: In an idealized GBM market with linear feedback $I = KV$, it is easy to show that the price $p(t)$ induces a GBM on the account value $V(t)$. To see this, we begin with a closed loop stochastic differential equation for the account value. In view of the discussion above, the stochastic increment for V satisfies

$$\begin{aligned} dV &= \frac{dp}{p}I - r(I - V)dt = \frac{dp}{p}KV - r(K - 1)Vdt \\ &= [K\mu - r(K - 1)]Vdt + K\sigma VdZ. \end{aligned}$$

Hence, when we divide by V, the induced account value dynamics follows a GBM process

$$\frac{dV}{V} = \mu' dt + \sigma' dZ,$$

with modified drift μ' and modified volatility σ' given by

$$\mu' \doteq K\mu - r(K-1); \qquad \sigma' \doteq K\sigma.$$

It is also important to note that *logarithmic wealth* $\log(V(t))$ is also a quantity which is monitored by traders. In this case, via straightforward application of Ito's Lemma, for example see [88], we modify the formulae above and obtain a standard Brownian Motion with increment

$$d(\log V) = \mu_* dt + \sigma_* dZ$$

where $\sigma_* \doteq \sigma' = K\sigma$ and

$$\mu_* \doteq K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2.$$

This result implies that the Brownian Motion is preserved when linear feedback on the account value is used to determine the investment level.

3.1 Drawdown Definitions

For a given continuous function in time, $V(t)$, we recall from Subsection 1.5.2 that the *maximum absolute drawdown* is defined by

$$D_{max}(V) \doteq \max_{0 \le s \le t \le T} V(s) - V(t).$$

When we replace V by $\log(V)$ above, we obtain $D_{max}\big(\log(V)\big)$, the logarithmic wealth version of absolute drawdown. The *maximum percentage drawdown of* V is similarly defined as

$$d_{max}(V) \doteq \max_{0 \le s \le t \le T} \frac{V(s) - V(t)}{V(s)}.$$

Since increments dV are obtained as percentages of the current account value $V(t)$ at time t, the denominator in the definition of d_{max} cannot vanish. Hence, this quantity is well defined. Since $V(t)$ cannot become zero, $D_{max}\big(\log(V)\big)$ is also well defined. As shown in Subsection 3.0.3, $V(t)$ is a GBM process and its sample paths are continuous; which makes the definition of the maximum absolute drawdown and maximum percentage drawdown applicable to $V(t)$.

3.1.1 Absolute Versus Percentage Drawdown: When a maximum absolute drawdown occurs, it is not necessarily the case that this corresponds to a maximum percentage drawdown and vice versa. We recall the function $V(t)$ shown in Figure 1.5.1 which is also provided below in Figure 3.1.1. It has two major drops: a drop from $V(1) = 0.2$ to $V(2) = 0.05$ and another drop from $V(10) = 10$ to $V(11) = 5$. From the plot, we obtain $D_{max}(V) = 5$ and $d_{max}(V) = 0.75$. Moreover, these drawdowns are different as far as their times of occurrence are concerned.

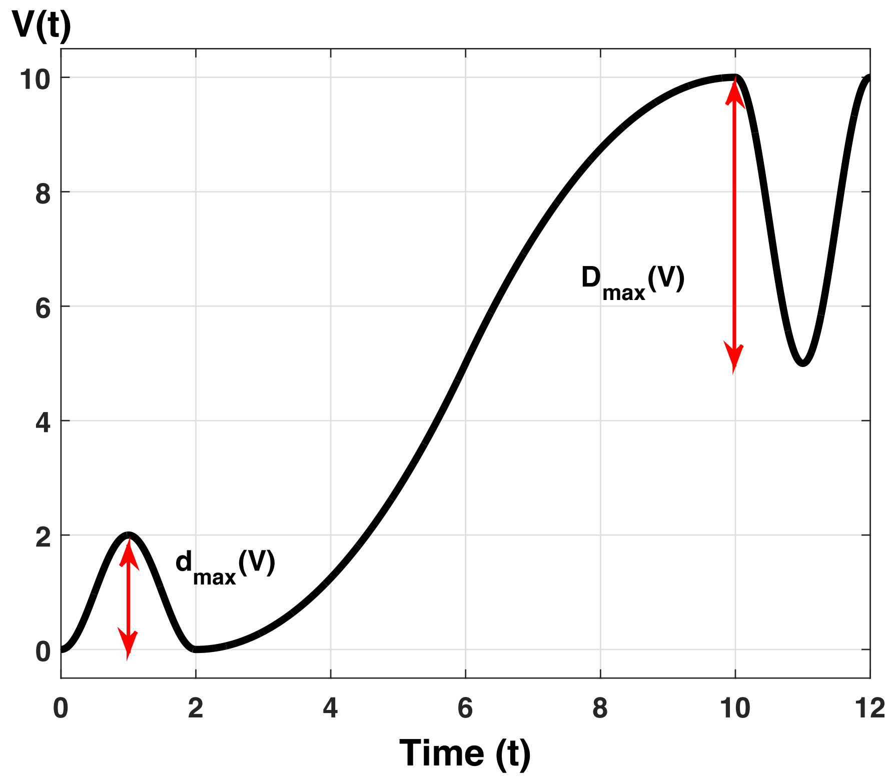

Figure 3.1.1: Maximum Drawdown Versus Maximum Percentage Drawdown

3.2 Main Result

In this section, we determine the formula for the expected maximum absolute drawdown of the logarithmic wealth, $\log(V)$, which is widely used in finance and also derive a formula for the upper bound on the expected maximum percentage drawdown of wealth, V, using the formula we obtained for logarithmic wealth. When the GBM model for the price generically has non-zero volatility σ, common sense reasoning dictates that a "gambler's ruin" type of situation presents itself. That is, when the terminal time T is sufficiently large, it becomes exceedingly likely that somewhere along the way, a drawdown which approaches 100% will occur. Furthermore, common sense also dictates that the larger the ratio μ/σ, the longer we expect it to take before one experiences a "bad run" with $V(t)$ tending to zero. The results presented in this section are seen to confirm this intuitive reasoning.

In the sequel, our plan is as follows: We present a preliminary lemma which connects $d_{max}(V(t))$ to $D_{max}(\log(V(t)))$ along sample paths of the closed loop system. Subsequently, we bring the Q-functions, introduced in [92], into our analysis in two ways: First, we modify these functions so that they apply in the feedback case rather than buy-and-hold. Second, we bring percentage drawdown into the picture with the help of the preliminary lemma to follow. Note that the lemma below can actually be given in a much more general form than what is used for our analysis. If $V(t)$ is any positive continuous on $[0, T]$ rather than GBM, the same proof can be used.

3.2.1 Preliminary Lemma: *Given any sample path for the account value $V(t)$, it follows that*

$$d_{max}(V) = 1 - e^{-D_{max}(\log(V))}.$$

Proof: Recalling that the sample path $V(t)$ is continuous and non-vanishing on $[0, T]$, $\log(V(t))$ is continuous and hence, $D_{max}(\log(V))$ is well defined as a maximum rather than a supremum. Now, for a sample path $V(t)$, let (s^*, t^*) be any pair which achieves $D_{max}(\log(V))$; i.e.,

$$D_{max}(\log(V)) = \log(V(s^*)) - \log(V(t^*)) = \log\left(\frac{V(s^*)}{V(t^*)}\right).$$

Now, since the $\log$ function is increasing, the same pair (s^*, t^*) maximizes $\frac{V(s)}{V(t)}$ and therefore minimizes $\frac{V(t)}{V(s)}$. In turn, this is equivalent to the pair (s^*, t^*) maximizing

$$1 - \frac{V(t)}{V(s)} = \frac{V(s) - V(t)}{V(s)}.$$

The last step is to recognize that

$$\begin{aligned} d_{max}(V) &= 1 - \frac{V(t^*)}{V(s^*)} = 1 - e^{-\log\left(\frac{V(s^*)}{V(t^*)}\right)} \\ &= 1 - e^{-D_{max}(\log(V))}. \ \square \end{aligned}$$

3.2.2 Introduction to Q Functions: In [92] and [110], a pair of real-valued functions $Q_p(x)$ and $Q_n(x)$ are introduced in the analysis of absolute drawdown for the Brownian Motion given by $dp = \mu dt + \sigma dZ$. These two functions involve rather complicated integrals which are numerically computed and stored as a table of values. In Figure 3.2.1, plots are given which summarize the data describing these two numerical functions.

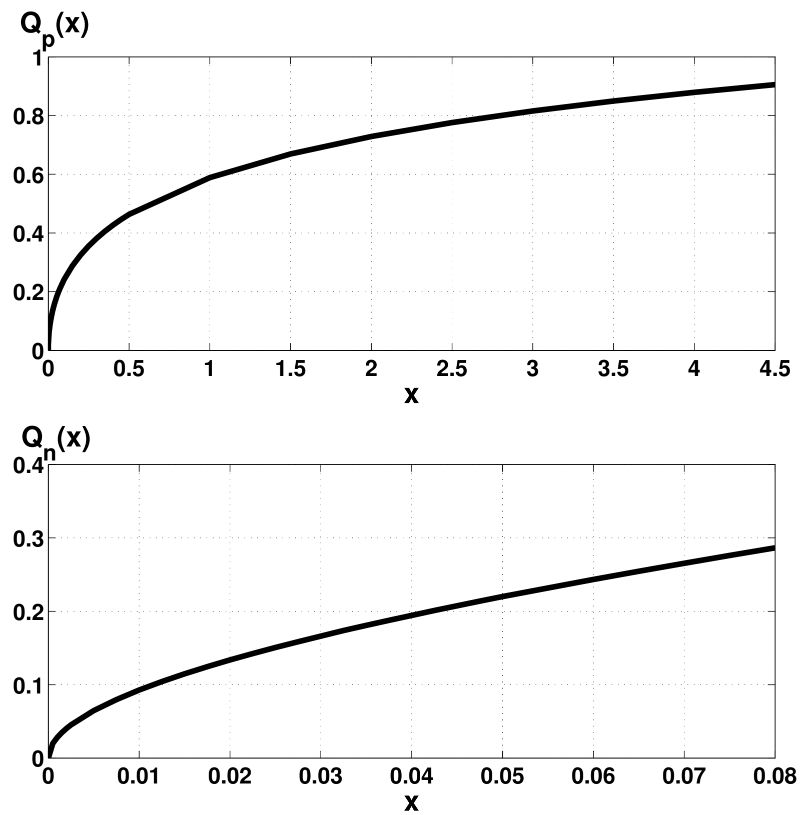

Figure 3.2.1: Plot of the Q Functions

For the special case when only the drawdown in price is concerned, as discussed in the beginning of the chapter, this corresponds to buy-and-hold in our formulation with initial account value $V(0) = p(0)$ and feedback gain $K = 1$ and investment $I(t) = V(t) = p(t)$. Hence, for this special case, the results in [92] tell us that

$$\mathbb{E}\left(D_{max}(V)\right) = \begin{cases} \frac{2\sigma^2}{\mu} Q_p\left(\frac{\mu^2 T}{2\sigma^2}\right) & \text{if } \mu > 0; \\ 1.2533\sigma\sqrt{T} & \text{if } \mu = 0; \\ \frac{-2\sigma^2}{\mu} Q_n\left(\frac{\mu^2 T}{2\sigma^2}\right) & \text{if } \mu < 0. \end{cases}$$

In addition to handling arbitrary feedback gains, we want to address the case when prices follow the more realistic percentage change model in dp/p rather an absolute change model given above for p. These issues are addressed below.

3.2.3 Main Result: In the theorem below, for simplicity of notation, we assume margin and interest rate $r = 0$. The more general case is discussed immediately after the theorem. We also provide some remarks about the drawdown formula obtained and address the asymptotic case when $T \to \infty$. In Section 3.3 we address the tightness of the upper bound via Monte Carlo analysis. This enables a comparison between our drawdown estimate and its true value.

3.2.4 Theorem: *For the feedback control $I = KV$ in the idealized GBM market with dynamics $dp/p = \mu dt + \sigma dZ$ and $r = 0$, the maximum absolute drawdown of logarithmic wealth has expected value*

$$\mathbb{E}\left(D_{max}(\log V)\right) = \begin{cases} \frac{2K\sigma^2}{\mu - \frac{1}{2}K\sigma^2} Q_p\left(\frac{\left(\mu - \frac{1}{2}K\sigma^2\right)^2 T}{2\sigma^2}\right) & \text{if } K < \frac{2\mu}{\sigma^2}; \\ 1.2533 K\sigma\sqrt{T} & \text{if } K = \frac{2\mu}{\sigma^2}; \\ -\frac{2K\sigma^2}{\mu - \frac{1}{2}K\sigma^2} Q_n\left(\frac{\left(\mu - \frac{1}{2}K\sigma^2\right)^2 T}{2\sigma^2}\right) & \text{if } K > \frac{2\mu}{\sigma^2} \end{cases}$$

with corresponding maximum percentage drawdown satisfying the condition

$$\mathbb{E}\big(d_{max}(V)\big) \leq 1 - e^{-\mathbb{E}\left(D_{max}(\log(V))\right)}.$$

Proof: Recalling the analysis given in Section 3.1, the stochastic differential equation governing $\log V$ is a Brownian Motion with drift and volatility given by

$$\mu_* \doteq K\mu - \frac{1}{2}K^2\sigma^2; \quad \text{and} \quad \sigma_* \doteq K\sigma.$$

Now, substitution of μ_* for μ, σ_* for σ and $\log V$ for V in $\mathbb{E}(D_{max}(V))$ above, we obtain

$$\mathbb{E}\left(D_{max}(\log(V))\right) = \begin{cases} \frac{2{\sigma_*}^2}{\mu_*} Q_p\left(\frac{{\mu_*}^2 T}{2{\sigma_*}^2}\right) & \text{if } \mu_* > 0; \\ 1.2533\sigma_*\sqrt{T} & \text{if } \mu_* = 0; \\ \frac{-2{\sigma_*}^2}{\mu_*} Q_n\left(\frac{{\mu_*}^2 T}{2{\sigma_*}^2}\right) & \text{if } \mu_* < 0. \end{cases}$$

Replacing μ_* and σ_* with their corresponding values in terms of μ, σ and $K > 0$, a straightforward calculation leads to the formula given for $\mathbb{E}\left(D_{max}(\log(V))\right)$. To complete the proof, we note the following: For an arbitrary sample path $V(t)$, recalling the Preliminary Lemma 3.2.1, we have

$$d_{max}(V) = 1 - e^{-D_{max}(\log(V))}.$$

Now, since $1-e^{-x}$ is concave, upon taking the expectation and using the Jensen's inequality [111], we obtain

$$\mathbb{E}\big(d_{max}(V)\big) \leq 1 - e^{-\mathbb{E}\big(D_{max}(\log(V)\big)}. \ \square$$

3.2.5 Non-Zero Margin and Interest Rates: For the case of non-zero margin and interest rates, via a lengthy but straightforward computation, we can readily modify the results in the theorem. We illustrate for $K > 1$, the case when margin arises. Indeed, let

$$\mu_* \doteq K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2;$$

$$K_p \doteq \frac{(\mu - r) + \sqrt{(\mu - r)^2 + 2r\sigma^2}}{\sigma^2}.$$

If $K_p \geq 1$, then there are again three regimes for the drawdown. With a simple calculation, the small-K regime which corresponds to $\mu_* > 0$ is equivalent to $K \in [1, K_p]$ and leads to

$$\mathbb{E}\big(D_{max}(\log(V))\big) = \frac{2K^2\sigma^2}{K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2} \times Q_p\Big(\frac{(K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2)^2 T}{2K^2\sigma^2}\Big).$$

In the second regime obtained with $\mu_* = 0$ or $K = K_p$,

$$\mathbb{E}\big(D_{max}(\log(V))\big) = 1.2533K\sigma\sqrt{T}.$$

Finally, for the third regime when $\mu_* < 0$ or $K > K_p$,

$$\mathbb{E}\big(D_{max}(\log(V))\big) = \frac{-2K^2\sigma^2}{K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2} \times Q_n\Big(\frac{(K\mu - r(K-1) - \frac{1}{2}K^2\sigma^2)^2 T}{2K^2\sigma^2}\Big).$$

In contrast, if $K_p < 1$, with $K > 1$ there is only one regime, $\mu_* < 0$, which is exactly the third regime above.

3.2.6 Remarks and Asymptotic Behavior: In the drawdown formulae above, the three regimes for the feedback gain K can be linked to a signal-to-noise type ratio μ/σ for the underlying price process. For a given feedback gain K, the smaller this ratio, the more drawdown we expect to see. For fixed μ and σ, if we allow K to increase, we expect to see larger and larger drawdowns. It is also instructive to consider the asymptotic versions of the results given in the theorem which are obtained as $T \to \infty$. Analogous to the arguments used in the proof of the theorem, we can readily modify the asymptotic analysis of functions $Q_p(x)$ and $Q_n(x)$ and apply these results to the feedback control problem being considered here. Namely, beginning with the asymptotic estimates

$$Q_p(x) \approx \frac{1}{4}\log x + 0.4988;\; Q_n(x) \approx x + \frac{1}{2},$$

for x suitably large, via a lengthy but straightforward calculation along the lines given in the proof of the theorem, as $T \to \infty$, the quantity $\mathbb{E}\left(D_{max}(\log(V))\right)$ is estimated in three regimes as follows: In the first regime, obtained with $K < 2\mu/\sigma^2$,

$$\begin{aligned}\mathbb{E}(D_{max}(\log(V)) \quad &\approx \quad \frac{4K\sigma^2}{2\mu - K\sigma^2} \\ &\qquad \times (0.63519 + 0.5\log T + \log\frac{\mu - 0.5K\sigma^2}{\sigma}).\end{aligned}$$

In the second regime, with critical value $K = 2\mu/\sigma^2$,

$$\mathbb{E}(D_{max}(\log(V)) \approx 1.2533K\sigma\sqrt{T}.$$

Finally, in the third regime, obtained with $K > 2\mu/\sigma^2$,

$$\mathbb{E}(D_{max}(\log(V)) \approx -(\mu K - 0.5K^2\sigma^2)T - \frac{K\sigma^2}{\mu - 0.5K\sigma^2}.$$

For all three regimes of K, using the formulae in the theorem, as expected, we see

$$\lim_{T\to\infty} \mathbb{E}(D_{max}(\log(V)) = \infty.$$

However, the specific value of K has a significant impact on the rate of convergence for this limit. For the small-K regime, a rate of $\frac{1}{T}$ is obtained from the formulae above. For the critical value case, the rate is $e^{-\sqrt{T}}$ and finally, for the large-K regime, the rate becomes e^{-T}.

3.3 Numerical Examples with Simulations

In this section, we provide plots of the drawdown functions indicating how risk evolves as a function of the duration of the trade T. We recall that when we work with percentage, the theorem in Section 3.2 provides an upper bound which arises when Jensen's Inequality is invoked. Hence, it is natural to ask whether the upper bound we obtain is "tight." We conducted Monte Carlo simulations in order to compare our upper bound with a Monte Carlo estimate of true drawdown. Indeed, we considered a stock with GBM model with time t measured in years, $T = 5$, annualized drift $\mu = 0.25$ and annualized volatility $\sigma = \sqrt{0.5} \approx 0.7071$. These values were picked intentionally so that the so-called critical value of K in the theorem is given by $K^* \doteq \frac{2\mu}{\sigma^2} = 1$.

For our simulations, one value of K was used for each of the three regimes in the theorem. More specifically, we took $K = 0.1$, $K = 1$ and $K = 2$ in our computations. For these three cases, the Monte Carlo estimates of the $\mathbb{E}\big(d_{max}(V)\big)$ along with the upper bounds are shown in Figure 3.3.1. For each value of T, our Monte Carlo estimate for the expected value of d_{max} was generated using 5000 sample paths. Our estimates appear to have converged quite well in that we do not see significant changes above the one thousand sample path level. In the three simulations below, we see that the "error" between the upper bound in the theorem versus the Monte Carlo estimate is different in each of the three regimes for K. In all cases, the error is at most a few percent.

3.3.1 Other Feedback Based Strategies: The drawdown concepts presented here not only apply to the linear feedback $I = KV$ but also to practically any other linear feedback control law which one might imagine. To illustrate, we consider another known strategy which uses a combination of two linear feedbacks, one long trade in combination with one short trade. We recall, this is the so-called Simultaneous Long-Short (SLS) feedback law described earlier in Subsection 1.4.3 and which has been studied in detail in [1–5, 7, 8]. Whereas this feedback control law in these papers is a mapping on the trading gain $g(t)$, so as to maintain consistency with the formulation considered here, we study a version of SLS which operates on $V(t)$.

Further recalling, this strategy can be viewed as the superposition of two independent trades as follows: The trader holds both a long investment $I_L(t) > 0$ and short investment position $I_S(t) < 0$ at the same time. In practice, these two positions can be "netted out" so that the overall investment

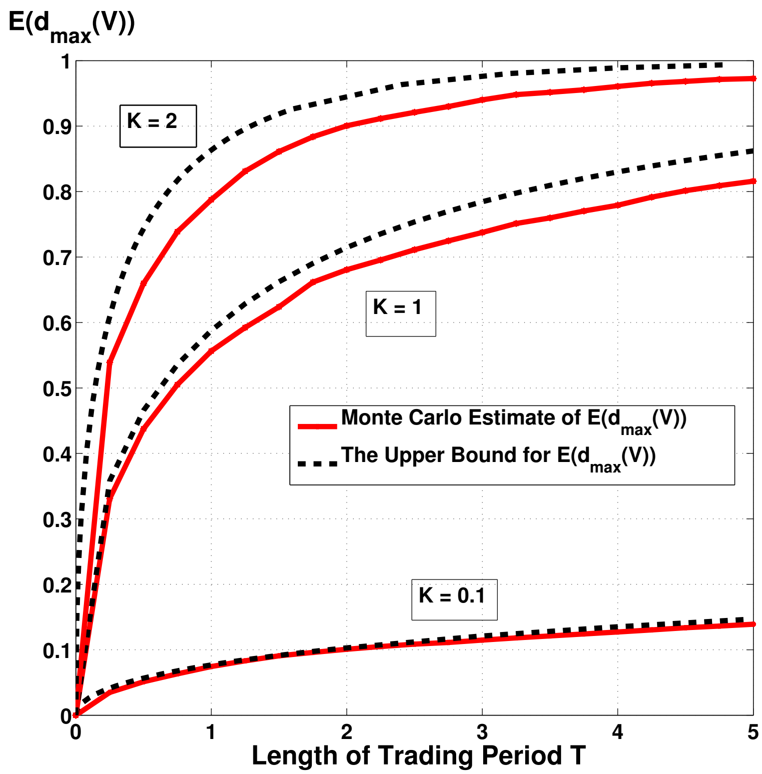

Figure 3.3.1: Bound on $\mathbb{E}\big(d_{max}(V)\big)$ for Different Values of K

is given by $I(t) = I_L(t) + I_S(t)$. Now, for feedback gain $K > 0$, these two investment components are given by $I_L(t) = KV_L(t)$ and $I_S(t) = -KV_S(t)$ with initial conditions $V_L(0) = V_S(0) = V_0/2$ and resulting account value evolving over time as $V(t) = V_L(t) + V_S(t)$. For this SLS system, the results in the literature, for example, see [4], [7] and Subsection 1.4.4 tell us that for all but the degenerate case when $\mu = 0$, the positive expectation condition $\mathbb{E}[V(T)] > V_0$ is guaranteed; that is, the Robust Positive Expectation Property holds. In addition, the mean, variance and skewness of the probability density function of $V(T)$ are increasing functions of K. We now analyze the drawdown of this scheme.

Since $I_L(t)$ and $I_S(t)$ are opposite in sign, it is natural to expect that the drawdown of V will be smaller than the drawdowns of V_L and V_S separately. Hence, it is also natural to conjecture that the SLS drawdown is always lower than the the one obtained in the simulation for the pure long case $I = KV$ above. Furthermore, as time goes on, a large "run-up" in one of these positions is exceedingly probable. Hence we expect to see the drawdown for SLS behave much the same as the pure long case. That is, we conjecture that $\mathbb{E}\big(d_{max}(V)\big)$ should go to 1 as time T tends to infinity. Finally, in view of the reasoning given above, we expect to see the rate of convergence for this SLS case to be slower than what we obtained with a purely long controller.

Indeed, a Monte Carlo simulation for the SLS controller was carried out for $K = 2$ using the same parameters as in the pure-long simulation above. In Figure 3.3.2, we see that the result is consistent with the conjectures given above. For T small, we see $\mathbb{E}(d_{max}(V))$ for the SLS case to be significantly below that obtained for the pure long case with asymptotic behavior as predicted.

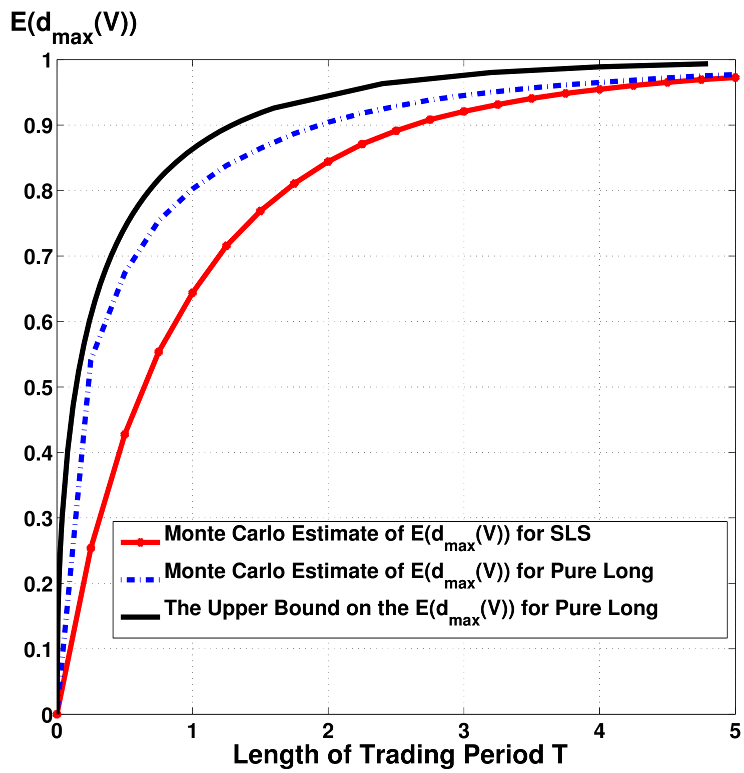

Figure 3.3.2: Estimates of $\mathbb{E}\big(d_{max}(V)\big)$, SLS and Pure Long for $K = 2$

3.4 Conclusion and Further Research

In the proof of the theorem in Section 3.2, we used Jensen's inequality. The results in [92] suggest an avenue of analysis which avoids introducing this inequality but may be computationally prohibitive. That is, $\log(V)$ is a standard Brownian motion, modifying the theory in [92] should make it possible to construct the probability density function for $D_{max}(\log(V))$. As a practical matter, however, in lieu of finding an exact representation of the probability density function, it appears much easier to carry out a Monte Carlo simulation and obtain a histogram to represent the density function for $D_{max}(\log(V))$ or even $d_{max}(V)$. This is illustrated in Figure 3.4.1 where the histogram for $d_{max}(V)$ is provided for the trading scenario described by the case $\mu = 0.5, \sigma = 0.5, T = 0.4$ and $K = 1$. The expected value of the maximum percentage drawdown for V is estimated to be $\mathbb{E}(d_{max}(V)) \approx 0.2558$. In contrast, the upper bound given in the theorem in Section 3.2, is approximately 0.27.

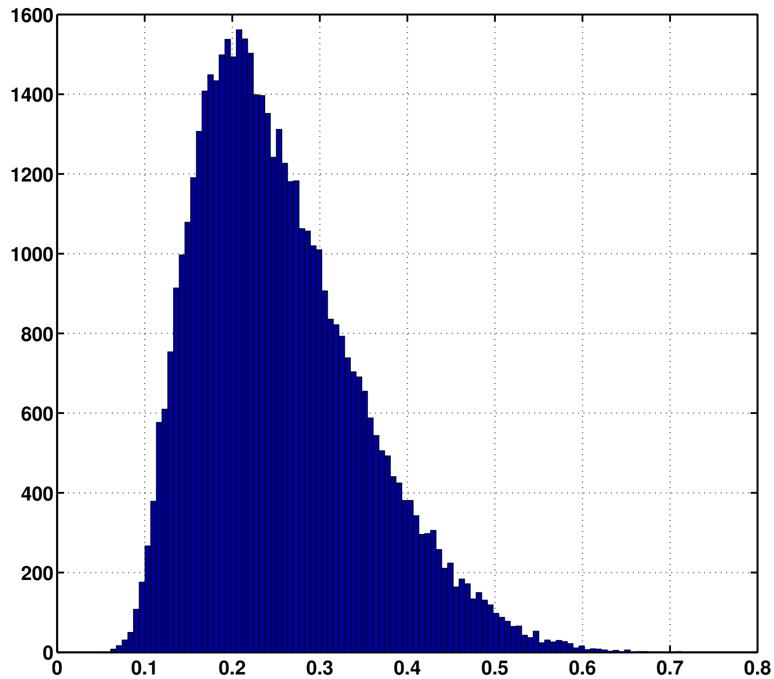

Figure 3.4.1: Histograms of $d_{max}(V)$

Finally, to conclude, we refer back to the discussion in the beginning of the chapter, involving the ongoing line of research on risk measures. Analogous to the information provided by a conditional value at risk measure, in the case of drawdown, a conservative investor might want to know how large the downside can be if $d_{max}(V) \geq \mathbb{E}(d_{max}(V))$ is experienced. Hence, one can define a conditional version of the percentage measure. That is, along a sample path $V(t)$, let

$$\mathcal{D}_{max}(V) \doteq \mathbb{E}\big(d_{max}(V)|d_{max}(V) \geq \mathbb{E}(d_{max}(V))\big) \geq \mathbb{E}(d_{max}(V)).$$

Finally, it would also be of interest to study the extent to which various measures of drawdown are compatible with the theory of coherent risk measures; e.g., see [105].

Chapter 4

PI Controllers: The Robust Positive Expectation Property

In this chapter, a generalization of the Simultaneous Long-Short (SLS) trading rule, described in Section 1.4, is studied.[1] This new strategy, called *Initially Long-Short (ILS)*, involves a controller which includes memory of the past performance. This is accomplished via use of an integrator as in classical control theory. The dynamic compensator which results is called a Proportional-Integral (PI) controller. More specifically, the main objective in this chapter is to generalize the results of Section 1.4.4 from static to dynamic feedback. In the static case, in combination with a "pure gain" which was the focal point of this thesis so far, the investment level at time t is given by

$$I(t) = I_0 + Kg(t),$$

where I_0 is the initial investment, K is the feedback gain and $g(t)$ is the cumulative gain-loss function up to time t; see 1.4.1 for details.

Our analysis begins by reducing the stochastic trading equations for the expectation of $g(t)$ to a classical second order system and finding the closed-form solution to prove that the Robust Positive Expectation Property still holds. Later in the chapter, we also consider a number of other issues such as the analysis of the variance of $g(t)$ and the monotonic dependence of $g(t)$ on the feedback gains. In addition, we provide simulations showing how the PI controller performs on a backtest with prices obtained from historical data. Finally, motivated by practical consideration described in the section on simulation, we provide a more general theorem which includes a exponentially weighting factor in the PI controller. More specifically, this new discounting scheme amounts to a modification of the investment rule to more heavily emphasize recent data.

[1]The results reported in this chapter have been published in [9].

4.0.1 The PI Controller and Issues we Address: To determine the investment level, we consider a classical Proportional-Integral (PI) controller

$$I(t) = I_0 + K_P g(t) + K_I \int_0^t g(\tau) d\tau,$$

where $K_P \geq 0$ and $K_I \geq 0$ are respectively the so-called proportional and integral gains. The block diagram for this investment rule is provided in Figure 4.0.1.

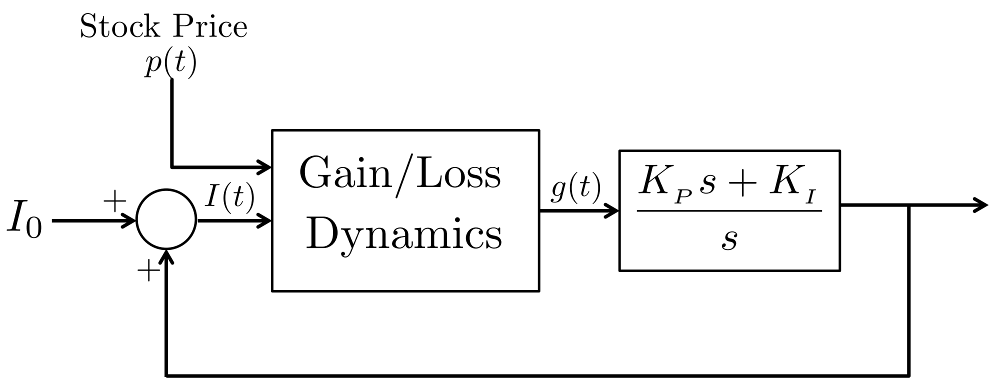

Figure 4.0.1: Block Diagram of Trading via PI Controller

Note that we use differentiator-free dynamics to avoid problems associated with price signals which typically include high-frequency components. Our goal is to analyze the behavior of the mean and variance of $g(t)$. Working with a long-short version of the PI controller above, our main result is that the Robust Positive Expectation Property, $\mathbb{E}[g(t)] > 0$ still holds except for the trivial break-even case with either both feedback gains $(K_P, K_I) = (0, 0)$ or the drift $\mu = 0$.

In the setting above, we first show that when dealing with $\mathbb{E}[g(t)]$, the stochastic equations for trading can be reduced to a classical second order differential equation. Hence, in addition to positivity of the expectation, many aspects of the behavior of $\mathbb{E}[g(t)] > 0$ such as damping, overshoot, oscillations and asymptotic convergence become straightforward to study. Finally, to demonstrate the application of the main ideas in this chapter, using historical data, we provide simulations to show how the controller performs in a real market when trades are conducted every two minutes.

4.1 Derivation of Dynamics for Expectation

We now consider an idealized market with prices $p(t)$ generated by the Geometric Brownian Motion (GBM) satisfying the stochastic differential equation

$$\frac{dp}{p} = \mu dt + \sigma dZ$$

where $Z(t)$ is a standard Wiener process, μ is the drift and $\sigma \geq 0$ is the volatility.

4.1.1 Interaction of Prices and PI Controller: It is important to note that neither μ nor σ is known to the trader. Furthermore, as previously stated, the feedback control law underlying the investment $I(t)$ is model-free in the sense that no attempt is made to estimate μ and σ on the fly. In order to differentiate between long and short positions, we use subscripts "L" and "S" for the investment function $I(t)$ and trading gain $g(t)$. Accordingly, with initial condition given by $I_L(0) = I_0 > 0$, PI trader who is initially long works with investment

$$I_L(t) = I_0 + K_P g_L(t) + K_I \int_0^t g_L(\tau) d\tau$$

where $K_P \geq 0$ and $K_I \geq 0$. Similarly, on the short side, the trader begins with $I_S(0) = -I_0$ and investment

$$I_S(t) = -I_0 - K_P g_S(t) - K_I \int_0^t g_S(\tau) d\tau.$$

The theory to follow allows us to consider three scenarios: Initially Long, Initially Short and *Initially Long-Short* (ILS). Our use of the word "initially" in describing these trades is based on a fundamental difference between static versus dynamic trading. That is, in the static case with $K_I = 0$, as seen in earlier work [4] and described in Chapter 1, the signs of $I_L(t)$ and $I_S(t)$ remain invariant over the course of the trade. However, when integrator action is included with $K_I \neq 0$, one can end up with either $I_L(t)$ or $I_S(t)$ changing sign; e.g., the initially long position can be "morphed" into a short. To conclude, in the ILS case, the control, being the sum of long and short positions, reduces to

$$I(t) = K_P(g_L(t) - g_S(t)) + K_I \int_0^t (g_L(\tau) - g_S(\tau)) d\tau.$$

Next, we consider the stochastic differential equation for the trading gain g_L noting that a nearly identical analysis applies to g_S. Indeed, noting that the incremental trading gain or loss dg_L is the percentage change in price dp/p times the amount invested $I_L(t)$, we have

$$dg_L = \frac{dp}{p} I_L = (\mu dt + \sigma dZ)\left(I_0 + K_P g_L(t) + K_I \int_0^t g_L(\tau)d\tau\right).$$

4.1.2 State Space Representation: We now create a state-space representation using the two-dimensional state vector

$$x(t) \doteq \begin{bmatrix} \int_0^t g_L(\tau)d\tau \\ g_L(t) \end{bmatrix},$$

with initial condition $x(0) = 0$. Next, we reduce the gain dynamics to the first order stochastic equation

$$\begin{aligned} dx_1 &= x_2 dt; \\ dx_2 &= (\mu dt + \sigma dZ)\left(I_0 + K_P x_2 + K_I x_1\right). \end{aligned}$$

To more clearly see the structure of these equations, we view the initial investment as a unit step input $u(t) \equiv I_0$ for $t \geq 0$ and express the increment above in the classical form

$$dx = (Ax + bu)dt + (Cx + du)dZ$$

with matrices having the structure

$$A \doteq \begin{bmatrix} 0 & 1 \\ \mu K_I & \mu K_P \end{bmatrix}; \quad b \doteq \begin{bmatrix} 0 \\ \mu \end{bmatrix};$$

$$C \doteq \begin{bmatrix} 0 & 0 \\ \sigma K_I & \sigma K_P \end{bmatrix}; \quad d \doteq \begin{bmatrix} 0 \\ \sigma \end{bmatrix}.$$

This is a linear system with multiplicative noise; for example, see [112] and [113]. Although a closed-form solution is generally not available, we see below that a tractable differential equation describing the expectation of the state can nevertheless be obtained.

4.1.3 Gain Expectation Dynamics: Denoting the unconditional expectation of the state, $x(t)$, at time t as

$$\bar{x}(t) \doteq \mathbb{E}[x(t)],$$

the differential equation for the expected-state dynamics is obtained by taking the expectation of both sides of the stochastic differential equation for $x(t)$. To accomplish this, a formal argument requires commutation of derivative and expectation operations above, for example, see [88]. Exploiting the zero-mean property of the Wiener process to eliminate the term multiplied by dZ, we arrive at

$$\begin{aligned} \frac{d\bar{x}}{dt} &= A\bar{x} + bu \\ &= \begin{bmatrix} 0 & 1 \\ \mu K_I & \mu K_P \end{bmatrix} \bar{x}(t) + \begin{bmatrix} 0 \\ \mu \end{bmatrix} I_0 \end{aligned}$$

with output of interest $y = \mathbb{E}[g_L(t)]$ given by

$$y = c^T \bar{x} = [0 \ \ 1]\bar{x}.$$

The system above is now straightforward to analyze using the classical analysis for second-order linear time-invariant systems. To this end, the transfer function from the investment to the trading gain is immediately calculated to be

$$H(s) = c^T(sI - A)^{-1}b = \frac{\mu s}{s^2 - \mu K_P s - \mu K_I},$$

with associated eigenvalues

$$\lambda_\pm = \frac{\mu K_P \pm \sqrt{\mu^2 K_P^2 + 4\mu K_I}}{2}.$$

The simplicity of the transfer function above makes it possible to analyze various scenarios. For example, oscillation and damping are readily studied. As a second example, imagine a trader who is initially long with $\mu < 0$, $K_I > 0$ and $K_P > 0$, then despite being on the "wrong side of the market," it is easy to verify using the Final Value Theorem that $lim_{t\to\infty}\mathbb{E}[g_L(t)] = 0$. That is, with the integrator, an initially "bad trade" eventually turns into a break-even situation.

4.2 Closed-Form Solution Possibilities

For the second order system above, the classical solution possibilities are readily enumerated and closed-form solutions for $\mathbb{E}[g_L(t)]$ can easily be obtained by considering two possibilities, $\mu > 0$ or $\mu < 0$ for the sign of the drift and for the discriminant, $\Delta \doteq \mu^2 K_P^2 + 4\mu K_I$, there are three cases: $\Delta < 0, \Delta = 0$ and $\Delta > 0$. We now demonstrate by showing the solution for two of the most important cases and then provide a convenient compact formula which covers all cases in one fell swoop.

4.2.1 The Oscillatory Case: Suppose $\mu < 0$, $\Delta < 0$, $K_I > 0$ and $K_P \geq 0$. Then using the formulae above, the expected value of the trading gain is a damped harmonic; i.e., the damping ratio $\zeta < 1$ is given by

$$\zeta = \frac{1}{2} K_P \sqrt{\frac{|\mu|}{K_I}},$$

the undamped natural frequency is $\omega_n = \sqrt{|\mu| K_I}$, and, inverting the Laplace transform, we easily obtain

$$\begin{aligned}
\mathbb{E}[g_L(t)] &= -\left(\frac{I_0}{K_I}\right) \frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin(\omega_n \sqrt{1-\zeta^2}\ \ t) \\
&= \frac{2\mu I_0}{\sqrt{4|\mu| K_I - \mu^2 K_P^2}} e^{\mu K_P t/2} \sin\left(\sqrt{(4|\mu| K_I - \mu^2 K_P^2)}\ \ \frac{t}{2}\right).
\end{aligned}$$

Perhaps the most important feature of the solution above is that it adapts to the trader's "error" in the assessment of the market's direction. By this, we mean the following: Suppose $K_P > 0$ and the trader begins with a long position $I_L(0) = I_0 > 0$ in a market which is drifting downward with $\mu < 0$. As losses build up, the integration action eventually forces $I_L(t) < 0$. That is, the trader finally "gets it right" in a falling market by switching from a long to a short position. To see this effect more clearly, we calculate the expected value of the investment, and, via a straightforward calculation, we obtain

$$\mathbb{E}[I_L(t)] = \frac{I_0}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \cos\left(\omega_n \sqrt{1-\zeta^2} t + \theta\right)$$

where

$$\theta \doteq \arctan\left(\frac{-\zeta}{\sqrt{1-\zeta^2}}\right).$$

The adaptation phenomenon described above is now seen in Figure 4.2.1 where $\mathbb{E}[g_L(t)]$ and $\mathbb{E}[I_L(t)]$ are plotted using sample parameter values $I_0 = 1,\ \mu = -3,\ K_P = .5$ and $K_I = 4$. A key observation is that three times over the duration of the trade, $I_L(t)$ switches from long to short and eventually, per discussion of the Final Value Theorem above, turns a losing trade into a break-even situation.

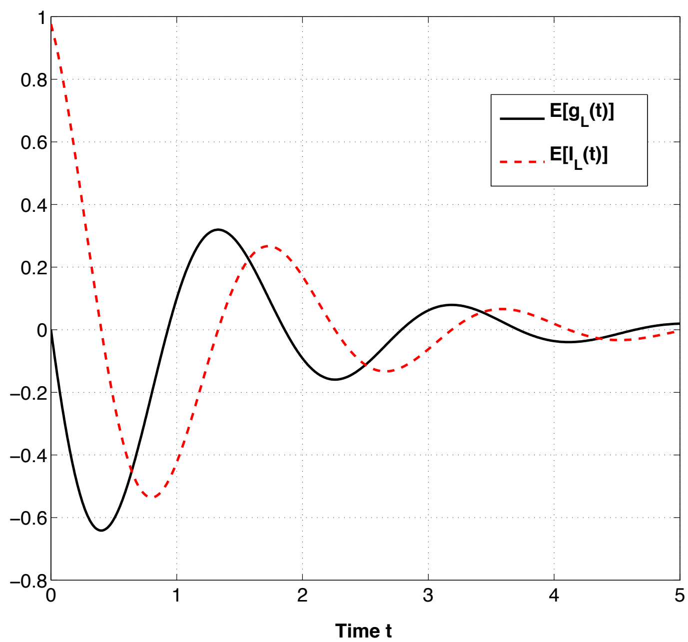

Figure 4.2.1: Trading Gain and Investment for the Oscillatory Case

4.2.2 The Purely Exponential Case: Continuing with the initially-long trader in a falling market with $\mu < 0$, when the discriminant Δ is positive, that is, $\mu^2 K_P^2 + 4\mu K_I > 0$, we obtain the solution

$$\mathbb{E}[g_L(t)] = \frac{\mu I_0}{2\alpha} e^{\mu K_P t/2}\left[e^{\alpha t} - e^{-\alpha t}\right]$$

where

$$\alpha \doteq \frac{\sqrt{\mu^2 K_P^2 + 4\mu K_I}}{2}.$$

We note that the initially-long trader is on the wrong side of the market but the controller adapts by switching from long to short as losses increase and ultimately breaks even.

4.2.3 Convenient Compact Solution Representation: The two cases considered above do not cover all the solution possibilities. Rather than enumerating them all, we simply provide a compact formula which covers all cases. Below, the understanding is that arguments under the square root sign can be negative. In such cases, using de Moivre's formula, we obtain the appropriate interpretation in terms of harmonics. Furthermore, since it is easily verified that the expected gain or loss is an even function of the drift μ, the formulae below are given for $\mu \geq 0$. Accordingly, for the case of PI control, we obtain

$$\mathbb{E}[g_L(t)] = \frac{\mu I_0}{\alpha} e^{\mu K_P t/2} \sinh(\alpha t);$$

$$\mathbb{E}[g_S(t)] = \frac{-\mu I_0}{\beta} e^{-\mu K_P t/2} \sinh(\beta t)$$

where

$$\beta \doteq \frac{\sqrt{\mu^2 K_P^2 - 4\mu K_I}}{2}.$$

4.2.4 Initially Long-Short (ILS) Controller: Motivated by results in the purely proportional feedback case, see Section 1.4.3 and [3] and [4], we consider a trading strategy that implements both a long and short version of the PI strategy, simultaneously. The expected gain from this strategy is simply the sum of the gain from the long and short strategies which is found to be

$$\mathbb{E}[g(t)] = \mu I_0 e^{-\mu K_P t/2} \left[e^{\mu K_P t} \frac{\sinh(\alpha t)}{\alpha} - \frac{\sinh(\beta t)}{\beta} \right].$$

Recalling the examples in the previous section, we know that a long controller which begins with $I_L(0) = I_0 > 0$, can eventually become a short over the course of the trade. Similarly, the short side can switch to long. Hence, we refer to this as an *Initially Long-Short (ILS)* PI controller. In earlier work with a static controller corresponding to the special case $K_I = 0$, see [4], this switch between long and short could not occur. Hence, this reference uses the terminology *Simultaneous Long-Short* (SLS) to describe the controller.

As previewed in Chapter 1 and established in [4], the statis SLS feedback control strategy achieves a positive expected trading gain under any GBM price process with $K\mu \neq 0$. It is reasonable to ask whether this same robustness property holds in the ILS case for the PI controller. We can see immediately the potential for such a result in the ILS case by setting $K_P = 0$. This is the so-called the pure integrator case and we see that the compact solution representation for this special case is

$$\mathbb{E}(g(t)) = \sqrt{\frac{|\mu|}{K_I}} I_0 \Big[\sinh(\sqrt{|\mu| K_I} t) - \sin(\sqrt{|\mu| K_I} t) \Big]$$

which is readily verified to be positive except for the trivial break-even case when $\mu = 0$ or $K_I = 0$. In the section to follow, we establish that this Robust Positive Expectation Property holds in the more general case when both K_P and K_I are in play. In addition we establish the monotonic dependence of the expected value of $g(t)$ on K_P and K_I.

4.3 Robust Positive Expectation Property for PI Controller

For the Initially Long-Short PI controller above, we now provide two theorems.

4.3.1 Theorem: (Robust Positive Expectation Property) *Consider the ILS PI controller with $K_I \geq 0$ and $K_P \geq 0$ in an idealized market with GBM prices. Then, except for the trivial break-even case obtained when either $\mu = 0$ or $(K_P, K_I) = (0, 0)$, the expected gain $\mathbb{E}[g(t)]$ is strictly increasing in t. Moreover, since $\mathbb{E}(g(0)) = 0$, it follows that*

$$\mathbb{E}(g(t)) > 0$$

for all $t > 0$.

Proof: To establish that the expected return is increasing with respect to t, we use the compact solution representation given in Section 4.2.4. Additionally, we use the notation,

$$f(x) \doteq \frac{\sinh(x)}{x}.$$

Recalling that the expected gain is an even function of μ, without loss of generality we assume that $\mu > 0$. To prove that the expected gain is increasing in time, we will show that its derivative is strictly positive. Indeed, differentiating with respect to time t and rearranging terms gives

$$\begin{aligned} \frac{d\mathbb{E}(g(t))}{dt} &= \frac{\mu^2 K_P I_0 t e^{-\mu K_P t/2}}{2} \left[e^{\mu K_P t} f(\alpha t) + f(\beta t) \right] \\ &\quad + \mu I_0 e^{-\mu K_P t/2} \left[e^{\mu K_P t} \cosh(\alpha t) - \cosh(\beta t) \right] \end{aligned}$$

where α and β are defined in the previous section. To show that this quantity is strictly positive, we consider the following two cases.

Case 1: β is real, that is $\mu^2 K_P^2 - 4\mu K_I \geq 0$. Then using the following facts proves that the derivative is positive:

1. $f(x) \geq 0$ for $x \geq 0$.

2. $e^{\mu K_P t} \geq 1$.

3. $\alpha > \beta$ and $\cosh$ is an increasing function.

Case 2: β is pure imaginary, that is $\mu^2 K_P^2 - 4\mu K_I < 0$. Then $\beta = \gamma j$ where $\gamma = \frac{\sqrt{4\mu K_I - \mu^2 K_P^2}}{2}$. Rewriting the derivative,

$$\begin{aligned}\frac{d\mathbb{E}(g(t))}{dt} &= \frac{\mu^2 K_P I_0 t e^{-\mu K_P t/2}}{2}\left[e^{\mu K_P t} f(\alpha t) + \text{sinc}(\gamma t)\right] \\ &\quad + \mu I_0 e^{-\mu K_P t/2}\left[e^{\mu K_P t}\cosh(\alpha t) - \cos(\gamma t)\right].\end{aligned}$$

Again the following facts prove the time derivative of the expected return to be strictly positive in this case:

1. $e^{\mu K_P t} > 1$.

2. $f(x) \geq 1$ for $x > 0$.

3. $|\text{sinc}(x)| \leq 1$ for all x.

4. $\cosh(x) \geq 1$ for all x.

5. $|\cos(x)| < 1$ for all x.

Since these two cases are mutually exclusive and exhaustive, this completes the proof. □

4.3.2 Theorem: (Monotonicity in the Control Parameters) *Consider the ILS PI controller with $K_I \geq 0$ and $K_P \geq 0$ in an idealized market with GBM prices. Then, except for the trivial break-even case of $\mu = 0$, for $t > 0$, the expected return $\mathbb{E}[g(t)]$ is increasing in K_P and in K_I.*

Proof: We begin from the convenient compact solution provided in the previous section. Rearranging, we obtain

$$\mathbb{E}\left(g(t)\right) = \mu I_0 t \left[e^{\mu K_P t/2} f(\alpha t) - e^{-\mu K_P t/2} f(\beta t)\right].$$

Recalling that the expected gain is an even function of μ, without loss of generality we assume $\mu > 0$. Consider the function $f(x)$ introduced in the proof of the previous theorem. Now writing the Taylor expansion for this function, we obtain

$$f(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n+1)!}; \quad f'(x) = \sum_{n=1}^{\infty} \frac{2nx^{2n-1}}{(2n+1)!}.$$

Taking the derivative of the expected gain and replacing $f(x)$ and $f'(x)$ by the expressions above, after further simplification and reordering of terms, we obtain

$$\begin{aligned} \frac{\partial \mathbb{E}(g)}{\partial K_I} &= \mu^2 I_0 t^3 e^{-\mu K_P t/2} \\ &\quad \times \sum_{n=1}^{\infty} \frac{n}{(2n+1)!} \left[e^{\mu K_P t} (\alpha t)^{2n-2} + (\beta t)^{2n-2})\right]. \end{aligned}$$

Noting that $\alpha^2 \geq |\beta^2|$ and the fact that $e^{\mu K_P t} > 1$, one may verify that each term in the sum is positive, and therefore, the derivative is positive. Similarly, for K_P, we obtain

$$\begin{aligned} \frac{\partial \mathbb{E}(g)}{\partial K_P} &= \frac{\mu^2 I_0 t^2}{2} e^{-\mu K_P t/2} \\ &\quad \times \left\{ \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \left[e^{\mu K_P t} (\alpha t)^{2n} + (\beta t)^{2n}\right] \right\} \\ &\quad + \frac{\mu^3 I_0 K_P t^3}{2} e^{-\mu K_P t/2} \\ &\quad \times \left\{ \sum_{n=1}^{\infty} \frac{n}{(2n+1)!} \left[e^{\mu K_P t} (\alpha t)^{2n-2} - (\beta t)^{2n-2})\right] \right\}. \end{aligned}$$

Again, using the same reasoning, each term in each of the sums is positive and therefore the derivative is positive. □

4.3.3 Remark: Note that while the theorem shows that the expected gain increases with increased parameter values, it does not address the issue of the variance of the gain or the increased risk associated with larger parameter values. Thus, in the next section, we provide an analysis of the covariance matrix that captures the risk and correlations between the long, short, and Initially Long-Short PI strategies.

4.4 Analysis of the Covariance Matrix

In this section, we analyze the variance and covariance of the ILS PI strategy. To begin, we combine the states for the initially-long and initially-short cases into

$$x \doteq \begin{bmatrix} x_L \\ x_S \end{bmatrix}$$

and write the combined dynamics for x as

$$dx = (Ax + bu)\,dt + (Cx + du)\,dZ$$

with

$$A \doteq \begin{bmatrix} 0 & 1 & 0 & 0 \\ \mu K_I & \mu K_P & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\mu K_I & -\mu K_P \end{bmatrix}; \quad b \doteq \begin{bmatrix} 0 \\ \mu \\ 0 \\ -\mu \end{bmatrix};$$

$$C \doteq \begin{bmatrix} 0 & 0 & 0 & 0 \\ \sigma K_I & \sigma K_P & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\sigma K_I & -\sigma K_P \end{bmatrix}; \quad d \doteq \begin{bmatrix} 0 \\ \sigma \\ 0 \\ -\sigma \end{bmatrix}.$$

4.4.1 Differential Equation for the Covariance Matrix: To analyze the covariance matrix of the state x, we take

$$\bar{x}(t) \doteq \mathbb{E}[x(t)]$$

and study the error $e \doteq x - \bar{x}$ and its covariance

$$P(t) \doteq \mathbb{E}[e(t)e^T(t)].$$

Now, similar to the previous initially-long analysis, a straightforward calculation leads to the stochastic increment for the error

$$de = Aedt + (Ce + C\bar{x} + du)dZ.$$

By Ito's lemma, we can calculate dee^T, which gives

$$\begin{aligned} dee^T &= Aee^T dt + ee^T A^T dt \\ &\quad +(Ce + C\bar{x} + du)(Ce + C\bar{x} + du)^T dt \\ &\quad +((Ce + C\bar{x} + du)e^T + e(Ce + C\bar{x} + du)^T)dZ. \end{aligned}$$

Taking the expectation of both sides leads to

$$\frac{dP}{dt} = AP + PA^T + CPC^T + (C\bar{x} + du)(C\bar{x} + du)^T.$$

This is a Lyapunov-type linear matrix differential equation in $P(t)$ and thus can be easily solved. In particular, we are interested in the variance of the individual components $g_L(t)$ and $g_S(t)$, P_{22} and P_{44} respectively, of the ILS strategy. In addition, using $P(t)$ we obtain the variance of the overall ILS trading gain $g(t) = g_L(t) + g_S(t)$ as

$$\text{var}[g(t)] = h^T P(t)h \text{ with } \; h \doteq [0, 1, 0, 1]^T.$$

The calculation of the variance, as prescribed above, is straightforward to implement numerically. Given the emphasis here on the Robust Positive Expectation Property, we do not provide an illustrative plot but make note of the fact that variance of $g(t)$ for the ILS strategy is lower than the sum of the variances of g_L and g_S. This is due to the fact that g_L and g_S are negatively correlated.

4.5 A Simulation Using Historical Data

In this section, we illustrate the application of the ILS PI feedback control strategy using a data set consisting of about $35,000$ price quotes for Apple (Ticker: AAPL) stock. The data covers 175 days from June 18, 2012 to February 28, 2013 with each new price quote following its predecessor by about two minutes.[2] In Figure 4.5.1, the daily closing prices are plotted.

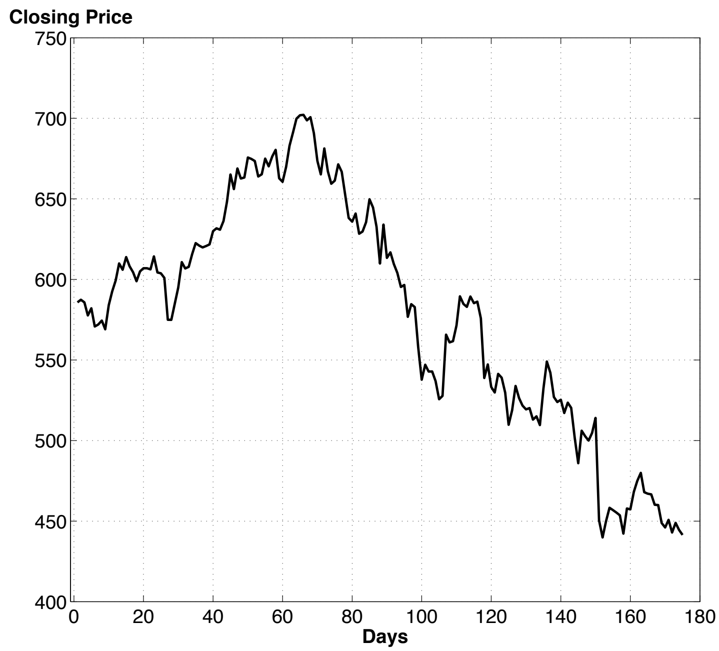

Figure 4.5.1: Closing Prices for AAPL

To more closely approximate the idealized market conditions underlying the theory in this chapter, we consider the case when the ILS position is initiated at the market open and vacated at the close; i.e., the trader is entirely in cash overnight. This increases the likelihood that price paths are more nearly continuous than would be the case if a position is carried overnight. Our interpretation is that each day provides a 200-point sample path from some random process governing the price of Apple. Figure 4.5.2 provides typical daily sample paths with the opening price normalized to $\$1$.

The first scenario we consider is the following: The investor begins with initial account value of $\$10,000$. Then, each day, the ILS position is initiated with $I_0 = 10,000$. We further assume no transaction costs and the ability to trade about every two minutes as a new data point arrives. In the first set of simulations, we held the proportional gain $K_P = 2$ fixed and varied the integrator gain

[2]The authors express their thanks to Mr. Amin Farmahini-Farahani for help with the acquisition and processing of intraday stock prices.

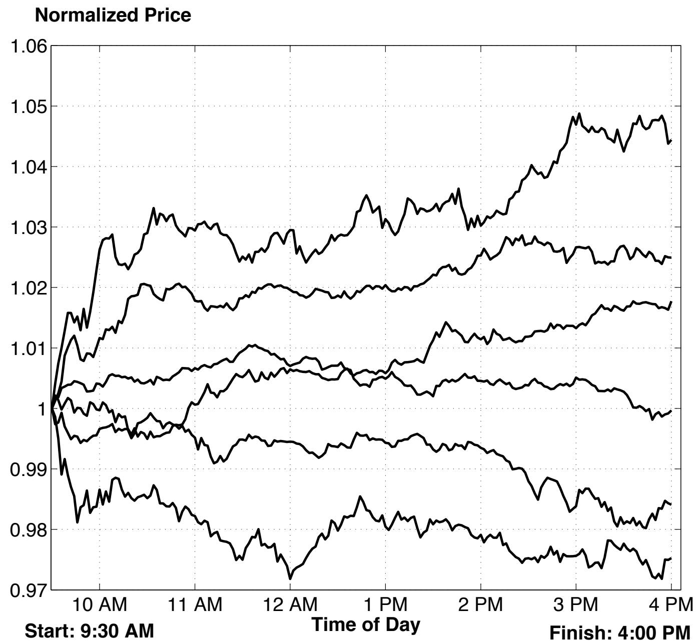

Figure 4.5.2: Sample Paths for AAPL

K_I	Ending Account Value	Average Investment
0	$10,098	$335
0.2	$11,950	$3,005
0.5	$14,726	$7,025
1	$19,357	$13,630
2	$28,686	$26,406

Table 4.1: The Ending Account Value and Average Investment with $K_p = 2$

in the interval $0 \leq K_I \leq 2$. The results are given in the table below. In Figure 4.5.3, the evolution of $g(t)$ is shown for $K_I = 2$.

Looking at the investment levels in Table 4.1, we raise the possibility of a difference which might arise between theory and practice. The trader may not have adequate cash reserves to fund the required investment level $I(t)$. For example, the gain $K_I = 2$ leads to an average daily investment of over $26,000. If the account value is much less than this investment level, there would be a requirement to pay margin interest or worse yet, a margin call by the broker could occur if the account is not adequately collateralized.

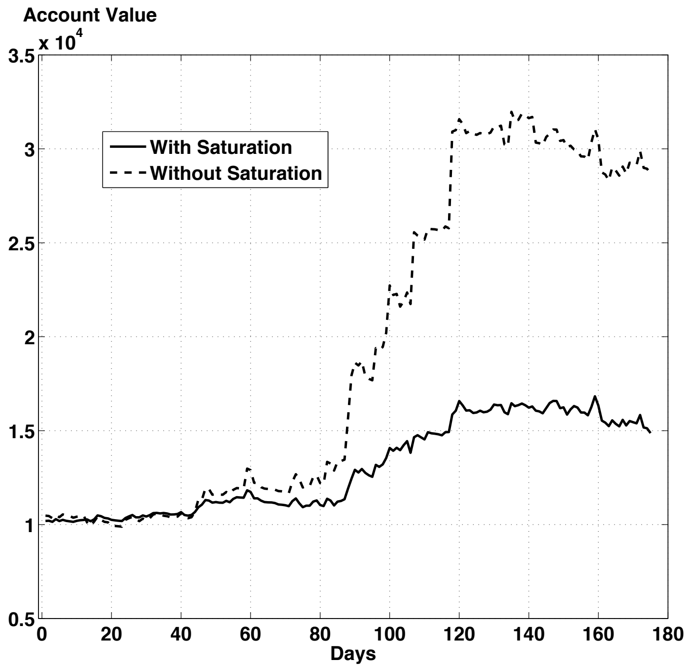

Figure 4.5.3: Cumulative Return for PI Controller

In view of these practical considerations, we also considered the case $K_I = K_P = 2$ taking the account value (wealth) of the trader into account in restricting the admissible investment level. If we denote the account value of the trader at time t by $V(t)$, a simple but practical constraint would be the following: Recalling that $I(t) = I_L(t) + I_S(t)$ is the net investment, we impose the constraint $|I(t)| \leq 2V(t)$. In other words, if the ILS control algorithm dictates an investment level no larger than $2V(t)$, we use the saturated control $I_{sat}(t) \doteq 2V(t)\text{sign } I(t)$. In Figure 4.5.3, with saturation included, the trading gain is seen to evolve more modestly versus the unconstrained case.

4.6 PI Controller with Exponentially-Weighted Moving Average

The simulations in the previous section demonstrate the possibility that the ILS investment levels can become "too large." From a practical implementation point of view, the broker's collateral requirements may restrict investment levels along the lines described in Subsection 3.0.2. This effect is closely related to what is known as *integral windup* problem, see [114] for details. Motivated by these practical considerations, in this section we provide an alternative discounting scheme. To this end, an immediate direction involves generalizing the results to allow for more emphasis on recent performance. In this regard, it is of interest to consider the so-called *exponentially weighted moving average* with control

$$I(t) = I_0 + K_P g(t) + K_I \int_0^t e^{-\gamma(t-\tau)} g(\tau) d\tau$$

where $\gamma \geq 0$ is chosen by the trader. Note that by choosing $\gamma = 0$, we recover integral feedback. The size of $\gamma \geq 0$ affects the rate at which the past performance is discounted.

The analysis for this more general case begins in much the same way as the $\gamma = 0$ case. That is, we start with the state-space vector x defined as

$$x(t) \doteq \begin{bmatrix} \int_0^t e^{-\gamma(t-\tau)} g(\tau) d\tau \\ g(t) \end{bmatrix},$$

and use the GBM price dynamics to develop a differential equation for $\bar{x} = \mathbb{E}[x(t)]$. We further combine two components, one long and one short, to obtain the *initially long-short* controller with exponentially weighted moving average. We show that with this new controller, the Robust Positive Expectation Property holds for a non-trivial GBM price process with the drift $\mu \neq 0$ and with controller gains $(K_P, K_I) \neq (0, 0)$.

4.6.1 The Feedback Controller and Overview of Main Result: Using the investment rule, $I(t)$ given in preceding section, we can extend it to define "long" and "short" components. To this end, the long and short investment at time t are given by

$$\begin{aligned} I_L(t) &= I_0 + K_p g_L(t) + K_I \int_0^t e^{-\gamma(t-\tau)} g_L(\tau) d\tau; \\ I_S(t) &= -I_0 - K_p g_S(t) - K_I \int_0^t e^{-\gamma(t-\tau)} g_S(\tau) d\tau, \end{aligned}$$

respectively. In the remainder of this chapter we analyze the expected value of the trading gain-loss function $g(t)$, resulting from the use of this new controller. More specifically, we start with the initially long component, $g_L(t)$, and use state-space model under GBM to obtain $\mathbb{E}[g_L(t)]$. Then, we describe how this analysis can be modified to obtain $\mathbb{E}[g_S(t)]$. Finally, writing the total gain-loss function as the sum of the two components; i.e., $g(t) = g_L(t) + g_S(t)$, we prove that the Robust Positive Expectation Property still holds; that is,

$$\mathbb{E}[g(t)] = \mathbb{E}[g_L(t) + g_S(t)] > 0$$

as long as $\mu \neq 0$ and $(K_P, K_I) \neq 0$.

4.7 Derivation of Dynamics for Expectation

The analysis provided here is carried out for the *long* component $I_L(t)$ while noting that a minor change in signs leads to the result for the *short* component $I_S(t)$. The starting point is the definition of the state as $x(t)$ given in the preceding section with initial condition $x(0) = 0$. Assuming Geometric Brownian Motion process for the price with drift μ and volatility σ; we reduce the gain dynamics to the first order stochastic equation

$$\begin{aligned} dx_1 &= (-\gamma x_1 + x_2)dt; \\ dx_2 &= (\mu dt + \sigma dZ)\left(I_0 + K_P x_2 + K_I x_1\right). \end{aligned}$$

To simplify dx_1, we use Leibniz rule to obtain

$$\begin{aligned} dx_1 &= d\big(e^{-\gamma t}\int_0^t e^{\gamma\tau} g_L(\tau)d\tau\big) \\ &= -\gamma e^{-\gamma t}dt(\int_0^t e^{\gamma\tau} g_L(\tau)d\tau) + e^{-\gamma t}e^{\gamma t} g_L(t)dt \\ &= (-\gamma x_1 - x_2)dt. \end{aligned}$$

In computing dx_1 above, we used the fact that differentiation involving the simple Lebesgue integral does not require any Ito correction. Now, with the initial investment as a unit step input $u(t) \equiv I_0$, for $t \geq 0$ we express the increment above as

$$dx = (A_\gamma x + bu)dt + (Cx + du)dZ$$

where

$$A_\gamma \doteq \begin{bmatrix} -\gamma & 1 \\ \mu K_I & \mu K_P \end{bmatrix}; \;\; b \doteq \begin{bmatrix} 0 \\ \mu \end{bmatrix};$$

$$C \doteq \begin{bmatrix} 0 & 0 \\ \sigma K_I & \sigma K_P \end{bmatrix}; \;\; d \doteq \begin{bmatrix} 0 \\ \sigma \end{bmatrix}.$$

Similar to the case of PI controller, we can now readily obtain the differential equation describing the expectation of the state.

4.7.1 Gain Expectation Dynamics: Following an approach similar to Section 4.1 and noting that the term $(Cx + du)dZ$ has no effect on the expectation dynamics, we obtain

$$\dot{\bar{x}} = A_\gamma \bar{x} + bu$$

with output $y(t) = \mathbb{E}[g_L(t)]$ given by

$$y = c^T \bar{x} = [0 \;\; 1]\bar{x}.$$

The second-order system above is now straightforward to analyze using the classical analysis. Indeed, the transfer function from the investment to the trading gain is immediately calculated to be

$$H(s) = c^T(sI - A_\gamma)^{-1}b = \frac{\mu(s+\gamma)}{s^2 + (\gamma - \mu K_P)s - \mu(K_I + \gamma K_P)},$$

with associated eigenvalues

$$\lambda_\pm = \frac{-(\gamma - \mu K_P) \pm \sqrt{(\gamma - \mu K_P)^2 + 4(\mu K_I + \mu K_P \gamma)}}{2}$$

Now with the unit step input which is $u = I_0/s$ in frequency domain, we obtain

$$G_L(s) \doteq [0 \;\; 1](sI - A_\gamma)^{-1}b = \frac{\mu(s+\gamma)I_0/s}{s^2 + (\gamma - \mu K_P)s - \mu(K_I + \gamma K_P)}.$$

where $G_L(s) \doteq \mathcal{L}\{\mathbb{E}[g_L(t)]\}$ is the Laplace transform of $g_L(t)$.

4.8 Closed-Form Solution Possibilities

A calculation similar to what we carried out in Section 4.2 leads to solutions for expected value of gain-loss function, $\mathbb{E}[g_L(t)]$ and expected investment level, $\mathbb{E}[I_L(t)]$ for different scenarios. In the sequel, we provide the calculations for the oscillatory case with the understanding that the other cases are simple extensions of earlier results. The oscillatory case we consider occurs when

$$\Delta \doteq (\mu K_p + \gamma)^2 + 4\mu K_I < 0.$$

Once this case is studied, we give a general closed-form solution which covers all the scenarios.

4.8.1 The Oscillatory Case: Suppose $\mu < 0$, $\Delta < 0$, $\gamma > 0$, $K_I > 0$ and $K_P \geq 0$. Then taking the inverse Laplace transform of $G_L(s)$ given above, we obtain the expected value of the trading gain as the damped harmonic

$$\mathbb{E}[g_L(t)] = -\frac{\gamma I_0}{K_I + \gamma K_P} - \frac{\mu r I_0 e^{\lambda t}}{\omega}\, cos(\omega t + \theta)$$

where

$$\omega \doteq \frac{\sqrt{4|\mu|K_I - (\mu K_P + \gamma)^2}}{2}; \qquad \lambda \doteq \frac{\mu K_P - \gamma}{2};$$

$$r \doteq \sqrt{\frac{K_I}{\gamma K_P + K_I}}; \qquad \theta \doteq \arctan(\frac{\lambda^2 + w^2 + \lambda\gamma}{w\gamma}).$$

Perhaps the most important feature of the solution above is that it adapts to the trader's "error" in the assessment of the market's direction. By this, we mean the following: The trader begins with a long position $I_L(0) = I_0 > 0$ in a market which is drifting downward with $\mu < 0$. As losses build up, the integration action eventually forces $I_L(t)$ to become negative after being initially positive. That is, the trader finally "gets it right" in a falling market by switching from a long to a short position. To see this effect more clearly, we calculate the expected value of the investment, and, via a straightforward calculation, we obtain

$$\mathbb{E}[I_L(t)] = \frac{I_0}{\omega}\, r_2 e^{\lambda t}\, \cos(\omega t + \theta_2)$$

where

$$r_2 \doteq \sqrt{|\mu|K_I}; \qquad \theta_2 \doteq -\arctan(\frac{\mu K_P + \gamma}{2\omega}).$$

Moreover, since $\lambda < 0$ we obtain the asymptotic quantities

$$\lim_{t\to\infty} \mathbb{E}[g_L(t)] = -\frac{\gamma I_0}{K_I + \gamma K_P}; \qquad \lim_{t\to\infty} \mathbb{E}[I_L(t)] = 0.$$

The adaptation phenomenon described above is now seen in Figure 4.8.1 where $\mathbb{E}[g_L(t)]$ and $\mathbb{E}[I_L(t)]$ are plotted using sample parameter values $I_0 = 1$, $\mu = -3$, $K_P = 0.5$, $K_I = 4$ and $\gamma = 0.25$. A key observation is that four times over the duration of the trade, $I_L(t)$ switches from long to short and eventually, per discussion of the Final Value Theorem, the "initially long" investment component vanishes.

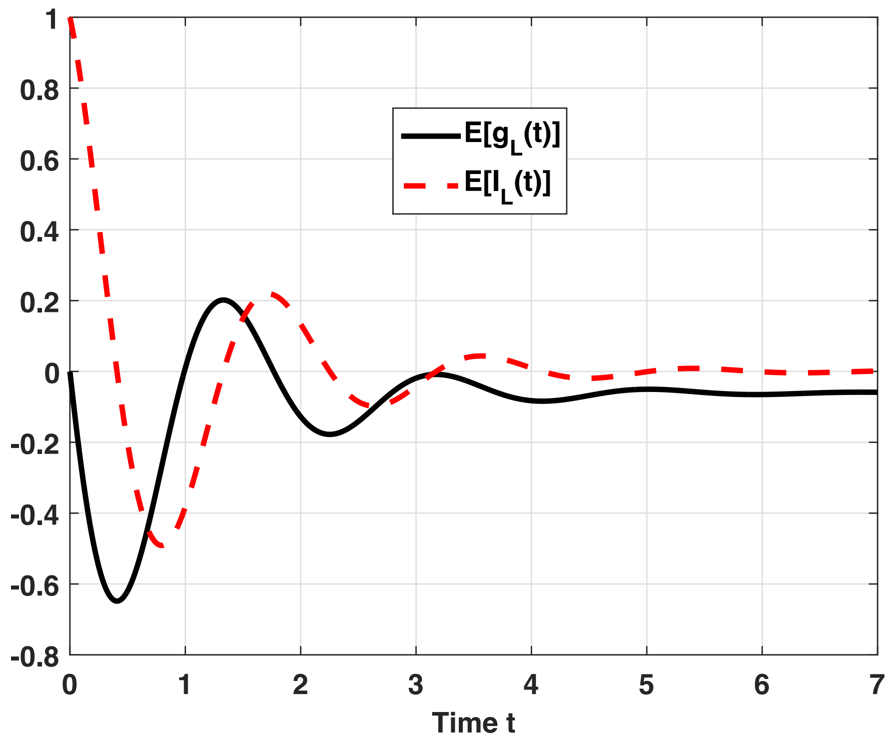

Figure 4.8.1: Trading Gain and Investment for the Oscillatory Case

4.8.2 Closed-Form Solution For $\mathbb{E}[g(t)]$: In this subsection, we provide a closed-form expression for the expected gain-loss function, $\mathbb{E}[g(t)]$ when a combination of long and short investment are used. That is with the setting similar to 4.2.4, we obtain the closed-form solution for

$$\mathbb{E}[g(t)] \doteq \mathbb{E}[g_L(t) + g_S(t)].$$

To this end, we first find the Laplace transform for the expected gain-loss function for the short component $\mathbb{E}[g_S(t)]$, which is

$$G_S(s) = \frac{-\mu(s+\gamma)I_0/s}{s^2 + (\gamma + \mu K_P)s + \mu(K_I + \gamma K_P)}.$$

Note that, for the special case of $\gamma = 0$, we recover the expression for the PI controller as described in Section 4.2. Similarly, we can find the Laplace transform for the expected gain-loss function for the long component $\mathbb{E}[g_L(t)]$.

Then taking the inverse Laplace transform of $G(s) \doteq G_L(s) + G_S(s)$ and further simplification leads to

$$\begin{aligned}\mathbb{E}[g(t)] &= \mu I_0\Big[e^{\lambda t}\,\frac{\sinh \alpha t}{\alpha} - e^{\eta t}\,\frac{\sinh \beta t}{\beta}\Big]\\ &\quad +\gamma\mu I_0\left[\frac{e^{\frac{\lambda+\alpha}{2}t}}{\alpha(\lambda+\alpha)}\sinh\left(\frac{\lambda+\alpha}{2}t\right) - \frac{e^{\frac{\eta+\beta}{2}t}}{\beta(\eta+\beta)}\sinh\left(\frac{\eta+\beta}{2}t\right)\right]\\ &\quad -\gamma\mu I_0\left[\frac{e^{\frac{\lambda-\alpha}{2}t}}{\alpha(\lambda-\alpha)}\sinh\left(\frac{\lambda-\alpha}{2}t\right) - \frac{e^{\frac{\eta-\beta}{2}t}}{\beta(\eta-\beta)}\sinh\left(\frac{\eta-\beta}{2}t\right)\right].\end{aligned}$$

where

$$\alpha \doteq \frac{\sqrt{(\mu K_P + \gamma)^2 + 4\mu K_I}}{2}; \qquad \beta \doteq \frac{\sqrt{(\mu K_P - \gamma)^2 - 4\mu K_I}}{2},$$

and

$$\lambda \doteq \frac{\mu K_P - \gamma}{2}; \qquad \eta \doteq \frac{-\mu K_P - \gamma}{2}.$$

For the special case when $\gamma = 0$, it is easy to show that this formula reduces to the one given in Section 4.2. Similar to the discussion in Section 4.2.3, rather than enumerating all the solution possibilities, we simply provided a compact formula which covers all the cases. The understanding is that arguments α and β can be complex. In such cases, using de Moivre's formula, we obtain the appropriate interpretation of complex hyperbolic functions in terms of harmonics.

4.9 Robust Positive Expectation Property for Exponentially-Weighted Case

In this section, we prove that for the combination of long and short components with exponentially weighting scheme with $\gamma \geq 0$ and $(K_P, K_I) \neq 0$ leads to the Robust Positive Expectation Property when a non-trivial Geometric Brownian Motion with drift $\mu \neq 0$ is the governing price process.

4.9.1 Theorem (Robust Positive Expectation): *Consider the PI controller with exponentially-weighted moving average, with $\gamma > 0$, $K_I \geq 0$ and $K_P \geq 0$ in an idealized market with GBM price process. Then, except for the trivial break-even case obtained when either $\mu = 0$ or $(K_P, K_I) = (0,0)$, the expected gain $\mathbb{E}[g(t)]$ is strictly increasing in t. Moreover, since $\mathbb{E}(g(0)) = 0$, it follows that for $t \geq 0$, we have*

$$\mathbb{E}(g(t)) > 0.$$

Proof: We start of with rewriting the Laplace transform for $\mathbb{E}[g_L(t)]$ as

$$G_L(s) = \underbrace{\frac{\mu I_0}{s^2 + (\gamma - \mu K_P)s - \mu(K_I + \gamma K_P)}}_{\doteq G_L^{(1)}(s)} + \underbrace{\frac{\gamma \mu I_0 / s}{s^2 + (\gamma - \mu K_P)s - \mu(K_I + \gamma K_P)}}_{\doteq G_L^{(2)}(s)}.$$

Similarly for the short component, we have

$$G_S(s) = \underbrace{\frac{-\mu I_0}{s^2 + (\gamma + \mu K_P)s + \mu(K_I + \gamma K_P)}}_{\doteq G_S^{(1)}(s)} + \underbrace{\frac{-\gamma \mu I_0 / s}{s^2 + (\gamma + \mu K_P)s + \mu(K_I + \gamma K_P)}}_{\doteq G_S^{(2)}(s)}.$$

Defining

$$\begin{aligned}
\mathbb{E}[g_L^{(1)}(t)] &\doteq \mathcal{L}^{-1}\{G_L^{(1)}(s)\}; \quad \mathbb{E}[g_L^{(2)}(t)] \doteq \mathcal{L}^{-1}\{G_L^{(2)}(s)\}; \\
\mathbb{E}[g_S^{(1)}(t)] &\doteq \mathcal{L}^{-1}\{G_S^{(1)}(s)\}; \quad \mathbb{E}[g_S^{(2)}(t)] \doteq \mathcal{L}^{-1}\{G_S^{(1)}(s)\},
\end{aligned}$$

it is easy to use the Initial Value Theorem to show that

$$\mathbb{E}[g_L^{(1)}(0)] = \mathbb{E}[g_S^{(1)}(0)] = 0.$$

Then noting that $1/s$ in frequency domain corresponds to integration in time domain leads to

$$\mathbb{E}[g_L^{(2)}(t)] = \gamma \int_0^t \mathbb{E}[g_L^{(1)}(\tau)] d\tau; \qquad \mathbb{E}[g_S^{(2)}(t)] = \gamma \int_0^t \mathbb{E}[g_S^{(1)}(\tau)] d\tau.$$

Now rearranging the terms in $\mathbb{E}[g(t)] = \mathbb{E}[g_L(t)] + \mathbb{E}[g_S(t)]$ leads to

$$\mathbb{E}[g(t)] = \underbrace{\mathbb{E}[g_L^{(1)}(t)] + \mathbb{E}[g_S^{(1)}(t)]}_{I} + \gamma \underbrace{\int_0^t \mathbb{E}[g_L^{(1)}(\tau)] + \mathbb{E}[g_S^{(1)}(\tau)] d\tau}_{II}.$$

With this arrangement, it would suffice to show that $I \doteq \mathbb{E}[g_L^{(1)}(t)] + \mathbb{E}[g_S^{(1)}(t)] > 0$ for all $t > 0$; i.e., I positive forces the integrand in II positive as well. Taking the inverse Laplace transform and using the simplification techniques in Section 4.3 yields

$$\mathbb{E}[g_L^{(1)}(t)] + \mathbb{E}[g_S^{(1)}(t)] = e^{-\gamma t/2} \underbrace{\left(\frac{\mu I_0}{\alpha} e^{\mu K_P t/2} \sinh(\alpha t) - \frac{\mu I_0}{\beta} e^{-\mu K_P t/2} \sinh(\beta t) \right)}_{III}$$

where

$$\alpha \doteq \frac{\sqrt{(\gamma - \mu K_P)^2 + 4\mu(K_I + \gamma K_P)}}{2} = \frac{\sqrt{(\mu K_P + \gamma)^2 + 4\mu K_I}}{2};$$
$$\beta \doteq \frac{\sqrt{(\gamma + \mu K_P)^2 - 4\mu(K_I + \gamma K_P)}}{2} = \frac{\sqrt{(\mu K_P - \gamma)^2 - 4\mu K_I}}{2}.$$

The exponential term $e^{-\gamma t/2}$ is trivially positive. The proof for the positivity of III is a straightforward modification of the proof for the case of $\gamma = 0$ in Section 4.3 with the new α and β given above. □

4.9.2 Asymptotic Behavior: In this subsection, we describe the asymptotic performance; i.e., when time $t \to \infty$. Since $\mathbb{E}[g(t)]$ is an even function of μ, without loss of generality, we assume $\mu > 0$. Moreover, since $\gamma > 0$, it is easy to show that the winning component of the gain-loss function; which is $\mathbb{E}[g_L(t)]$ in this case, will tend to $+\infty$. That is, since

$$\alpha > \frac{|\mu K_P - \gamma|}{2}$$

then $\mathbb{E}[g_L^1(t)]$ and $\mathbb{E}[g_L^2(t)]$ will both tend to $+\infty$ as $t \to \infty$. In order to study the performance of the losing side, the "short" component" in this case, we start with the Laplace transform of the corresponding expected value of gain-loss function given by

$$G_S(s) = \frac{-\mu(s+\gamma)I_0/s}{s^2 + (\gamma + \mu K_P)s + \mu(K_I + \gamma K_P)}.$$

In order to use the Final Value Theorem, we can simply verify that roots of the denominator of $G_S(s)$ has negative real parts and moreover $(K_P, K_I) \neq (0,0)$ and $\gamma > 0$ rule out multiple poles at the origin. Hence, we can apply the Final Value Theorem to obtain

$$\lim_{t\to\infty} \mathbb{E}[g_S(t)] = \lim_{s\to 0} sG_S(s) = -\frac{\gamma I_0}{K_I + \gamma K_P}.$$

It is instructive to consider two special cases: First, when $\gamma \to 0$, with $K_I > 0$, $\mathbb{E}[g_S(t)] \to 0$ which is similar to the result we obtained for PI controller in Section 4.2. Second, when $K_I = 0$ then $\mathbb{E}[g_S(t)] \to -I_0/K_P$ which is the result in [3].

4.10 Analysis of the Covariance Matrix

The analysis of the covariance is pretty similar to the one we carried out earlier in Section 4.4. In fact, the same formula holds with the matrix A replaced with A_γ

$$A_\gamma \doteq \begin{bmatrix} -\gamma & 1 & 0 & 0 \\ \mu K_I & \mu K_P & 0 & 0 \\ 0 & 0 & -\gamma & 1 \\ 0 & 0 & -\mu K_I & -\mu K_P \end{bmatrix}.$$

That is, we obtain the Lyapunov-type linear matrix differential equation

$$\frac{dP}{dt} = A_\gamma P + P A_\gamma^T + CPC^T + (C\bar{x} + du)(C\bar{x} + du)^T,$$

where C, d are defined earlier in 4.4 and $u = I_0$. This differential equation can readily be solved and the variance of the gain-loss function, $g(t) = g_L(t) + g_S(t)$ is found using

$$\text{var}[g(t)] = h^T P(t) h; \quad h \doteq [0\ 1\ 0\ 1]^T.$$

4.11 Simulation Revisited

In this section, we revisit the simulation provided earlier in Section 4.5. In this section, we consider the historical prices for Apple (Ticker: AAPL) again and apply the PI controller with exponentially moving average. Using the data in Section 4.5, similar to the previous simulation we take $I_0 = 10000$, $K_P = 2$. Moreover, the $\gamma = 0.01$ is picked such that the ending account values and average investments are comparable. The K_I is varied again and the resulting average investments and ending account values are reported in Table 4.2. The discounting effect of exponentially weighted average is clear from the smaller average investment levels as depicted in the table which is deemed as a potential approach to address collateral requirement. However, less investment implies less risk and hence smaller reward which is evident in the ending account values reported.

Another interesting possibility would be to study the effect of this exponentially discounting scheme on the maximum percentage drawdown which was discussed earlier in Chapter 3. In Table 4.3, the maximum percentage drawdown is reported for the PI controller with and without the exponentially weighting. The use of the weighting scheme lead to smaller maximum percentage drawdown which is desired.

K_I	Acc. Value $\gamma = 0.01$	Acc. Value γ=0	Avg. Invest. $\gamma = 0.01$	Avg. Invest. $\gamma = 0$
0	\$10,098	\$10,098	\$335	\$335
0.2	\$11,087	\$11,950	\$1,918	\$3,005
0.5	\$12,569	\$14,726	\$4,320	\$7,025
1	\$15,031	\$19,357	\$8,282	\$13,630
2	\$19,912	\$28,686	\$15,934	\$26,406

Table 4.2: The Ending Account Value and Average Investment with $K_p = 2$

K_I	PI Controller	With Exp. Weighted Average
0	0.36%	0.36%
0.2	2.65%	1.68%
0.5	5.03%	3.27%
1	7.36%	5.16%
2	11.43%	8.89%

Table 4.3: The Effect of Exponentially Weighting on Expected Drawdown

4.12 Conclusion and Further Research

In this chapter, we introduced memory into the feedback-based trading rule. The resulting PI controller was analyzed and we proved that the Robust Positive Expectation Property still holds for the benchmark of GBM price process. Motivated by practical considerations, we extended the analysis to include an exponentially discounting in the PI controller. Again, we obtained a closed-form solution for the expected gain-loss function, $\mathbb{E}[g(t)]$ and established positivity in a manner similar to the one used for the classical PI controller.

The reader should not erroneously misconstrue the results in this chapter to mean that the Initially Long-Short PI controller is necessarily a good strategy in practice. The main point of this chapter is that the analysis is tractable using classical control-theoretic tools. A positive expected value does not necessarily mean that the probability of winning is significant. In fact, based on previous work for the static case, it is known that the probability density function for gains and losses can be highly skewed and that individual sample paths might exhibit significant drawdown; see [7] and [8]. The positive expectation result is useful but in practice, it should also be counterbalanced by including considerations of risk.

By way of future research, it would be of interest to formulate and solve an *optimal gain selection problem* for K_P and K_I which takes both risk and return considerations into account. In other words, we view the results in this chapter as analysis tools which will be helpful going forward as opposed to a recipe for “winning” in the stock market. In the case of PI controller with exponentially weighted moving average, the parameter γ is also a design parameter which may be included in any optimization. The issue of optimization of control parameters is pursued in the final chapter.

Chapter 5

Discrete-Time Controller With Delay: The Robust Positive Expectation Property

In contrast to the results given to date for feedback-based trading in continuous time, this chapter considers the discrete-time case. For a "low-frequency" trader such as a typical small investor, discrete-time results are more realistic. Note that discrete-time results also apply to high-frequency trading because the discretization interval Δt can be as small as we wish. Said another way, in the discrete-time setting, the same method of analysis applies whether the trader updates the investment level once every second or once every month. Another reason for consideration of discrete time is that a result in this domain readily lends itself to empirical backtesting using historical data. In contrast, results obtained in continuous time need to be discretized before their use.

With the motivation above in mind, in this chapter, to demonstrate the type of analysis which is possible in our control-theoretic setting, a new linear feedback type investment rule is introduced and formulated in discrete time.[1] This trading rule involves a *controller with delay* and is motivated by a desire to include weighting of recent performance to obtain the investment level. Once introduced, it is proven that the Robust Positive Expectation Property holds for this delay system; e.g., see Section 1.4.5 where this property is introduced. Consequently, we further discuss that this result can be reduced to the special case of *controller with no delay* which is essentially the discrete-time version of *Simultaneous Long-Short* (SLS) controller. Finally, the strategy is backtested using historical price data.

5.1 Setup For Discrete-Time Formulation

Recalling the setup for the linear feedback controller in continuous time, see Section 1.1.1, the discrete-time counterpart for the investment level at stage k is

$$I(k) = I_0 + Kg(k)$$

[1] Initial results of the work reported in this chapter have been published in [10].

with $I_0 > 0$ being the initial investment, $K \geq 0$ the feedback gain and $g(k)$ being the gain-loss function satisfying $g(0) = 0$. In this setting, with $p(k)$ being the price at stage k, the return on the stock price is given by

$$\rho(k) \doteq \frac{p(k+1) - p(k)}{p(k)}$$

for $k = 0, 1, 2, ..., N-1$. In a purely theoretical framework, these returns are typically generated by some stochastic process such as Geometric Brownian Motion (GBM). For such an example, the robustness framework of this thesis dictates that the trader does not have apriori knowledge of the GBM parameters μ and σ. That is, $K \geq 0$ above is not a function of μ or σ.

Given that the single-stage gain or loss is obtained by multiplying the percentage change in the stock price, the return $\rho(k)$, by the amount invested, we obtain

$$g(k+1) = g(k) + I(k)\rho(k).$$

For the case of linear feedback, this dynamic equation in $g(k)$ is solvable in closed form; e.g., see [10] and [115]. That is, beginning from initial condition $g(0) = 0$ and linear feedback $I(k)$ above, a lengthy but straightforward calculation leads to sample path solution

$$g(k) = \frac{I_0}{K}\left[\prod_{i=0}^{k-1}(1 + K\rho(i)) - 1\right].$$

The associated investment is then readily obtained as

$$I(k) = I_0 \prod_{i=0}^{k-1}(1 + K\rho(i)).$$

It is important to note that similar to the case of PI controller described in Section 4.1.1, the investment can change sign during the course of the trade. That is, unlike the static case in continuous time, as seen in earlier work [4], in which the sign of investment remains invariant, the investment can change sign in the discrete-time setting. For example, with $I_0 > 0$, the initially-long investment, $I(k)$ above, can be "morphed" into a short.

More specifically, this happens at stage k if $1 + K\rho(k) < 0$. That is a trader who is too "aggressive" using a gain K which is too high compared to $\rho(k)$ can be switched into a short position with $I(k+1) < 0$. A similar switching phenomenon can occur from short to long. With this consideration in mind, depending on the sign of I_0, similar to the convention used in Chapter 4, we refer to $I(k)$ as either Initially-Long or Initially-Short.

5.1.1 Investment, Interest and Collateral Considerations: With linear feedback $I(k)$ above and $V(k)$ being the account value at stage k, if $I(k) < V(k)$, the trader receives interest on excess cash at the so-called risk-free rate of return. Another possibility is that $I(k) > V(k)$. In this case, the trader borrows money from the broker and will be charged margin interest. The results presented in this chapter are for the idealized case obtained when the risk-free interest and margin interest rates are taken to be zero. Accordingly, for a single stock, we have $V(k) = V_0 + g(k)$. By way of future work, we envision a modification of the analysis similar to that in [5] to account for non-zero interest rates. Note that for cases when a trader uses no leverage, the formulae which are obtained for $g(k)$ represent a lower bound for the case when interest on idle cash is included.

5.1.2 Idealized Frictionless Markets: Similar to other results obtained throughout this thesis, the analysis in this chapter assumes an "idealized frictionless market." To briefly summarize, in such a market, there is "perfect liquidity" allowing the trader to transact as many shares as desired at the instantaneous price $p(k)$. The trader is a price taker in the sense that the stock price remains constant at $p(k)$ during the course of the transaction. In this market, it is also assumed that no transaction costs such as brokerage commissions and exchange fees are imposed and that the trading strategy meets collateral requirements of the broker so that all trades are admissible; i.e., no transactions are "stopped" and no liquidation occurs due to a failure to meet margin requirements; see Section 1.2 for more details.

In practice, the broker generally imposes a restriction on the allowed size of $I(k)$ based on the assets in the account. To illustrate, one way to model this restriction is to incorporate a constraint $|I(k)| \leq \gamma V(k)$ into the formulation with $\gamma = 2$ being rather typical. In carrying out the analysis to follow, we assume that the trader has adequate resources which can be deployed so that this trading restriction is not encountered.

5.2 Controller Delay Problems

The motivation for the study of a controller with delay is derived from the fact that the previously described linear feedback controller $I(k) = I_0 + Kg(k)$ may react rather slowly to a market losing its trend with prices moving "sideways." To elaborate, we consider the following scenario: Imagine a bull market with a large "run-up" in the price $p(k)$ from $p(0) = p_0$ to $p(k_*) \doteq p_*$ for some $k_* > 0$. We now take $g_* \doteq g(k_*)$ to be the trading gain which results at this bull market high and imagine the market "stalling" for $k \geq k_*$. For simplicity, say $p(k) = p_*$ for some period beyond $k = k_*$.

In the scenario above, the key point to note is that for $k \geq k_*$, with price fixed at p_*, the linear feedback controller maintains the bull market "high" investment level $I(k) = I_0 + Kg_*$ even when the market fails to progress and $g(k)$ remains constant. The motivation for the new controller described below is that many investors would argue that $I(k)$ should not be maintained at an "undeservedly high" level when the market moves sideways over an extended period of time. After some "waiting period," some investors might consider it time to "lighten up" on investment level $I(k)$ as price $p(k)$ begins to stall .

In order to be more responsive to the considerations above, in this chapter, we propose a new linear feedback trading rule to address this type of issue; i.e., at stage k, we consider a controller which focuses on recent performance via inclusion of a delay element. The controller includes a term $g(k - m)$ where m is some pre-specified look-back period. Then, at time k, the investment level is given by

$$I(k) = I_0 + K\left[g(k) - g(k - m)\right].$$

Notice that the investment level $I(k)$ is attenuated as the "winning power" goes away; i.e., as $g(k)$ ceases to increase. In addition, if the market reverts to profitability, the controller will respond to this change and begin increasing $I(k)$. Later in this chapter, we will generalize the analysis to follow to the SLS case.

5.2.1 Trading Dynamics for the Delay System:

Beginning with the gain-loss equation

$$g(k + 1) = g(k) + I(k)\rho(k),$$

from Section 5.1, upon substitution of the controller above, using a standard "ploy" which is common for delay systems, we can derive a linear time-varying state equation to describe the dynamics.

That is, defining state vector

$$x(k) \doteq \begin{bmatrix} g(k) \\ g(k) - g(k-1) \\ g(k-1) - g(k-2) \\ \vdots \\ g(k-m+1) - g(k-m) \end{bmatrix},$$

after some algebra, the gain-loss dynamics reduce to

$$x(k+1) = A(\rho(k))x(k) + b\rho(k)I_0$$

where the matrix $A(\rho(k))$ has dimension $(m+1) \times (m+1)$ and is given by

$$A(\rho(k)) \doteq \left[\begin{array}{c|cccc|c} 1 & K\rho(k) & K\rho(k) & \dots & K\rho(k) & K\rho(k) \\ 0 & K\rho(k) & K\rho(k) & \dots & K\rho(k) & K\rho(k) \\ \hline \mathbf{0}_{(m-1)\times 1} & & \mathbf{I}_{(m-1)\times(m-1)} & & & \mathbf{0}_{(m-1)\times 1} \end{array}\right]$$

and

$$b \doteq \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}^T.$$

Now, using initial conditions $g(0) = g(-1) = \cdots = g(-m) = 0$, a lengthy but straightforward calculation leads to solution at time $k = N$ given by

$$x(N) = \sum_{k=0}^{N-2} \left(\Big(\prod_{i=0}^{N-k-2} A(\rho(N-i-1)) \Big) b\rho(k)I_0 \right) + b\rho(N-1)I_0.$$

Beginning with the formula above, in order to get a handle on the expected gains or losses, we assume a stochastic process which governs the returns $\rho(k)$ and that these random variables are independent with common mean

$$\mathbb{E}(\rho(k)) = \mu.$$

Then, since $x(N)$ is obtained by multiplication of independent matrices $A(\rho(i))$, it is easy to verify that it will be a sum of products of independent random variables leading to

$$\mathbb{E}(x(N)) = \left(\sum_{k=0}^{N-1} A^k(\mu) \right) b\mu I_0,$$

where

$$A(\mu) \doteq A(\rho(k))|_{\rho(k)=\mu}.$$

That is, the matrix $A(\mu)$ is the matrix $A(\rho(k))$ with returns $\rho(k)$ replaced by $\mu = \mathbb{E}[\rho(k)]$. In the sequel, an extension of the formula above, combining a long and a short component, is used to obtain a robust positive expectation property result for an SLS-type controller with delay.

5.3 Robust Positive Expectation Property for Delay System

The Initially Long-Short (ILS) feedback controller *with delay* is the focus of this section. Below, we provide the main result in this chapter: a robust positive expectation theorem for the class of delay systems under consideration. In this setting, for a pre-specified delay of $m > 0$, the investment for the two components at step k are given by

$$\begin{aligned} I_L(k) &\doteq I_0 + K\big(g_L(k) - g_L(k-m)\big); \\ I_S(k) &\doteq -I_0 - K\big(g_S(k) - g_S(k-m)\big), \end{aligned}$$

where $I_0 > 0$ is the initial investment in dollars, $K \geq 0$ is the feedback gain and $g(k)$ is the trading gain-loss function at stage k. The subscripts "L" and "S" denote the initially-long and initially-short investment and gain-loss components respectively. The total gain-loss function given by

$$g(k) = g_L(k) + g_S(k)$$

with dynamics for these components described by

$$g_L(k+1) = g_L(k) + \rho(k) I_L(k); \quad g_S(k+1) = g_S(k) + \rho(k) I_S(k).$$

5.3.1 Theorem: (Robust Positive Expectation) *Consider the ILS feedback controller with delay $m > 0$ as described above. Assume further that the returns $\rho(k)$ are independent with common expected value $\mathbb{E}\big(\rho(k)\big) = \mu$ for all k. Then, except for the break-even case obtained when either $K\mu = 0$ or $N = 1$, at any stage $N > 1$ we have*

$$\mathbb{E}\big(g(N)\big) > 0,$$

and $\mathbb{E}\big(g(N)\big)$ is an increasing function of $|\mu|$.

Proof: For the degenerate break-even case $K\mu = 0$, using the updates for $g_L(k)$ and $g_S(k)$, it is straightforward to show that by induction that $\mathbb{E}\big(g(k)\big) = 0$ for all k. Also for the case of $N = 1$, since the net investment at time zero is $I(0) = I_L(0) + I_S(0) = 0$, it follows that $\mathbb{E}\big(g(1)\big) = 0$. Henceforth, we assume that $K\mu \neq 0$ and $N > 1$ in the remainder of the proof.

To complete the proof, it suffices to show that $\mathbb{E}\big(g(N)\big)$ is a polynomial in μ^2 with positive coefficients which is not constant with respect to μ. Starting with the state vector for the long component and noting the formula for $\mathbb{E}\big(x(N)\big)$ in Section 5.2.1, we have

$$\mathbb{E}\big(x_L(N)\big) = \left(\sum_{k=0}^{N-1} A^k(\mu)\right) b\mu I_0.$$

Similarly for the case of initially-short trade, by replacing I_0 with $-I_0$ and K with $-K$ in the $\mathbb{E}\big(x(N)\big)$ formula, we obtain

$$\mathbb{E}\big(x_S(N)\big) = -\left(\sum_{k=0}^{N-1} A^k(-\mu)\right) b\mu I_0.$$

Now since $\mathbb{E}\big(g(N)\big)$ is the first entry of $\mathbb{E}\big(x_L(N) + x_S(N)\big)$, letting

$$c \doteq \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}^T,$$

we obtain

$$\begin{aligned} \mathbb{E}\big(g(N)\big) &= c^T\, \mathbb{E}\big(x_L(N) + x_S(N)\big) \\ &= c^T \left(\sum_{k=0}^{N-1} A^k(\mu) - A^k(-\mu)\right) b\mu I_0. \end{aligned}$$

Now letting $A^k_{i,j}(\mu)$ denote the (i,j)-th entry of the matrix $A^k(\mu)$, the expression above reduces to

$$\mathbb{E}\big(g(N)\big) = \left(\sum_{k=0}^{N-1} A^k_{1,1}(\mu) - A^k_{1,1}(-\mu)\right) \mu I_0 + \left(\sum_{k=0}^{N-1} A^k_{1,2}(\mu) - A^k_{1,2}(-\mu)\right) \mu I_0. \quad (*)$$

Claim: For all k, the first column of $A^k(\mu)$ is $a^k(\mu) = c$.

Proof: By induction, since the case $k = 0$ is trivial, we now assume that this claim holds for $k \geq 1$ and must prove that it holds for $k + 1$. Indeed, writing

$$A^{k+1}(\mu) = A(\mu) \times A^k(\mu)$$

and noting that both $A(\mu)$ and $A^k(\mu)$ have c as their first columns, it follows that $A^{k+1}(\mu)$ does too.

By the claim above, we have $A^k_{1,1}(\mu) = 1$ for all k and a straightforward use of change of variables leads to $A^k_{1,1}(-\mu) = 1$ too. Now using $A^k_{1,1}(\mu) = A^k_{1,1}(-\mu)$ in $(*)$ above, we have

$$\mathbb{E}\big(g(N)\big) = \left(\sum_{k=0}^{N-1} A^k_{1,2}(\mu) - A^k_{1,2}(-\mu)\right) \mu I_0.$$

To further reduce $\mathbb{E}\big(g(N)\big)$, we note that entries of the matrix $A(\mu)$ are affine in μ with non-negative coefficients. This implies that the entries of the matrix $A^k(\mu)$, such as $A^k_{1,2}(\mu)$ which is of interest, are polynomials in μ with non-negative coefficients. Hence, $A^k_{1,2}(-\mu)$ is also a polynomial in μ with the same even part and oppositely-signed odd part. Putting the facts together, we conclude that $A^k_{1,2}(\mu) - A^k_{1,2}(-\mu)$ is a polynomial in μ with non-negative coefficients and only odd degrees of μ present for all $k \geq 0$. After the multiplication by the extra μI_0 term in the $\mathbb{E}\big(g(N)\big)$ formula, we are left with a polynomial with non-negative coefficients and with only even powers of μ present. In other words, $\mathbb{E}\big(g(N)\big)$ is a polynomial in μ^2 with non-negative coefficients. To prove that this polynomial is not constant with respect to μ, recalling that for $N > 1$ and $K\mu \neq 0$, we use the fact that $\mathbb{E}\big(g(N)\big)$ is lower-bounded by the sum of the first two polynomial terms which is easily obtained to be $2K\mu^2 I_0$. Then, we have

$$\mathbb{E}\big(g(N)\big) \geq 2K\mu^2 I_0 > 0.$$

This not only completes the proof for the Robust Positive Expectation Property but also shows that $\mathbb{E}\big(g(N)\big)$ is an increasing function of $|\mu|$. □

5.3.2 Remark: Recalling the discussion in Section 5.1, the two investment components, labeled as "long" and "short," can switch signs as trade progresses. To demonstrate that this is possible we consider the *initially*-long investment $I_L(N)$. Now with

$$e \doteq \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & \cdots & 1 \end{bmatrix}^T,$$

we calculate

$$\begin{aligned} I_L(N) &= I_0 + K\big(g_L(N) - g_L(N-m)\big) \\ &= I_0 + Ke^T x_L(N) \\ &= I_0 \left(1 + Ke^T \sum_{k=0}^{N-2} \Big(\prod_{i=k+1}^{N-1} A(\rho(i)) b\rho(k)\Big) + b\rho(N-1)\right) \end{aligned}$$

where $x_L(N)$ is the state vector for the long component as defined in the beginning of the proof of the theorem above. Now, the intended long investment $I_L(N)$ will be morphed to short if $g_L(N) - g_L(N-m) < -I_0/K$. Indeed, an equivalent condition in terms of the returns $\rho(i)$, can be obtained using the last equation above. The following example demonstrates this possibility.

5.3.3 Example (Morphing of Investment): Consider the long investment given above with initial investment $I_0 = 1$, feedback gain $K = 2$ and delay $m = 1$. Moreover, suppose we encounter the 3-stage sample path given by $\rho(0) = 0.1$ and $\rho(1) = -0.8$. Then the corresponding investment level will be $I_L(0) = 1,\ I_L(1) = 1.2,\ I_L(2) = -1.12$. For this case with $N = 2$ and $m = 1$ we have $g_L(2) - g_L(1) = -1.06$ which is less than $I_0/K = -0.5$ and accounts for the morphing of an initially-long position into a short position.

5.3.4 Discrete-Time Version of Simultaneous Long-Short Result: The focal point of this section is the discrete-time SLS-type controller with *no delay* along the lines of [10] and [115]. Our objective is to establish discrete-time results presented in this chapter reduce to SLS analogous of existing results in the literature such as [1] and [4].

To see that the Robust Positive Expectation Property for this delay-free case can be established, using Theorem 5.3.1, we assume $K\mu \neq 0$ and $N > 1$ and take the delay amount to be $m = N$. Taking note that initial conditions $g(-N) = g(-N+1) = \cdots = g(0) = 0$ are used in the theorem, we obtain the desired result.

It is also possible to obtain this result without resorting to delay systems. Starting with the same set of assumptions; i.e., returns $\rho(k)$ being independent with common expected value $\mathbb{E}[\rho(k)] = \mu$, and combining the closed-form solution for the long and short components of the gain-loss function using the formulae in Section 5.1, [10] and [115] with $K \geq 0$, we obtain

$$\mathbb{E}\big(g(N)\big) = \mathbb{E}\big(g_L(N) + g_S(N)\big) = \frac{I_0}{K}\left[(1+K\mu)^N + (1-K\mu)^N - 2\right].$$

Expanding the powers above and noting cancellations of odd-order terms, it follows that $\mathbb{E}\big(g(N)\big)$ is a polynomial in μ^2 with positive coefficients. Hence, except for the break-even cases when either $N = 1$ or $\mu K = 0$, the Robust Positive Expectation Property, $\mathbb{E}\big(g(N)\big) > 0$, holds.

Per our earlier discussion in Sections 5.1 and 5.3.2, it is important to note that in the setting above, the investment components can change signs during the course of the trade. Consistent with the convention used in Chapter 4 and earlier in this section, in discrete-time, we refer to this sort of strategy as *Initially Long-Short (ILS)*.

5.4 Backtesting Using SPDR S&P 500 Trust ETF

This section is dedicated to backtesting of the ILS controller with delay using historical prices. In the backtest to follow, we consider the exchange-traded fund with ticker SPY which tracks the S&P 500 index. In order to study performance under both bullish and bearish market trends, we consider the eight-year period beginning January, 1, 2003 and ending on December, 31, 2010. More specifically, we use the bi-weekly, dividend-adjusted price for this ETF shown in Figure 5.4.1. As seen in the figure, this time period covers both the major bull market high in 2007 and the period of the "crash" in 2008 and 2009.

5.4.1 The Simulation: Beginning with the controller in Section 5.3 with investment level

$$\begin{aligned} I_L(k) &= I_0 + K\left[g_L(k) - g_L(k-m)\right]; \\ I_S(k) &= -I_0 - K\left[g_S(k) - g_S(k-m)\right], \end{aligned}$$

in the simulations to follow, we take initial investment $I_0 = 1$, delay $m = 10$ and feedback gain $K = 2$. Assuming the investor has no initial idle cash on hand at $k = 0$, the initial account value is $V_0 = V(0) = 1$. Then, as time evolves, the total account value $V(k)$ is obtained by summing the contributions of the "long" and "short" components. That is, the gains and losses are summed up and we obtain

$$V(k) = V_0 + g_L(k) + g_S(k).$$

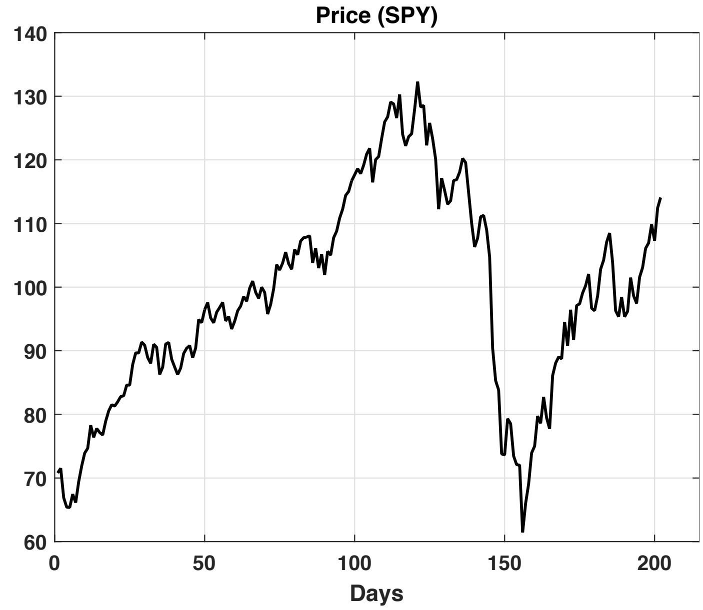

Figure 5.4.1: Price for (SPY) from 2003 to 2011

In the plots below, we compare our results to the performance of the ILS linear feedback controller with no delay which uses the same gain K and the same initial investment. In Figure 5.4.2, the resulting account values are shown. As seen in the figure, the controller with delay suffers a lower percentage loss during the crash which occurs around $k = 120$ and bounces back more quickly when the market rebounds around $k = 160$.

In Figure 5.4.3, the net investment level, $I(k) = I_L(k)+I_S(k)$ is provided. As shown, the controller with delay responds more quickly to the market crash by becoming negative. It also rebounds faster when the new bull market returns around $k = 160$. Finally, Figure 5.4.4 shows the so-called “Log-Investment Ratio.” This quantity is the natural logarithm of the ratio of absolute net investment for the controller with delay to the absolute net investment for the controller with no delay. This quantity being negative for more than 88% of the time means the strategy with delay is putting less money at risk. Moreover, the spike around $k = 150$, the crash period, is attributable to the fact that the controller with no delay is almost out of the trade while the controller with delay is making profit by taking short position; see Figure 5.4.3.

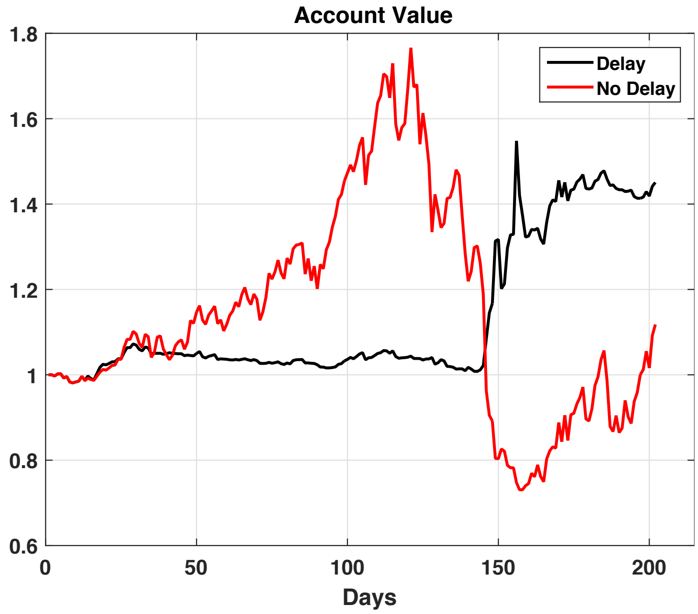

Figure 5.4.2: Account Values: With and Without Delay

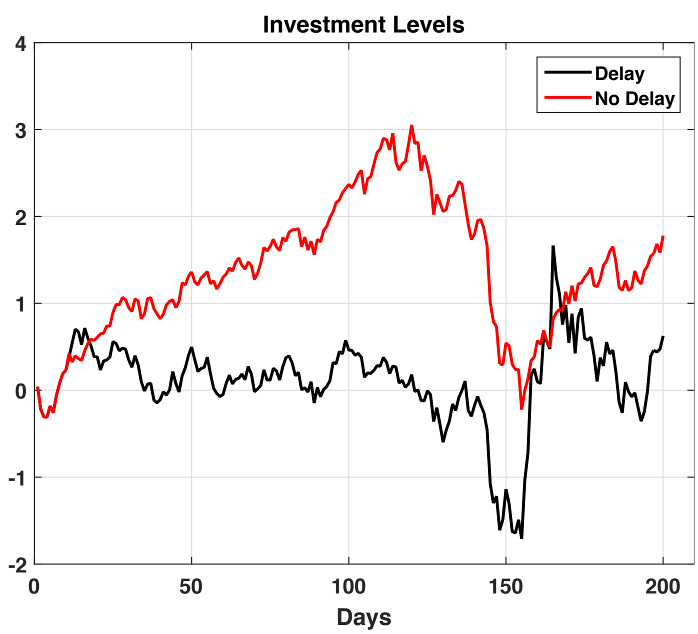

Figure 5.4.3: Investment Level: With and Without Delay

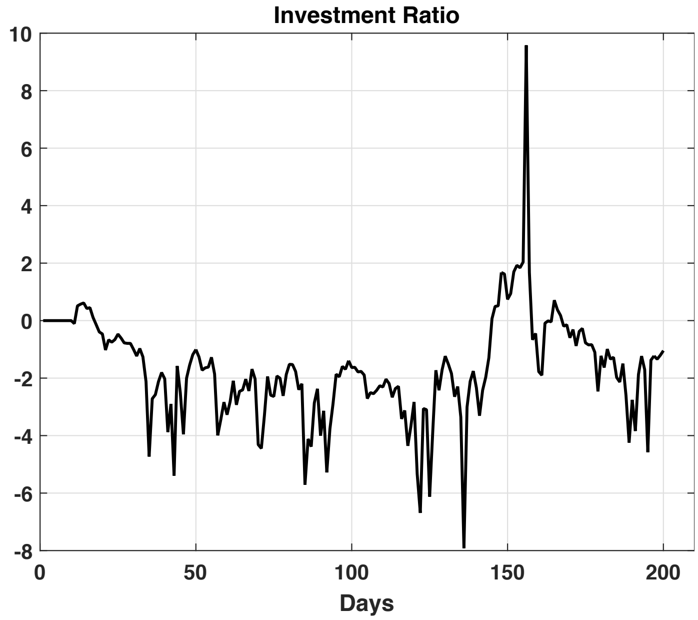

Figure 5.4.4: Logarithm of Invest. Ratio of With Delay to No Delay

5.5 Conclusion and Further Research

In this chapter, we considered a discrete-time stock trading strategy including both delay and linear feedback. We proved that an Initially Long-Short (ILS)-type version of the controller with delay has the Robust Positive Expectation Property. We also showed that this result reduces to a "standard" type of SLS controller with no delay.

There are many possibilities for building upon the results in this chapter. One important direction involves consideration of various performance metrics beyond the expected value. Another research possibility involves studying the dynamics of a more general two-gain controller given by

$$I(k) = I_0 + K_1 g(k) + K_2 g(k - m).$$

Notice when $K_1 = -K_2 = K$, we recover delay analysis in Section 5.2 and when $K_2 = 0$, we recover a "classical" linear feedback. Even more generally, one can consider an ARMA-like dynamic controller with multiple delays given by

$$I(k) = I_0 + \sum_{i=0}^{m} K_i g(k - i).$$

Note that in this case, when the discretization interval is suitably small with all K_i equal, this approximates a PI controller; see Chapter 4 for details.

Chapter 6

When the Expected Value is not Expected: Problems Requiring Emphasis on Conservatism

Our research in this thesis has resulted in some interesting "offshoots." One of them is the development of a new reward-risk pair which we called *Conservative Expected Value* (CEV) and *Conservative Semi-Variance* (CSV) and are the focus of this chapter.[1] Much of the motivation for development of this pair comes from the fact that a highly-skewed probability distribution for the gain-loss function $g(t)$ might result when feedback is used. We envision the theory to follow to be applicable to areas other than stock trading. Accordingly, the exposition in this chapter is more general than the "finance-flavored" focus of Chapters 1-5.

The takeoff point in this chapter is a random variable X for which large positive values are desired. When its probability distribution is highly skewed, the possibility of a long fat-tailed distribution can lead to an expected value, $\mu = \mathbb{E}[X]$, which is unduly optimistic. For the reverse case when small values of X are desired, the ideas in this chapter are applied to $-X$. This issue of over-optimism in the expected value is particularly important when a mission-critical random variable is involved. For example, when considering earthquake intensity or flood levels, reporting an understated expected value to a technically unsophisticated general public would be considered by many as highly undesirable.

The *Conservative Expected Value* of X, denoted by $\mathbb{CEV}(X)$ is a new definition provided in this chapter. It is a new metric which we argue is particularly useful when risk aversion must be highly emphasized. When $\mathbb{E}[X]$ does not represent what can reasonably be expected, then $\mathbb{CEV}(X)$, while being conservative, is defined in such a way so as not to be unduly pessimistic. In classical analysis, enhancement of a calculation often includes the variance $\sigma^2 = \text{var}(X)$. However, when large X-values are desired, this may further distort one's overview of the risk at hand; e.g., if X is profit, values above the mean should not be penalized. For example, this is the case for the

[1]The results reported in this chapter have been published in [116] and [117].

gain-loss function; i.e., $X = g(t)$ resulting from trading via linear feedback; e.g., see Chapter 2. With these one-sidedness considerations in mind, we introduce as new reward-risk pair, the CEV and the so-called *Conservative Semi-Variance* of X, denoted by $\mathbb{CSV}(X)$. Whereas the CEV definition is entirely new, the CSV definition is similar in flavor to various risk metrics used in finance; e.g., see [105] and [118]. More specifically, in finance, it is standard to use a *semi-variance measure* to penalize downside variations; e.g., see [90] and [91]. This measure motivates the definition of the CSV.

In this chapter, we also illustrate calculation of the (CEV,CSV) pair for a number of classical probability distributions and we describe a number of properties of these metrics which suggest that this new theory is mathematically rich. Finally, we demonstrate the potential for application via three examples, first example is motivated by the skewing effect in trading via feedback-based trading rule which was the focal point in Chapter 2 and the other two are numerical examples coming from real-world data.

6.1 Introduction

The focal point in this chapter is a random variable X for which large positive values are desirable and downside risk is of paramount concern. As previously stated, the results to follow are also applicable when "smallness" of X is considered to be the desirable outcome. In this case, we apply the results to follow to the negated random variable $-X$. For example, if the random variable of interest corresponds to earthquake magnitude, see Section 6.2.3, the smaller its value, the better. For a random variable X, the first focal point in this chapter is the classical expected value, denoted by $\mu = \mathbb{E}[X]$. For cases when the underlying probability distribution for X is either highly skewed or untrustworthy in nature, the expected value may be unduly optimistic and should not be viewed as "expected." In such a scenario, a long, possibly fat tail for the distribution can make $\mathbb{E}[X]$ so different from what one can reasonably expect so as to make this measure misleading and of questionable worth. In many cases, it may be exceedingly probable that the outcome will be much worse than $\mathbb{E}[X]$. Returning to the earthquake example, when "expected magnitude" is reported to a technically unsophisticated public, given that protective measures are taken in proportion to the perceived risk, it is important to provide measures which are both simple to understand and which emphasize conservatism.

Another motivating example comes from finance literature, see [119–121] where extreme, unlikely "crash" events are studied in the context of investment and conservatism is of paramount concern. Similar issues also arise in large deviation theory, see [122], where such events are studied. With the above providing initial context, the first main objective in this chapter is to introduce a new measure, called the *Conservative Expected Value* which we denote by $\mathbb{CEV}(X)$. Essentially, this metric discounts the classical expected value by including heavy emphasis on downside risk which is important for the type situations we have in mind. At the same time, as explained in the sequel, while the CEV is intended to be conservative in nature, it is defined in such a way as to avoid being unduly pessimistic. Given the preferences for large X versus small, there is a danger that complementing any analysis by introducing the classical variance, $\sigma^2 = \text{var}(X)$, will distort the picture of the risk at hand. For example, an investor with very high risk aversion may get a distorted picture of risk because large positive values of trading gains X, while contributing to the variance, are considered to be good outcomes. It is only the downside risk associated with large negative X which is of major concern.

With all of the considerations above in mind, this chapter departs from the voluminous body of literature dealing with the classical mean-variance paradigm; e.g., see classical work such as [15] and [98] in finance, [123–125] in signal processing and [126] in network analysis. To summarize, our aim in the sequel is to provide a conservative alternative to classical return-risk analysis which uses the pair (μ, σ^2). To get a meaningful characterization of the likely reward versus risk, we do not deal with the technical complexities associated with higher order moments or the entire probability distribution for X; e.g., see [102] and [127]. When providing summarizing information, our goal is to provide new measures which are intended to be simple surrogates for the classical mean and variance and which address situations when a conservative assessment is essential.[2]

[2]Another motivation for this work is derived from "distributional robustness" considerations. In this situation, even in the absence of skew, the need for discounting the classical mean-variance pair is derived from the fact that the underlying probability density function $f_X(x)$ for X may not be well known. For example, one might envision a sphere of uncertainty in an appropriate vector space centered on f_X; e.g., see [128]. Without a parameterization for this distribution, discounting must be based entirely on the "nominal" distribution for X as done in this chapter.

The type of situation described above is epitomized by the celebrated St. Petersburg Paradox first pointed out by Bernoulli as early 1738; see [129] where his work was reprinted. A probability mass function for payoff X, say in dollars, is described by $X = 2^k$ with probability $1/2^{k+1}$ and k ranging over the non-negative integers. While the expected value $\mathbb{E}[X]$ is readily verified to be infinite, the probability of receiving a specified "large" payoff is very small. Furthermore, even when variance considerations are introduced, the unboundedness of the risk leads to an ambiguous characterization of the bet at hand. In this regard, many authors after Bernoulli have argued that no rational individual would pay more than two or three dollars for a ticket to play such a game; e.g., see [130–132]. Simply put, in many cases the expected value appears to be an "unrealistic expectation."

The inconsistency between the information provided by the expected value and decision-based behavior, presumably based on maximization of utility, has been studied in the literature; e.g., see Allias, [133] and Ellsberg, [134], where scenarios of the sort described above are studied from both a mathematical and behavioral point of view. In the presence of skewness, the probability that X is far below $\mathbb{E}[X]$ can be very high. Hence, $\mathbb{E}[X]$ can be quite misleading as a predictor of the way a "bet" might be assessed by individuals making decisions. For example, if using flood data to provide guidelines on how high above sea level to locate one's home, given the potential catastrophe involved, a conservative expected value for the flood level may be better suited for risk-averse individuals versus the standard expected value. Another class of problems where these considerations are paramount involves component reliability [135] for so-called mission-critical systems; e.g., assessment of lifetime for components in a nuclear plant. To summarize, the starting point in this chapter involves situations for which the expected value is arguably not expected and it may be very costly to act on unduly optimistic information.

6.1.1 Risk Considerations: As already indicated, any discussion along the lines above would be incomplete without bringing risk metrics into the picture. In other words, to simply assess whether a probability distribution is attractive or not, it does not suffice to look solely at the expected value $\mathbb{E}[X]$. Even after adjustment for skewness, a bet with a large expected value may nevertheless be deemed unattractive based on risk considerations. For example, imagine a hedge fund manager providing summarizing portfolio gain-loss prospects to a group of technically unsophisticated investors via the moments of a highly-skewed probability distribution; e.g., see [7, 99, 136, 137]. The information provided should be easy to understand leading to a simple picture of the risk and rewards, and, analogous to the classical (μ, σ^2) pair, it should be independent of utility considerations; i.e., each user of the information can apply their own utility functions in making judgments about attractiveness of the situation and associated random variables under consideration.

6.1.2 Salient Features of Our Work: With the scenarios above providing motivation, our goal in this chapter is to develop a new alternative to the classical mean-variance pair. To this end, the applications which we have in mind are characterized by some salient features: First, as previously stated, we assume that larger values of X are preferred to smaller values. Second, we assume that a worst-case value α_X for X is finite. For example, when considering the lifetime of a hardware component in an engineering system, it would be reasonable to take $\alpha_X = 0$ since X cannot be negative. It is also important to note that the theory to follow is readily adapted to cases when smaller values of X are preferred to larger values. For example, when we deal with earthquake intensity in the section to follow, with X representing earthquake intensity, we apply our theory to $-X$. The third salient feature, as already stated, is that a high degree of risk aversion is in play. For example, when a mission-critical hardware component's lifetime is involved, a simple system maintenance policy based on the classical expected value might lead to insufficient frequency of replacement. In order to "robustify" against skew and possible inaccuracies in the probability distribution, a conservative analysis which addresses risk aversion demands discounting the classical expected value — better safe than sorry when the cost of failure is high. The fourth and final salient feature in the theory to follow involves a distinction between downside and upside risk. The risk measure which we develop does not penalize possible outcomes which are larger than what is expected. Motivated by the work of Sortino and others, see [90, 91, 138], we define a semi-variance type measure which only penalizes X-values which are below what is expected.

6.1.3 Plan For the Sequel: In Section 6.2, for a random variable X with finite worst-case value α_X, we define the so-called *Conservative Expected Value* which is denoted by $\mathbb{CEV}(X)$. Once defined, some elaborative remarks are provided and we provide a motivating example which is illustrative of the applications we have in mind. In Section 6.3, the CEV is calculated for some classical probability distributions and results are compared to their classical expected-value counterparts. On the risk side of the analysis, in Section 6.4, we define the so-called *Conservative Semi-Variance*, denoted by $\mathbb{CSV}(X)$ which, for some of the classical probability distributions, is compared to the classical variance in Section 6.5. These new measures are obtained using a discounting mechanism leading to the resulting $(\mathbb{CEV}(X), \mathbb{CSV}(X))$ pair more accurately reflecting the degree of "attractiveness" of the reward-risk situation when high skew may be present. In Section 6.6, we describe a byproduct of the CEV framework — a tightening of the classical Markov and Chebyshev inequalities. Noting that our definition leads to $\mathbb{CEV}(X) \leq \mathbb{E}[X]$, given $\epsilon > 0$, instead of the classical Markov inequality $P(X > \epsilon) \leq \mathbb{E}[X]/\epsilon$, we obtain the tighter version

$$P(X > \epsilon) \leq \frac{\mathbb{CEV}(X)}{\epsilon}.$$

Also seen in Section 6.6, the basic risk metrics defined by CEV and CSV are not *sub-additive* and hence can not be coherent; see [105, 118, 139]. In Section 6.7, we consider the application of CEV and CSV analysis most closely related to the theme of this thesis — namely stock trading via linear feedback. Subsequently, more application examples involving S&P 500 stock index returns and household income data are also provided to demonstrate how the new theory can be used for probability distributions estimated from real-world data. Finally, in Section 6.8, a discussion of possible future research directions is provided.

6.2 The Conservative Expected Value

As discussed earlier in Section 6.1, for a random variable X for which large values are desired, the possible long, fat right-sided tail can be problematic. For such a scenario the highly-skewed distribution may result in a large expected value, $\mu = \mathbb{E}[X]$, which can be deemed an unduly optimistic indicator of performance. When the probability is small that X exceeds $\mathbb{E}[X]$, it is arguable that the expected value is not expected. Furthermore, enhancing the analysis to include the classical variance, $\sigma^2 = \text{var}(X)$, can add to this confusion because this measure penalizes the variation of random variable X in both directions even though upside variation is desirable.

In the analysis to follow, we take the leftmost support-point of probability distribution for random variable X, denoted by α_X, to be known and finite. That is, with $F_X(x)$ denoting the cumulative distribution function for random variable X, we assume

$$\alpha_X \doteq \inf\{x : F_X(x) > 0\} > -\infty.$$

This assumption on the "worst-case" outcome is realistic in many applications: For example, the worst-case loss of a linear feedback trading rule when the price is driven by Geometric Brownian Motion is given in Theorem 1.4.6; that is $\alpha_X = g^*$. Additional examples involve the price of a stock or the lifetime of a component in a system which are both non-negative random variables; that is $\alpha_X = 0$.

To motivate the definition to follow, we imagine a risk-averse individual who is rewarded or penalized based on X. If this person has a "minimal acceptable target" for X, denoted by γ, a conservative approach to the analysis of this gamble X would be to shift the probability mass associated with $X \le \gamma$ to α_X, and to shift the probability mass associated with $X > \gamma$ to γ. This procedure maps the original random variable to a Bernoulli random variable X_γ described by

$$X_\gamma \doteq \begin{cases} \alpha_X & \text{with probability} \quad P(X \le \gamma) = F_X(\gamma); \\ \gamma & \text{with probability} \quad P(X > \gamma) = 1 - F_X(\gamma). \end{cases}$$

In the sequel, we let $f_X(x)$ be the probability density function for X. Then, this mass-shifting, shown in Figure 6.2.1, lessens the effect of large skewness by discounting the long possibly-fat tail.

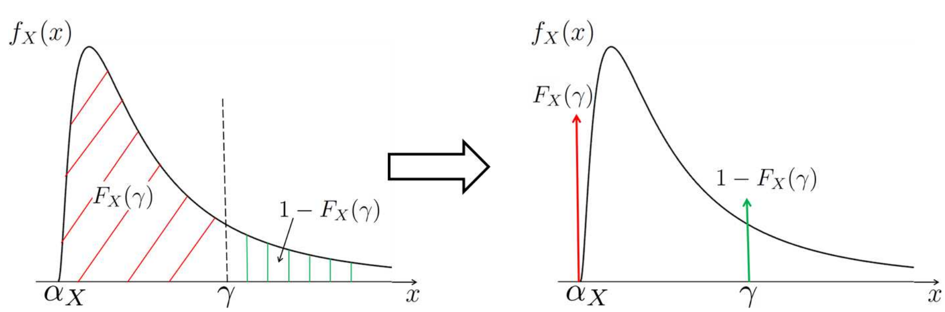

Figure 6.2.1: Mapping of Random Variable X to Bernoulli Random Variable X_γ

After this conservative procedure is carried out, one is left with a simple Bernoulli random variable X_γ with expected value

$$\mathbb{E}[X_\gamma] = \alpha_X F_X(\gamma) + \gamma\big(1 - F_X(\gamma)\big).$$

This user-dependent mass-shifting procedure forces

$$\mathbb{E}[X_\gamma] \leq \mathbb{E}[X].$$

In the definition of the CEV to follow, the user-dependence on γ is reduced via a maximization process which we now explain.

Maximizer of $\mathbb{E}[X_\gamma]$: Suppose $\gamma = \gamma^*$ is a maximizer of $\mathbb{E}[X_\gamma]$. Then, we claim that any $\gamma < \gamma^*$ is "inefficient" in the following sense: Since $\mathbb{E}[X_\gamma] \leq \mathbb{E}[X_{\gamma^*}]$, it follows that the pair $\big(\gamma, \mathbb{E}[X_\gamma]\big)$ is *dominated* by $\big(\gamma^*, \mathbb{E}[X_{\gamma^*}]\big)$. Hence for $\gamma < \gamma^*$, it is arguable that even individuals with differing utility functions will reject γ in favor of γ^*. Said another way, $\gamma < \gamma^*$ is too conservative. On the other hand, for an individual seeking a $\gamma > \gamma^*$, there is a trade-off between the minimal acceptable target and the expected outcome. In addition to inefficiency considerations, as we see in Section 6.4, the choice of γ^* will lead to the most conservative downside risk assessment.

6.2.1 Conservative Expected Value Definition: Given the random variable X, the *Conservative Expected Value (CEV)* is defined to be

$$\mathbb{CEV}(X) \doteq \sup_{\gamma} \mathbb{E}[X_\gamma].$$

Equivalently,

$$\begin{aligned}
\mathbb{CEV}(X) &= \sup_{\gamma} \alpha_X F_X(\gamma) + \gamma(1 - F_X(\gamma)) \\
&= \sup_{\gamma \geq \alpha_X} \gamma + (\alpha_X - \gamma) F_X(\gamma).
\end{aligned}$$

6.2.2 Remark: Since we already know that $\mathbb{E}[X_\gamma] \leq \mathbb{E}[X]$, by taking the supremum with respect to γ, it is immediate that

$$\mathbb{CEV}(X) \leq \mathbb{E}[X].$$

In summary, for reasons previously given, the CEV "discounts" the classical expected value. Rather than beginning with a structured utility function, $\mathbb{CEV}(X)$ is based on the idea that the value α_X is known and that an individual wants to discount the misleading effect of long, fat-tailed probability distribution as part of conservative analysis. This point of view is similar to the approach taken to define wins and losses in the celebrated paper on Prospect Theory by Kahneman and Tversky in [130].

6.2.3 Motivating Example Illustrating Need For Conservatism: This example comes from an important line of research in the Geophysical Systems literature which is dedicated to predictive modeling of earthquakes; e.g., see [140–143]. In this setting, a widely-used measure for earthquake intensity is the *moment magnitude* which is denoted by m_W. The larger the value of m_W, the more severe the damage is expected to be. Given the potential catastrophes involved, one issue in the literature involves the extent to which existing models understate the risk to the public. This is important because it has an effect on the development of safety policies and insurance rates; e.g., see [141]. In this regard, decisions based on the "expected" magnitude, $\mathbb{E}[m_W]$, of a earthquake has been blamed for recent tragic failures; e.g., see [144] and [145] where this issue is studied for the case of Tohoku earthquake and Tsunami in 2011. One approach, model improvement, involves finding a "robust" estimate of the magnitude; e.g., see [146]. Below, we consider the handling of this issue using the CEV. To provide a first illustration of our theory, recalling that our random variable X is such that large outcomes are desirable, we take $X \doteq -m_W$. Therefore, if m_{max} denotes the maximum possible magnitude obtained from historical data, we take $\alpha_X = -m_{max}$ in the calculation to follow.

6.2.4 CEV Analysis Example: In this example, we use the historical data for earthquake magnitudes, m_W which occurred in California between 1800 and 1985. This data, reported in logarithmic scale, is obtained from [147] and includes about $57,000$ measurements. Using the histogram representing this data in Figure 6.2.2, the expected value, variance and skewness of the resulting empirical distribution are

$$\mathbb{E}[m_W] \approx 5.08; \quad \text{var}(m_W) = 1.35; \quad S(m_W) \approx 0.35$$

respectively. Interestingly, even though logarithmic data is often used in statistical analysis to remove the skewness, for example, see [148], as evidence by our calculation, there still appears to be considerable "leftover" skewness which we claim can lead to "understatement" of expected earthquake magnitude.

Using the data, it is easy to obtain $P\big(m_W > \mathbb{E}[m_W]\big) \approx 0.49$. Basically, this says that there is a nearly 50-50 chance that earthquake intensity will exceed 5.08. Given that distributions of earthquake magnitude is arguably not trustworthy, we now calculate our more conservative estimate

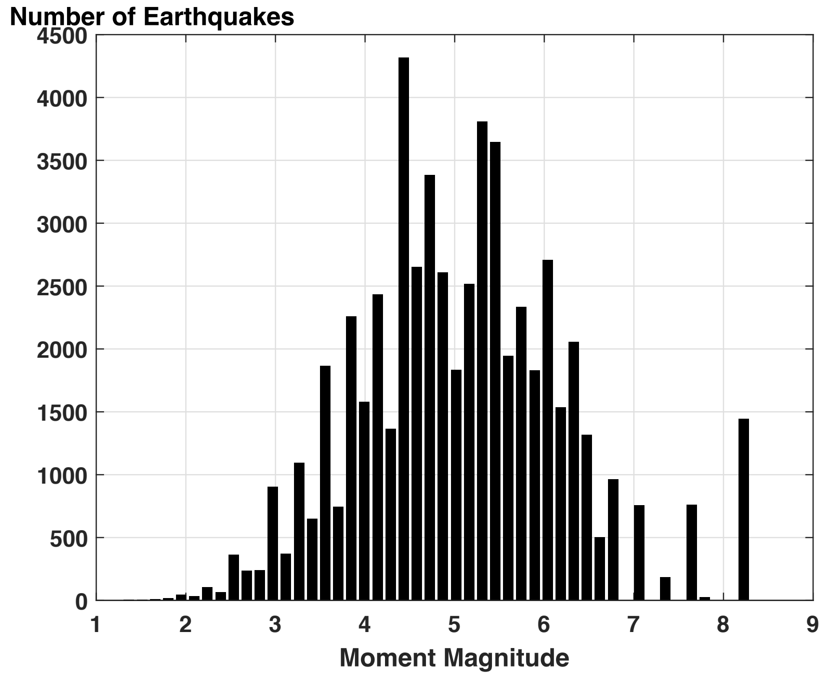

Figure 6.2.2: Distribution of Earthquake Magnitude in California

working with the random variable $X = -m_W$. Indeed, since a closed-form expression is generally not available for data-based distributions, the CEV is found by performing a γ-sweep and plotting $\mathbb{E}[X_\gamma]$; see Figure 6.2.3. Selecting the maximizing point $\gamma^* \approx -5.60$, we obtain

$$\mathbb{CEV}(X) = \mathbb{E}[X_\gamma^*] \approx -6.37.$$

In other words, our more conservative estimate for the expected earthquake intensity is 6.37 rather than 5.08. In addition we compute

$$P\big(m_W > 6.37\big) \approx 0.12.$$

In summary, for this mission-critical application which includes model distrust, our *conservative* approach leads to intensity exceeding the CEV only 12% versus 49% of the time using classical analysis.

6.2.5 Geometric Interpretation of CEV: In this subsection, a geometric interpretation of the new measure, $\mathbb{CEV}(X)$, is provided to show how it is related to classical expected value $\mathbb{E}[X]$. Without loss of generality, we assume that $\alpha_X = 0$ in the remainder of this subsection. Now given the random variable X and its cumulative distribution function $F_X(x)$, we begin with the

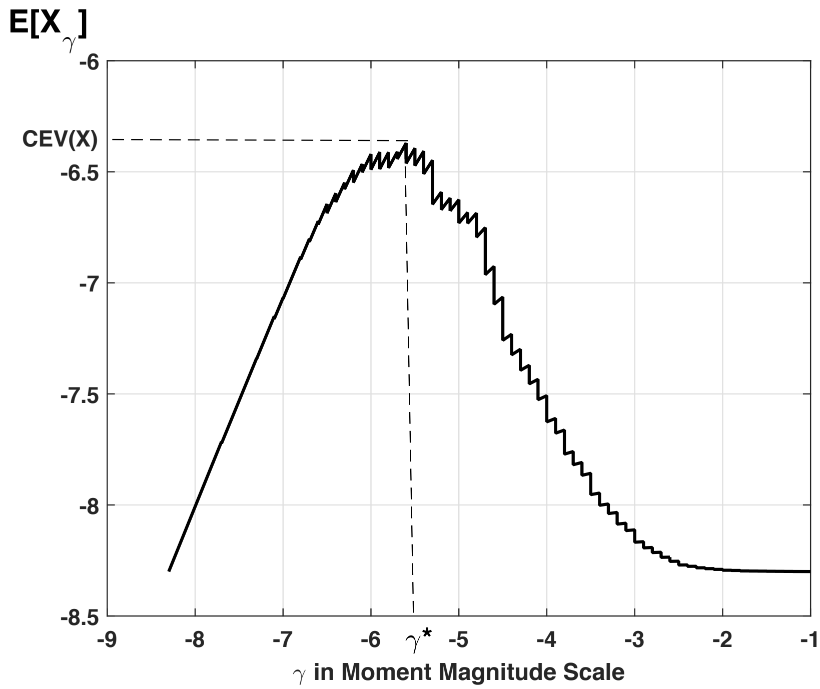

Figure 6.2.3: $\mathbb{E}[X_\gamma]$ Versus γ for the Earthquake Problem

well-known fact, for example, see [87], that

$$\mathbb{E}[X] = \int_0^\infty \bar{F}_X(x)dx,$$

where $\bar{F}_X(x) \doteq 1 - F_X(x)$ is the so-called *complementary cumulative distribution function*. The formula above implies that the area under the plot of $\bar{F}_X(x)$ gives $\mathbb{E}[X]$. Now expressing the CEV as

$$\mathbb{CEV}(X) = \sup_\gamma \gamma \bar{F}_X(\gamma),$$

we note that $\gamma \bar{F}_X(\gamma)$ is the area of a rectangle with base $[0, \gamma]$ and height $\bar{F}_X(\gamma)$. Since $\bar{F}_X(\gamma)$ is a non-increasing function of γ, this rectangle will be below the plot for $\bar{F}_X(\gamma)$. In other words, $\mathbb{CEV}(X)$ is obtained as the area of the largest possible "inscribed" rectangle under the complementary CDF.

This geometry is illustrated in Figure 6.2.4 for X uniformly distributed on $[0, 1]$ with $\bar{F}_X(x) = 1-x$ for $x \in [0, 1]$; the area of shaded rectangle is $\mathbb{CEV}(X) = 0.25$ while the sum of the shaded and striped area is $\mathbb{E}[X] = 0.5$. Similarly for an exponential distribution with complementary CDF $\bar{F}_X(x) = e^{-x}$ for $x \geq 0$, the maximally inscribed rectangle in Figure 6.2.5 has area $\mathbb{CEV}(X) = e^{-1}$ whereas $\mathbb{E}[X] = 1$.

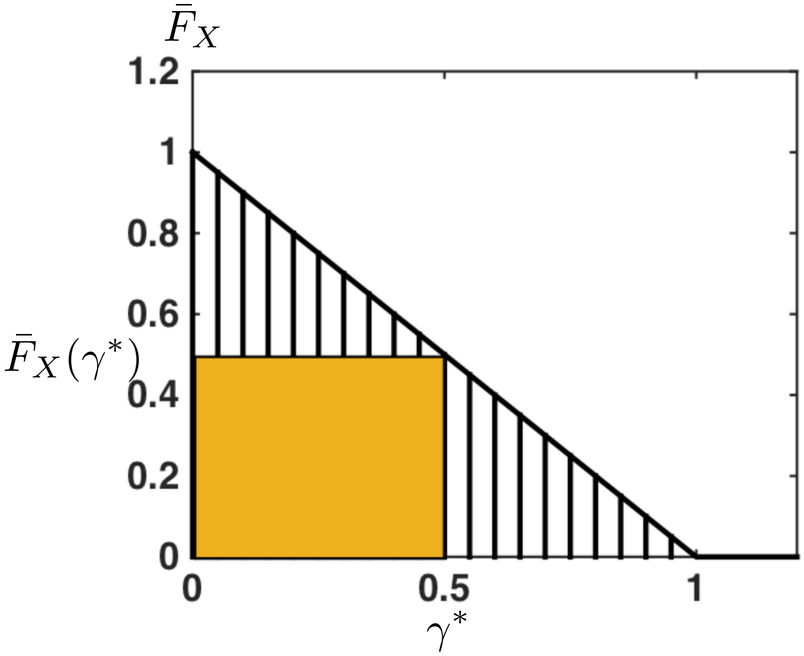

Figure 6.2.4: Expected Value (Area of Triangle) and CEV (Shaded Area) for Unif$[0, 1]$

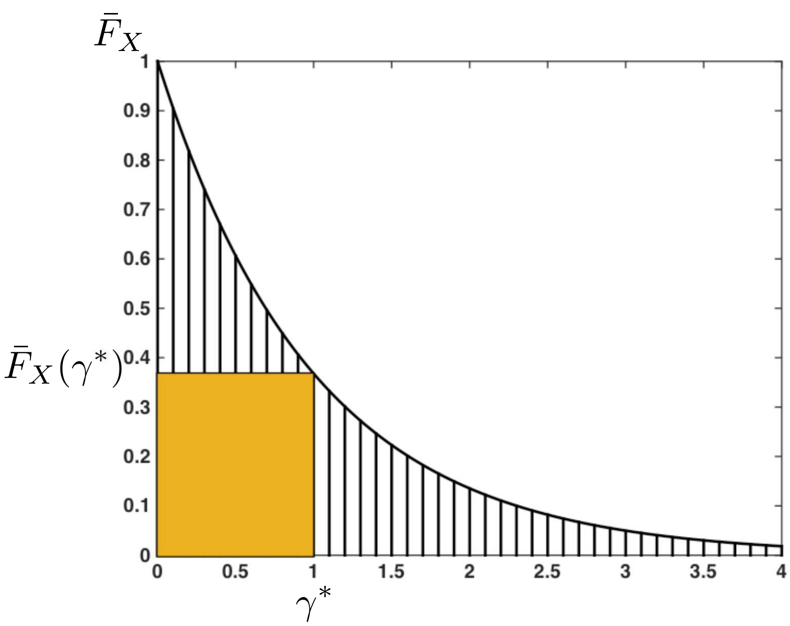

Figure 6.2.5: Expected Value (Area under Curve) and CEV (Shaded Area) for Exp(1)

6.3 Examples of CEV

In this section, the CEV is calculated for some classical probability distributions and results are compared with the classical expected value $\mathbb{E}[X]$. Beginning with the trivial case when X is itself a Bernoulli random variable, it follows immediately from the definition that $\mathbb{CEV}(X) = \mathbb{E}[X]$.

6.3.1 Uniform Distribution: For X uniformly distributed in $[0, 1]$, a straightforward calculation leads to

$$\mathbb{E}\left[X_\gamma\right] = \begin{cases} \gamma - \gamma^2; \quad 0 \leq \gamma \leq 1; \\ 0; \quad \gamma > 1. \end{cases}$$

Hence, $\mathbb{E}[X_\gamma]$ is maximized at $\gamma = 0.5$ and we obtain $\mathbb{CEV}(X) = 0.25$ which compares with classical expected value, $\mathbb{E}[X] = 0.5$. This generalizes to X uniformly distributed over $[\alpha_X, b]$. For this case, we obtain

$$\mathbb{CEV}(X) = \frac{3\alpha_X + b}{4}$$

which compares with $\mathbb{E}[X] = (\alpha_X + b)/2$.

6.3.2 Weibull Random Variable: For the random variable X, a generalization of the exponential distribution often used in reliability studies, having cumulative distribution function

$$F_X(x) = 1 - e^{-(\lambda x)^\alpha}$$

with $\alpha,\ \lambda > 0$ and for $x \geq 0$, a straightforward calculation leads to

$$\mathbb{E}[X_\gamma] = \gamma e^{-(\lambda\gamma)^\alpha}$$

for $\gamma \geq 0$. Then, the maximizer of $\mathbb{E}[X_\gamma]$, obtained by differentiation, is $\gamma^* = (1/\alpha)^{1/\alpha}/\lambda$, and this leads to

$$\mathbb{CEV}(X) = \frac{(1/\alpha)^{1/\alpha}}{\lambda} e^{-1/\alpha}.$$

Comparing this to classical expected value

$$\mathbb{E}[X] = \frac{1}{\lambda}\Gamma\Big(\frac{1}{\alpha} + 1\Big),$$

one can consider the *percentage discounting*, PD(X) of $\mathbb{E}[X]$; i.e.,

$$\text{PD}(X) \doteq \frac{\mathbb{E}[X] - \mathbb{CEV}(X)}{\mathbb{E}[X]} = 1 - \frac{(1/\alpha)^{1/\alpha} e^{-1/\alpha}}{\Gamma\Big(\frac{1}{\alpha} + 1\Big)}.$$

which is shown in Figure 6.3.1. The exponential random variable, obtained as a special case

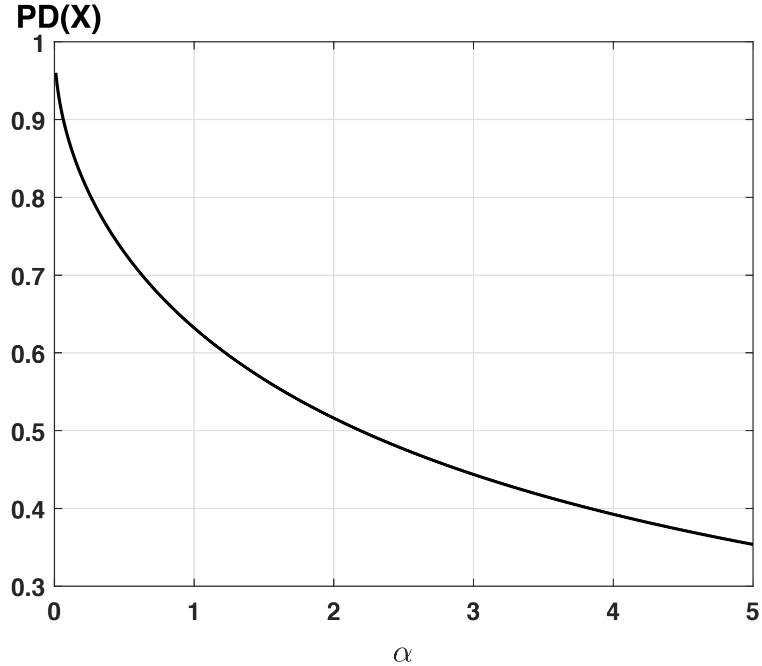

Figure 6.3.1: Percentage Discounting for Weibull Random Variable

with $\alpha = 1$ above, with skewness $S = 2$, has

$$\mathbb{CEV}(X) = \frac{e^{-1}}{\lambda},$$

which compares to $\mathbb{E}[X] = 1/\lambda$. A Rayleigh random variable, another special case, is now found with $\alpha = 2$ and $\lambda = 1/(\sqrt{2}\sigma)$ which has skewness $S = 0.63$. Via a lengthy calculation to maximize $\mathbb{E}[X_\gamma]$ we obtain

$$\mathbb{CEV}(X) = \frac{\sigma}{\sqrt{e}} \simeq 0.60\sigma$$

which compares with $\mathbb{E}[X] = \sqrt{\frac{\pi}{2}}\sigma \approx 1.25\sigma$.

6.3.3 Pareto Random Variable: For $\alpha_X > 0$, $\beta > 1$ and $x > \alpha_X$, beginning with

$$F_X(x) = 1 - \left(\frac{\alpha_X}{x}\right)^\beta,$$

with undefined skewness for $\beta \in [2, 3]$ and possibly high skewness given by

$$S = \frac{2(1+\beta)}{\beta - 3}\sqrt{\frac{\beta - 2}{\beta}}$$

for $\beta > 3$, we calculate

$$\mathbb{E}[X_\gamma] = \alpha_X \left[1 - \left(\frac{\alpha_X}{\gamma}\right)^\beta\right] + \gamma \left(\frac{\alpha_X}{\gamma}\right)^\beta.$$

Then, differentiation with respect to γ leads to the maximizer

$$\gamma^* = \frac{\beta\alpha_X}{\beta - 1},$$

and

$$\mathbb{CEV}(X) = \frac{\beta\alpha_X}{\beta - 1}\left[1 + \frac{\left(1 - \frac{1}{\beta}\right)^\beta - 1}{\beta}\right]$$

which compares with

$$\mathbb{E}[X] = \frac{\beta\alpha_X}{\beta - 1}.$$

6.3.4 Log-Normal Random Variable: With Φ being the cumulative distribution function for the standard normal random variable $\mathcal{N}(0, 1)$, we consider the random variable X having cumulative distribution function

$$F_X(x) = \Phi\Big(\frac{\ln x - \mu}{\sigma}\Big),$$

for $x \geq 0$, $\sigma > 0$ with skewness given by

$$S = (e^{\sigma^2} + 2)\sqrt{e^{\sigma^2} - 1}.$$

A straightforward calculation leads to

$$\mathbb{E}[X_\gamma] = \gamma\left(1 - \Phi\Big(\frac{\ln\gamma - \mu}{\sigma}\Big)\right)$$

whose supremum is found via a single-variable line-search over $\gamma \in\ [0, \infty)$. For example, when $\mu = \sigma = 1$, with skewness $S \approx 6.18$, this calculation leads to maximizer $\gamma^* \approx 3.59$ and $\mathbb{CEV}(X) \approx 1.40$ which compares to $\mathbb{E}[X] \approx 4.48$. For the special case when $\sigma = \sqrt{2/\pi}$, with skewness $S \approx 3.67$, the optimization problem can be solved analytically giving $\gamma^* = e^\mu$. Putting back to the formula for $\mathbb{E}[X_\gamma]$ yields to $\mathbb{CEV}(X) = 0.5e^\mu$ which compares to ordinary expected value $\mathbb{E}[X] = e^{1/\pi}e^\mu \approx 1.37e^\mu$. In this case, we see more than 63% discounting of the ordinary expected value.

6.3.5 Remark: For $a \geq 0$, it is easy to see that $\mathbb{CEV}(aX) = a\mathbb{CEV}(X)$; see also Subsection 6.6.3. Accordingly, suppose the CEV of $X_1 \sim \text{lognormal}(0, \sigma)$ is found through a line-search. Then, it is easy to show that $X \sim \text{lognormal}(\mu, \sigma)$ has $\mathbb{CEV}(X) = e^{\mu}\mathbb{CEV}(X_1)$.

6.4 Risk Considerations and Conservative Semi-Variance

Recalling in the introduction, we now describe a companion risk metric for the CEV which addresses one-sidedness; i.e., since large positive values of X are desirable, we only penalize downside variations. This concept of one-sidedness arises in many applications, most notably in the context of finance where risk metrics are studied; e.g., see [105] and [118]. Perhaps the most popular risk metric which is common in many fields such as engineering, finance and statistics is the *variance*; e.g., see [15] and [124]. For problems with asymmetric probability distributions, it is standard in the literature to consider higher order moments; e.g., see [123] and [99]. Since the variance is "blind" to the direction of the variation, in many applications such as finance where one-sided deviations from the expected value can be good, it is standard to use a *semi-variance measure*; e.g., see [90] and [91]. This motivates the definition below.

6.4.1 Conservative Semi-Variance Definition: Given a target value γ, the *semi-variance* for the random variable X is defined as

$$SV(X,\gamma) \doteq \int_{x<\gamma} (x-\gamma)^2 f_X(x)dx = \mathbb{E}\big[(X-\gamma)^2 \mathbf{1}_{\{X<\gamma\}}\big],$$

where the "$\mathbf{1}$" denotes the indicator function

$$\mathbf{1}_{\{X<\gamma\}}(x) \doteq \begin{cases} 0 & \text{if} \quad X \geq \gamma; \\ 1 & \text{if} \quad X < \gamma. \end{cases}$$

The semi-variance looks at the downside variance; i.e., the deviation in "bad direction." Given the previously defined CEV, we make the natural choice for $\gamma = \mathbb{CEV}(X)$ above and obtain *the Conservative Semi-Variance*;

$$\mathbb{CSV}(X) \doteq SV\big(X, \mathbb{CEV}(X)\big).$$

6.4.2 Remark: Our view is that a "worse-than-expected" outcome is one which is below $\mathbb{CEV}(X)$. Hence, variation below this quantity counts toward the risk. The definition above is further substantiated by the fact that the $SV(X,\gamma)$ is an increasing function of γ. Accordingly, the choice of $\gamma=\gamma^*$ maximizes the $SV(X,\mathbb{E}[X_\gamma])$ and therefore results $\mathbb{CSV}(X)$ a *conservative* risk assessment.

6.5 Examples of CSV

In this section, the CSV is calculated and compared with the classical variance for the probability distributions considered earlier in Section 6.3. Beginning again with the trivial case when X is itself a $0-1$ Bernoulli random variable $P(X=1)=p$, we obtain $\mathbb{CSV}(X)=p(1-p)^2$ which compares with var$(X)=p(1-p)$. Recalling that $\mathbb{CEV}(X)=\mathbb{E}[X]$ for this extreme case, we have $\mathbb{CSV}(X)<$ var(X) due to the one-sided emphasis on downside risk.

6.5.1 Uniform Distribution: For random variable X uniformly distributed on $[0,1]$, starting with $\mathbb{CEV}(X)=0.25$ previously found, we obtain

$$\mathbb{CSV}(X)=\int_0^{\mathbb{CEV}(X)}\big(x-\mathbb{CEV}(X)\big)^2dx=\frac{1}{192},$$

which compares with var$(X)=1/12$. For the more general case with X uniformly distributed over $[\alpha_X,b]$, we obtain

$$\mathbb{CSV}(X)=\frac{(b-\alpha_X)^2}{192},$$

which compares with var$(X)=(b-\alpha_X)^2/12$.

6.5.2 Weibull Random Variable: For the Weibull random variable X of Subsection 6.3.2, we now use parameters $\alpha=\lambda=2$. Beginning with $\mathbb{CEV}(X)\approx 0.21$, a straightforward calculation leads to $\mathbb{CSV}(X)\approx 0.001$, which compares with var$(X)\approx 0.05$. For the special case with $\alpha=1$, X is an exponential random variable and $\mathbb{CEV}(X)=e^{-1}/\lambda$. Then, writing the definition of CSV gives

$$\mathbb{CSV}(X)=\int_0^{e^{-1}/\lambda}\big(x-\frac{e^{-1}}{\lambda}\big)^2\lambda e^{-\lambda x}dx\approx\frac{0.015}{\lambda^2}.$$

which is comparable to var$(X)=1/\lambda^2$. Another special case, called Rayleigh distribution, is obtained with $\alpha=2$ and $\lambda=1/(\sqrt{2}\sigma)$. Now beginning with $\mathbb{CEV}(X)\approx 0.60\sigma$, a lengthy calculation yields

$$\mathbb{CSV}(X) = \sigma^2\Big(e^{-1} - 2e^{-1/(2e)} - 2\sqrt{\frac{2\pi}{e}}\phi(e^{-0.5}) + \sqrt{\frac{2\pi}{e}} + 2\Big) \simeq 0.011\sigma^2,$$

which is comparable to variance $\text{var}(X) \simeq 0.43\sigma^2$.

6.5.3 Pareto Random Variable: For the Pareto random variable with $\alpha_X > 0$ and $\beta > 1$, beginning with $\mathbb{CSV}(X)$ obtained in Subsection 6.3.3, a lengthy calculation leads to

$$\mathbb{CSV}(X) = \begin{cases} (2\log\frac{5}{4} - \frac{7}{32})\alpha_X^2 \approx 0.01\alpha_X^2 & \text{for} \quad \beta = 2, \\ \Big[\frac{-2(r^{2-\beta}-1)}{\beta(1-\beta)(2-\beta)} + \frac{(r-1)^2}{\beta} + \frac{2(r-1)}{\beta(1-\beta)}\Big]\beta\alpha_X^2 & \text{for} \quad \beta \in (1,\infty)\setminus\{2\}. \end{cases}$$

where

$$r \doteq \frac{\beta}{\beta-1}[1 + \frac{-1 + \big(1-\frac{1}{\beta}\big)^\beta}{\beta}],$$

and CSV compares to

$$\text{var}(X) = \begin{cases} \infty & \text{for} \quad \beta \in (1,2], \\ \frac{\beta\alpha_X^2}{(\beta-1)^2(\beta-2)} & \text{for} \quad \beta > 2. \end{cases}$$

In Figure 6.5.1, the CSV is compared with the variance when $\alpha_X = 1$ and β varying.

6.5.4 Log-Normal Random Variable: For the log-normal random variable $X \sim$lognormal(μ,σ) in Subsection 6.3.4, using $\sigma = \sqrt{2/\pi}$ for illustrative purposes, a lengthy calculation of CSV yields $\mathbb{CSV}(X) \approx 0.007e^{2\mu}$ which compares to $\text{var}(X) \simeq 1.6823e^{2\mu}$. Similar to the remark in Subsection 6.3.4, it is trivial to see that for $a \geq 0$, we have $\mathbb{CSV}(aX) = a^2\mathbb{CSV}(X)$, see also Subsection 6.6.3. Hence for the more general random variable $X \sim$ lognormal(μ,σ), we obtain $\mathbb{CSV}(X) = e^{2\mu}\mathbb{CSV}(X_1)$ where $X_1 \sim$ lognormal$(0,\sigma)$.

6.6 Properties of the CEV and CSV

In this section, a number of basic properties of the CEV and CSV are established.

6.6.1 Bounds on the CEV and CSV: In the lemma below, simple bounds on the $\mathbb{CEV}(X)$ and $\mathbb{CSV}(X)$ are given. Regarding the tightness of the given bounds, one can verify that the lower bound on CEV is achieved by a uniform random variable and upper bound is achieved by a Bernoulli random variable.

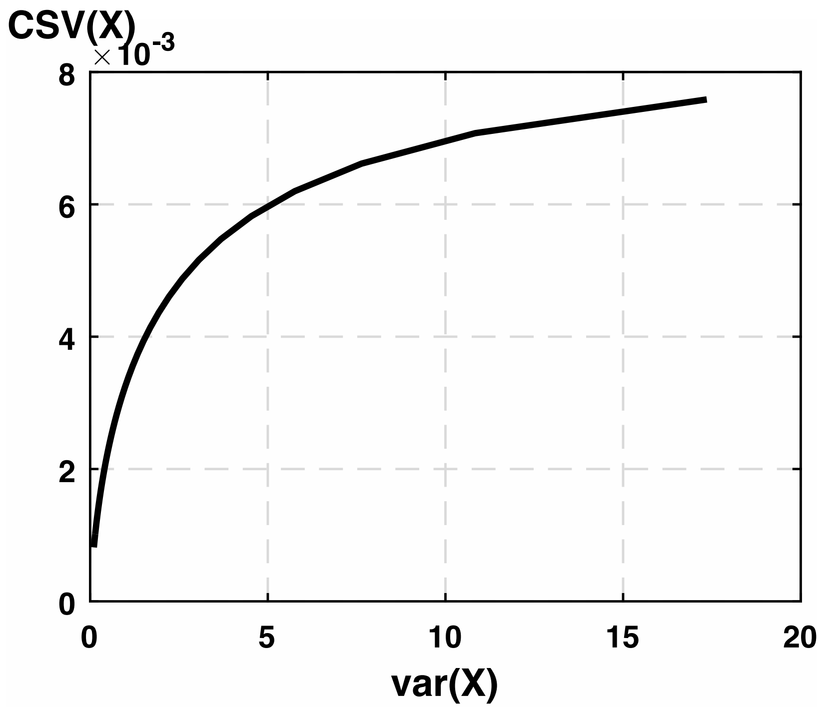

Figure 6.5.1: $\mathbb{CSV}(X)$ versus var(X) for Pareto Random Variable with $\alpha_X = 1$

6.6.2 Lemma: *Let X be a random variable with finite leftmost support point α_X. Then*

$$\frac{\text{median}(X) + \alpha_X}{2} \leq \mathbb{CEV}\,(X) \leq \mathbb{E}\,(X)\,,$$

and

$$0 \leq \mathbb{CSV}\,(X) \leq \text{var}(X).$$

Proof: The upper bound on $\mathbb{CEV}(X)$ was obtained in Section 6.2. The lower bound is obtained using the special choice $\gamma = \text{median}\,(X)$ in combination with the fact that $F_X(\gamma) = 0.5$ for this case. The lower bound on the CSV follows from the definition. For the upper bound, we introduce the function

$$g(\gamma) \doteq \int_{\alpha_X}^{\gamma} (x-\gamma)^2 f_X(x)dx$$

and take the derivative to obtain

$$\frac{dg}{d\gamma} = 2\int_{\alpha_X}^{\gamma} (\gamma - x) f_X(x)dx \geq 0.$$

Now, since $\mathbb{CEV}(X) \leq \mu \doteq \mathbb{E}[X]$, we can write

$$\mathbb{CSV}(X) = g\big(\mathbb{CEV}(X)\big) \leq g(\mu),$$

and noting that

$$g(\mu) = \int_{\alpha_X}^{\mu} (x-\mu)^2 f_X(x)dx \leq \int_{\alpha_X}^{\infty} (x-\mu)^2 f_X(x)dx = \text{var}(X),$$

the proof is complete. □

6.6.3 Affine Linearity: In the following lemma, it is shown that $\mathbb{CEV}(X)$ has an affine linearity property.

6.6.4 Lemma: *Given constants $a \geq 0$ and b, for a random variable X with finite leftmost support point α_X, the CEV satisfies*

$$\mathbb{CEV}(aX+b) = a\mathbb{CEV}(X) + b$$

and the CSV satisfies

$$\mathbb{CSV}(aX+b) = a^2\mathbb{CSV}(X).$$

Proof: It suffices to prove that $\mathbb{CEV}(aX) = a\mathbb{CEV}(X)$ and $\mathbb{CEV}(X+b) = \mathbb{CEV}(X) + b$. Indeed for random variable $Y \doteq aX$, noting that the left support point is $\alpha_Y = a\alpha_X$;

$$\begin{aligned}
\mathbb{CEV}(Y) &\doteq \sup_{\gamma} \mathbb{E}(Y_\gamma) \\
&= \sup_{\gamma} \{\gamma + (\alpha_Y - \gamma) F_Y(\gamma)\} \\
&= \sup_{\gamma} \{\gamma + (a\alpha_X - \gamma) F_X(\gamma/a)\} \\
&= \sup_{\theta} \{a\theta + a(\alpha_X - \theta) F_X(\theta)\} \\
&= a\mathbb{CEV}(X).
\end{aligned}$$

For the second part of the proof, we now take $Y \doteq X + b$ with leftmost support point $\alpha_Y = \alpha_X + b$

$$\begin{aligned}
\mathbb{CEV}(Y) &\doteq \sup_{\gamma} \mathbb{E}(Y_\gamma) \\
&= \sup_{\gamma} \{\gamma + (\alpha_Y - \gamma) F_Y(\gamma)\} \\
&= \sup_{\gamma} \{\gamma + (\alpha_X + b - \gamma) F_X(\gamma - b)\} \\
&= \sup_{\theta} \{\theta + (\alpha_X - \theta) F_X(\theta)\} + b \\
&= \mathbb{CEV}(X) + b.
\end{aligned}$$

For the CSV, starting with the definition and using the affine linearity of CEV above,

$$\begin{aligned}\mathbb{CSV}(aX+b) &= \mathbb{E}\Big[\big(aX+b-\mathbb{CEV}(aX+b)\big)^2\mathbf{1}_{\{aX+b<\mathbb{CEV}(aX+b)\}}\Big]\\ &= \mathbb{E}\Big[a^2\big(X-\mathbb{CEV}(X)\big)^2\mathbf{1}_{\{X<\mathbb{CEV}(X)\}}\Big]\\ &= a^2\mathbb{CSV}(X).\quad\square\end{aligned}$$

6.6.5 Remark: Using the affine linearity property for the CEV, it is easy to see that the percentage discounting, introduced in Section 6.3.2, is invariant to scaling; that is, for a random variable X and any scalar $a>0$

$$\text{PD}(aX)=\frac{\mathbb{E}[aX]-\mathbb{CEV}(aX)}{\mathbb{E}[aX]}=\frac{a\mathbb{E}[X]-a\mathbb{CEV}(X)}{a\mathbb{E}[X]}=\text{PD}(X).$$

This result is consistent with the fact that the skewness is also independent of scaling. Two examples illustrating this invariance are the Weibull random variable considered in Subsection 6.3.2 with scaling parameter λ and the Pareto random variable in Subsection 6.3.3 with scaling factor α_X. Finally, if random variable X is log-normally distributed, see Subsection 6.3.4, then the percentage discounting is a function of σ only and is independent of the parameter μ.

6.6.6 Average of i.i.d Random Variables: In the theorem to follow, we consider the average $\mathcal{X}_n$ of n independent and identically distributed (i.i.d.) random variables, and show that the $\mathbb{CEV}(\mathcal{X}_n)$ tends to the common expected value, μ, as $n\to\infty$.

6.6.7 Theorem: *For positive integers k, let X_k be a sequence of i.i.d. random variables with finite mean $\mathbb{E}\left(X_k\right)= \mu$, finite variance σ^2 and finite leftmost support point, $\alpha_{X_k}=\alpha_X$. Then, with partial sum averages given by*

$$\mathcal{X}_n\doteq\frac{1}{n}\sum_{k=1}^{n}X_k,$$

it follows that

$$\lim_{n\to\infty}\mathbb{CEV}(\mathcal{X}_n)=\mu$$

and

$$\lim_{n\to\infty}\mathbb{CSV}(\mathcal{X}_n)=0.$$

Proof: For each n, we first note that α_X must be the leftmost support point of $\mathcal{X}_n$. Now, noting Lemma 6.6.4, since $\mathbb{CEV}(\mathcal{X}_n - \alpha_X) = \mathbb{CEV}(\mathcal{X}_n) - \alpha_X$, without loss of generality, we assume that $\alpha_X = 0$ and $\mu \geq 0$ in the remainder of the proof. Now, along the sequence $\mathcal{X}_n$, recalling Lemma 6.6.2,

$$\mathbb{CEV}(\mathcal{X}_n) \leq \mathbb{E}(\mathcal{X}_n) = \mu.$$

Next, we construct a lower bound for $\mathbb{CEV}(\mathcal{X}_n)$ using a one-sided Chebyshev inequality. Indeed, since $\mathcal{X}_n$ has finite mean μ and bounded variance $\sigma_n^2 = \sigma^2/n$, for $\epsilon > 0$ and each n, the Chebyshev inequality

$$P(\mathcal{X}_n \leq (1-\epsilon)\mu) \leq \frac{\sigma_n^2}{\sigma_n^2 + \epsilon^2\mu^2}$$

is satisfied. Hence, for any $\gamma \in [0, \mu)$, letting $\epsilon = (\mu - \gamma)/\mu$ and noting that $\epsilon > 0$, via the inequality above, we obtain

$$P(\mathcal{X}_n > \gamma) \geq \frac{(\mu-\gamma)^2}{\sigma_n^2 + (\mu-\gamma)^2}.$$

Using this inequality leads to a lower bound for the CEV; i.e.,

$$\mathbb{CEV}(\mathcal{X}_n) = \sup_{\gamma} \gamma P(\mathcal{X}_n > \gamma) \geq \sup_{\gamma \in [0,\mu)} \gamma \frac{(\mu-\gamma)^2}{\sigma_n^2 + (\mu-\gamma)^2}.$$

For large enough n such that $\mu > (1/n)^{0.25}$, for the specific choice $\gamma = \mu - (1/n)^{0.25}$,

$$\sup_{\gamma} \gamma \frac{(\mu-\gamma)^2}{\sigma_n^2 + (\mu-\gamma)^2} \geq \left(\mu - \left(\frac{1}{n}\right)^{0.25}\right) \cdot \frac{\frac{1}{\sqrt{n}}}{\sigma_n^2 + \frac{1}{\sqrt{n}}}$$

Since μ is an upper bound for $\mathbb{CEV}(\mathcal{X}_n)$ and further noting that $\sigma_n^2 = \sigma^2/n$; for large enough n, we can combine the inequalities above to obtain

$$\mu \geq \mathbb{CEV}(\mathcal{X}_n) \geq \left(\mu - \left(\frac{1}{n}\right)^{0.25}\right) \frac{1}{\frac{\sigma^2}{\sqrt{n}} + 1}.$$

Now letting $n \to \infty$, it is easy to show that the right-hand side tends to μ and hence,

$$\lim_{n \to \infty} \mathbb{CEV}(\mathcal{X}_n) = \mu.$$

The proof of the second part is straightforward. Using the bounds in Lemma 6.6.2,

$$0 \leq \mathbb{CSV}(\mathcal{X}_n) \leq \text{var}(\mathcal{X}_n) = \frac{\sigma^2}{n}.$$

Now letting $n \to \infty$, this inequality forces $\mathbb{CSV}(\mathcal{X}_n) \to 0$. $\square$

6.6.8 Convexity Property of CEV: Consider a random variable X whose probability density function is a convex combination of the probability density functions of n random variables, $X_1, X_2, \ldots, X_n$; i.e.,

$$f_X(x) = \sum_{i=1}^{n} \lambda_i f_{X_i}(x)$$

where $\lambda_i \geq 0$, $\sum_{i=1}^{n} \lambda_i = 1$ and $f_{X_i}(x)$ is the probability density function for X_i. To illustrate how the situation above arises, consider the case for the random variable describing the output of a system which can switch among n different states. Suppose, the state is modelled by a random variable θ such that $P(\theta = i) = \lambda_i$, for values of $i = 1, 2, \ldots, n$ and further assume that the output of the system, X, conditioned on the state is modelled by a set of random variables X_i; that is,

$$f_X(x|\theta = i) \doteq f_{X_i}(x).$$

This implies that, the probability density function for X is a convex combination of the f_{X_i} given above. In the lemma below, an upper bound on the CEV of X is given in terms of the convex combination of the $\mathbb{CEV}(X_i)$.

6.6.9 Lemma: *Let the probability density function f_X of the random variable X be the convex combination of the probability density functions f_{X_i} of the n random variables $X_1, X_2, \ldots, X_n$. Then X has a conservative expected value satisfying*

$$\mathbb{CEV}(X) \leq \sum_{i=1}^{n} \lambda_i \mathbb{CEV}(X_i).$$

Proof: Without loss of generality, we assume that $\alpha_{X_1} \leq \alpha_{X_2} \leq \ldots \leq \alpha_{X_n}$. Using the definition of X, it is easy to show that $\alpha_X = \alpha_{X_1}$. Now we calculate

$$\begin{aligned}
\mathbb{CEV}(X) &= \sup_{\gamma} \gamma + (\alpha_X - \gamma) F_X(\gamma) \\
&= \sup_{\gamma} \gamma + (\alpha_X - \gamma) \sum_{i=1}^{n} \lambda_i F_{X_i}(\gamma) \\
&\leq \sum_{i=1}^{n} \sup_{\gamma} \lambda_i (\gamma + (\alpha_{X_i} - \gamma) F_{X_i}) \\
&= \sum_{i=1}^{n} \lambda_i CEV(X_i). \;\; \square
\end{aligned}$$

6.6.10 Finiteness of the CEV: To begin, we note that the inequality $\mathbb{CEV}(X) \leq \mathbb{E}[X]$ allows for the possibility that $\mathbb{CEV}(X)$ can be finite even though $\mathbb{E}[X]$ is infinite. This is indeed the case for the well-known St. Petersburg Paradox discussed in Section 1. For the random variable X with probability density function given by $X = 2^k$ with probability $p = 1/2^{k+1}$ for non-negative integers k, we recall that $\mathbb{E}[X] = \infty$ even though studies suggest that \$3 is the "ticket price" one might reasonably pay to play this game; see [131]. The shortcoming of expected value is due to the fat tail of the probability distribution. We now argue that the CEV provides more realistic assessment of the situation at hand.

Now, to calculate $\mathbb{CEV}(X)$, for $k \geq 0$ and $\gamma \in [2^k, 2^{k+1})$, noting that $\alpha_X = 1$, a simple calculation yields

$$\mathbb{E}[X_\gamma] = 1 + \frac{\gamma - 1}{2^{k+1}}$$

We see that $\mathbb{E}[X_\gamma]$ varies linearly from $1.5 - 1/2^{k+1}$ to $2 - 1/2^{k+1}$. It is now easy to show that $\mathbb{CEV}(X) = 2$ is attained as γ tends to infinity. In addition it is also straightforward to obtain the $\mathbb{CSV}(X) = 0.5$.

In view of the example above, it is natural to ask if the CEV can ever be infinite. Indeed the answer is "yes" and this is illustrated by the random variable X with probability density function $f_X(x) = 1/(2x^{\frac{3}{2}})$ for $x \geq 1$. For $\gamma \geq 1$, straightforward calculation leads to

$$\mathbb{E}[X_\gamma] = \sqrt{\gamma} - \frac{1}{\sqrt{\gamma}} - 1,$$

Now as $\gamma \to \infty$, we obtain $\mathbb{E}[X_\gamma] \to \infty$ which implies that $\mathbb{CEV}(X) = \infty$.

6.6.11 CEV and Improvement of Markov Inequality: For a non-negative random variable X and a given $\epsilon > 0$, the classical Markov inequality tells us that

$$P(X > \epsilon) \leq \frac{\mathbb{E}[X]}{\epsilon}.$$

In the lemma below, we see that this bound still holds with $\mathbb{E}[X]$ replaced by $\mathbb{CEV}(X)$. Since $\mathbb{CEV}(X) \leq \mathbb{E}[X]$, our new bound is tighter than the Markov inequality.

6.6.12 Lemma: *For a non-negative random variable X and given constant $\epsilon > 0$, it follows that*

$$P(X > \epsilon) \leq \frac{\mathbb{CEV}(X)}{\epsilon}.$$

Proof: Since the result trivially holds for $\epsilon \leq \mathbb{CEV}(X)$, we henceforth assume that $\epsilon \geq \mathbb{CEV}(X)$. For simplicity, we first consider the case where $\alpha_X = 0$ and later generalize to $\alpha_X \geq 0$. Now since $\alpha_X = 0$, for any $\epsilon > 0$

$$\mathbb{CEV}(X) = \sup_{\gamma} \gamma P(X > \gamma) \geq \epsilon P(X > \epsilon).$$

Equivalently,

$$P(X > \epsilon) \leq \frac{\mathbb{CEV}(X)}{\epsilon}$$

which completes the proof. Now, more generally, if $\alpha_X > 0$, by defining $Y \doteq X - \alpha_X$, for $\epsilon > \mathbb{CEV}(X)$, we will have

$$P(X > \epsilon) = P(Y > \epsilon - \alpha_X) \leq \frac{\mathbb{CEV}(Y)}{\epsilon - \alpha_X} = \frac{\mathbb{CEV}(X) - \alpha_X}{\epsilon - \alpha_X} \leq \frac{\mathbb{CEV}(X)}{\epsilon},$$

where the last inequality holds because $\alpha_X > 0$ and

$$f(x) \doteq \frac{\mathbb{CEV}(X) - x}{\epsilon - x}$$

is a decreasing function of x when $\epsilon > \mathbb{CEV}(X)$. $\square$

6.6.13 Remark: The same approach can be taken to improve the upper bound given by Chebyshev inequality, a generalization of the Markov inequality. That is; for a given random variable X with expected value μ, it can be shown that

$$P(|X - \mu| > \epsilon) \leq \frac{\mathbb{CEV}[(X - \mu)^2]}{\epsilon^2}.$$

In Figures 6.6.1-6.6.3, the actual probability and the upper bounds given by Markov inequality and the one introduced above is shown for different random variables. The tightness of the new bound is evidenced by the fact that it touches the actual probability at a point.

The improved bound above can also be related to the Percentage Discounting PD(X) introduced in Section 6.3. Letting $\epsilon = \mathbb{E}[X]$ and assuming $\alpha_X = 0$ for simplicity, we observe that

$$P(X > \mathbb{E}[X]) \leq \frac{\mathbb{CEV}(X)}{\mathbb{E}[X]} = 1 - \text{PD}(X).$$

This inequality implies that if the discounting in calculation of $\mathbb{CEV}(X)$ is heavy; that is if PD$(X) \to 1$, then the probability of the random variable X exceeding the expected value is very small. That is, expected value $\mathbb{E}[X]$ is an unduly optimistic indicator.

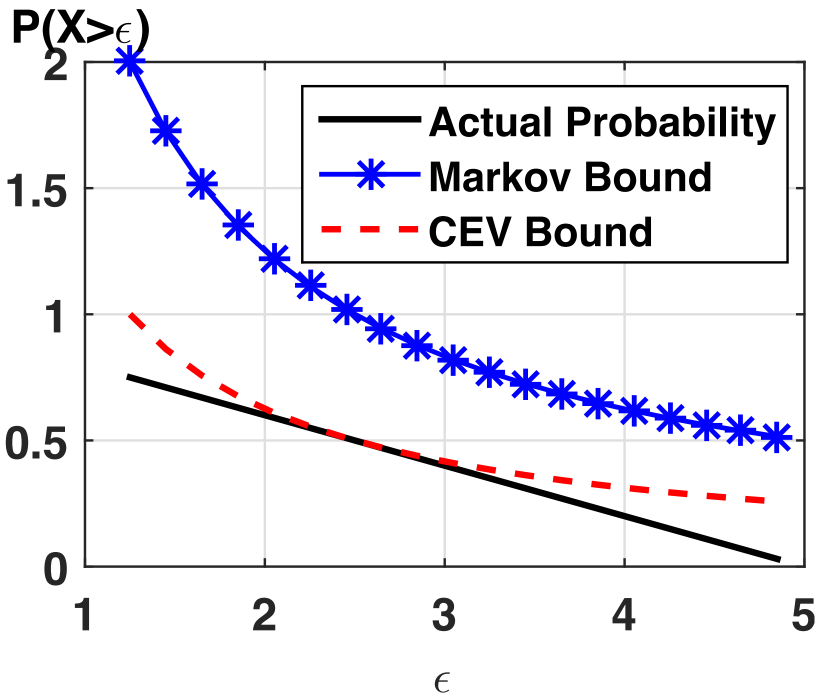

Figure 6.6.1: The Bounds on $P(X > \epsilon)$ when X Uniformly Distributed

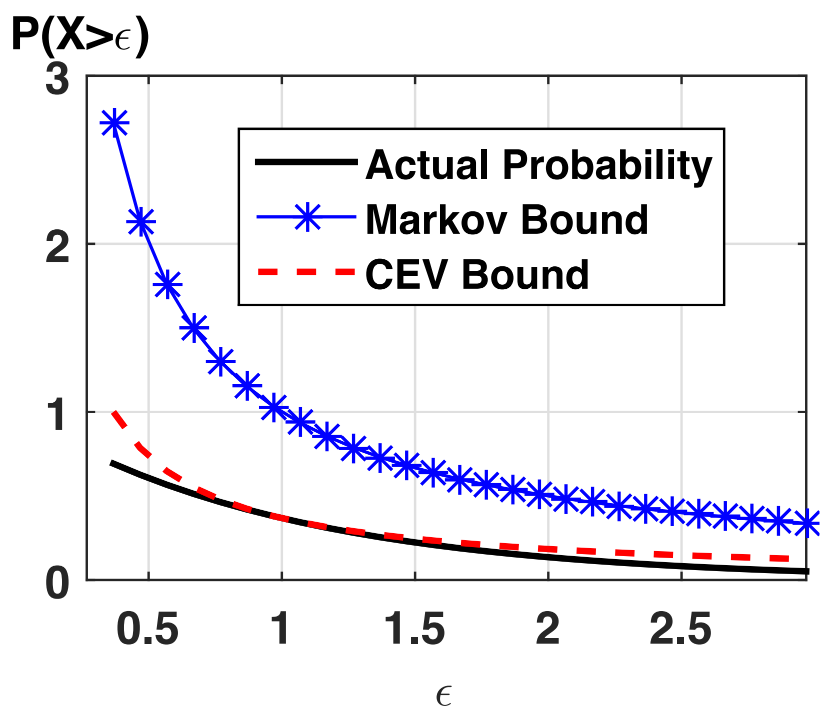

Figure 6.6.2: The Bounds on $P(X > \epsilon)$ when X is Exponential with $\lambda = 1$

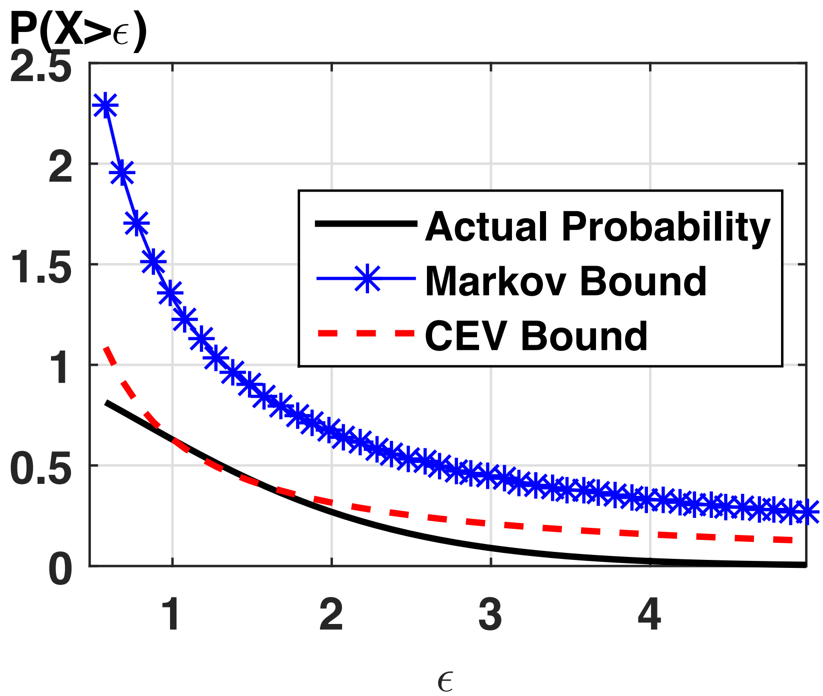

Figure 6.6.3: The Bounds on $P(X > \epsilon)$ when X is Weibull$(4, 1.5)$

6.6.14 CEV and CSV as Risk Metrics: For a given random variable X, the function $\rho(X)$ defined as $\rho(X) \doteq -\mathbb{CEV}(X)$ satisfies the conditions to be a risk measure; that is, $\rho(0) = 0$, $\rho(X + r) = \rho(X) - r$ for all $r \in \mathbb{R}$ and for any two random variable X_1 and X_2 such that $X_1 \leq X_2$ then $\rho(X_2) \leq \rho(X_1)$. However, similar to widely-used measures such as Value-at-Risk, it is not sub-additive and hence not *coherent*; see [105] for details. That is, the inequality $\rho(X + Y) \leq \rho(X) + \rho(Y)$ fails to be satisfied for some combinations of random variables X and Y; e.g., take X and Y both to be $0 - 1$ Bernoulli random variable with probability $p = 0.5$.

Another possibility for a metric is $D(X) \doteq \sqrt{\mathbb{CSV}(X)}$, which we call it *Conservative Semi-Deviation*, which satisfies all but one of the conditions to be a *Deviation Risk Measure*; see [139] for details. More specifically, we have $D(X + r) = D(X)$ for any $r \in \mathbb{R}$, $D(\lambda X) = \lambda D(X)$ for any $\lambda > 0$, $D(0) = 0$, and $D(X) > 0$ for all non-constant X while $D(X) = 0$ for any constant X. Although, simple examples are available to show that sub-additivity can not be guaranteed, it is interesting to note that the sub-additivity inequality for $D(X)$ using a *specific* pair (X, Y) is satisfied whenever the same inequality is violated for $\rho(X)$.

6.7 Potential for Application

In this section, to further demonstrate potential for application for our new theory, we provide three examples. The first example is motivated by the "skewing effects" related to the use of feedback when trading in financial markets as described in Chapter 2; see [6] and [7] for more details. In Subsection 6.7.1, we compare the classical expected value with the CEV of the resulting gain-loss function $g(t)$ when the price benchmark is the Geometric Brownian Motion.

The other two examples involve real data which are rather skewed. Working with such skewed distributions arises in different applications such as system reliability [135], finance [7], statistics [149] and psychology [150] where a similar approach based on CEV and CSV may be helpful to carry out the analysis. Per discussion in Section 6.2.3, since a closed-form expression is generally not available for data-based distributions, the CEV and CSV are found using a γ-sweep which is straightforward to implement.

6.7.1 Example: Use of Linear Feedback: Given the discussion in Chapter 2 and earlier in this chapter, suffice it to say, when a feedback control is used to modify an investment position, the resulting probability distribution for profits and losses can be highly skewed. For example, if $K > 0$ is the gain of a linear stock-trading controller, the resulting skewness $S(K)$ for profits and losses can increase dramatically with K and can easily become so large as to render many existing forms of risk-return analysis of questionable worth. Said another way, the long tail of the resulting highly-skewed distribution can lead to a large expected profit but the probability of an "adequate" profit may be quite small. Another negative associated with high skew is that there can be a significant probability of large drawdown in an investor's account; e.g., see Chapter 3 and reference [8].

To provide a concrete illustration of the issues raised above, we consider the following linear feedback controller described in Section 1.4.2. The amount invested $I(t)$, at time t, is given by

$$I(t) = I_0 + Kg(t),$$

where I_0 is the initial investment, K is the feedback gain and $g(t)$ is the cumulative gain-loss up to time t. When Geometric Brownian Motion (GBM) is used to drive the stock prices, the random variable $g(t)$ turns out to be a shifted and scaled log-normal distribution which can be highly skewed with an expected value which may be misleading in terms of the prospect for success.

To illustrate the scenario above, suppose time $t = 1$ represents one year and assume GBM process parameters $\mu = 0.25$ and $\sigma = 0.5$, where μ is the annualized drift and σ is the annualized volatility. Furthermore, assume initial investment $I_0 = 1$ representing one dollar and feedback gain $K = 4$. Then, via a simple modification of the results in [4] and Section 1.4.2, the probability density function for the gains and losses, $g(t)$, at $t = 1$, is given by

$$f(x) = \frac{1}{\sqrt{\frac{\pi}{2}}(1+4x)}\, e^{-\frac{\left(\log(1+4x)+1\right)^2}{8}}$$

for $x > -0.25$; see Figure 6.7.1.

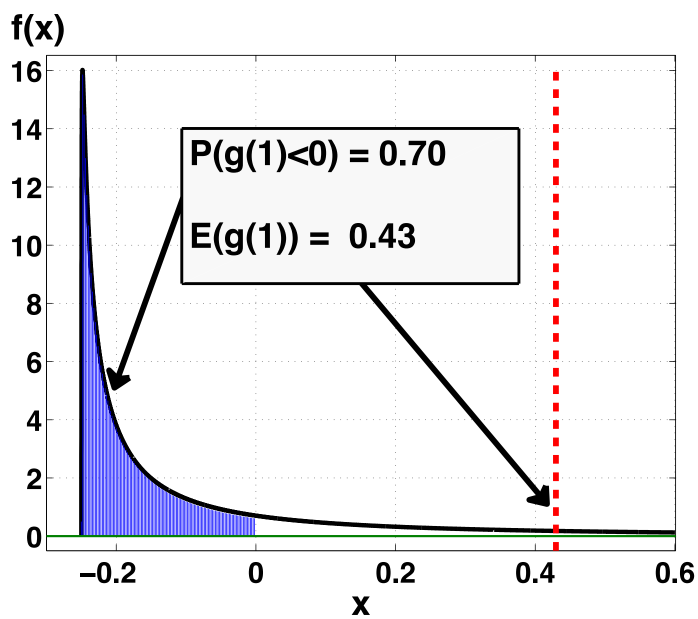

Figure 6.7.1: Trading Profit-Loss: The Probability Density Function

As seen in this figure, the expected value is $E[g(1)] \approx 0.43$ which is shown via the vertical dashed line. This expected value represents a raw return of 43% on an investment of one dollar. However, as seen in the figure, the probability of loss, the shaded area, is $p_{\text{LOSS}} \approx 0.70$. In other words, the expected return is quite attractive but it is highly probable that a losing trade will occur. To this end, the CEV is obtained to be $\mathbb{CEV}\big(g(1)\big) \approx -0.12$. This negative value indicates an expectation of loss from a conservative perspective. Comparing CEV to the classical expected value, $\mathbb{E}[g(1)] \approx 0.43$, shows how the long tail of the distribution is discounted. On the risk side, the Conservative Semi-Deviation is obtained to be $D\big(g(1)\big) \approx 0.07$ while the standard deviation is approximately 4.97.

6.7.2 Example: Annual Return for Standard and Poor's 500 Stock Index: As a second example, we begin with the empirical distribution of the annual percentage return for the S&P 500 stock index. The histogram in Figure 6.7.2 is obtained using the dividend-adjusted daily closing values, available at [151], in conjunction with a 252-day sliding window so as to obtain annual returns. From the data, we obtain $\mathbb{E}[X] \approx 0.14$ with probability of exceeding $\mathbb{E}[X]$ slightly greater than 0.5. That is $P(X > 0.14) \approx 0.53$.

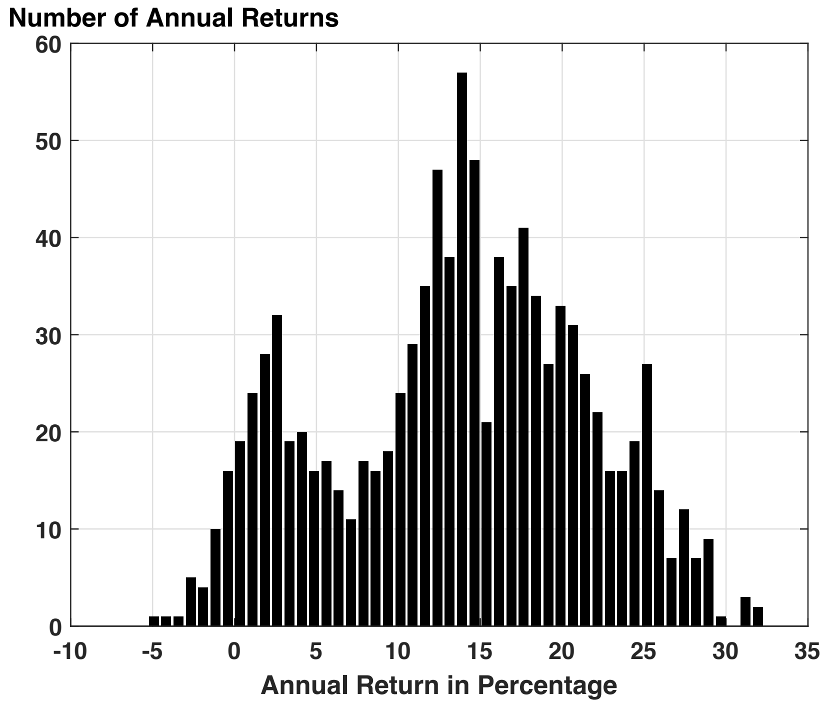

Figure 6.7.2: Histogram of Annual Returns for S&P 500, 2010-2015

One scenario under which a more conservative estimate of the mean is desirable occurs when conservative investors are reliant on index funds to either generate fixed-income in retirement or plan for the future. In this case, it is of interest to modify the expected value to obtain the CEV which is arguably a more realistic expectation for one who is risk-averse. Carrying out a γ-sweep, using the plot of $\mathbb{E}[X_\gamma]$ shown in Figure 6.7.3, we obtain $\mathbb{CEV}(X) \approx 0.06$. In addition we obtain $P(X > 0.06) \approx 0.8$. That is, the CEV analysis indicates that a conservative investor should expect only a 6% return. In contrast to the standard deviation $\sigma \approx 0.08$, for this data, the *Conservative Semi-Deviation* is found to be $D(X) \approx 0.02$. Hence, our estimate is "tighter" and more reliable for investors not interested in delving into the details related to one-sided risk.

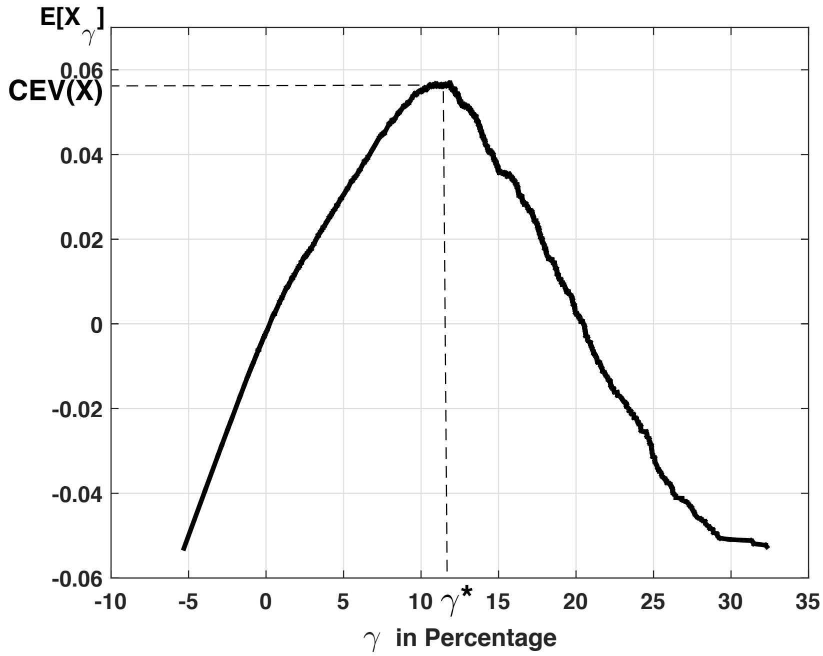

Figure 6.7.3: $\mathbb{E}[X_\gamma]$ Versus γ for the Annual Return Problem

Period	$\mathbb{E}[X]$	$\mathbb{CEV}(X)$	σ	$\sqrt{\mathbb{CSV}(X)}$
1-day	0.1427	-3.9230	0.1510	0.04
5-day	0.1393	-1.6123	0.1414	0.0346
1-month	0.1383	-0.5135	0.1208	0.0361
6-month	0.1401	-0.0184	0.1005	0.0361
1-year	0.1369	0.0571	0.0781	0.02

Table 6.1: Statistics of the Returns of S&P 500 for Different Time Horizons, 2010-2015

The same sliding window approach which was used above can be used to find the returns for different time horizons; i.e., daily, weekly, monthly, etc. For example, considering Table 6.1, where statistics are reported using the same data with different time horizons, it is interesting to note that the CEV is positive only for the case of annual returns. This can be interpreted as a warning to a conservative investor: In order to reasonably expect a positive return, it is better to own the stocks for a year or more.

6.7.3 Example: U.S. Household Annual Income 2011: The distribution of household annual income for 2011, according to US Census data given in [152], is shown in Figure 6.7.4. The statistical sample covers 121,084 households and is reported in each income bracket. The two isolated "spikes" at levels \$220,171 and \$426,271 correspond to wider income brackets used for high-income individuals. This distribution with its high skew, $S \approx 2.89$, makes this scenario a candidate for application of our theory. By way of additional motivation for use of the Conservative Expected Value, in this case, we note that "mean household income" is widely reported to the public at large.

The issue which we consider here is the extent to which the mean household income of \$69,677 is reflective of what is to be expected. In this regard, the histogram indicates that only about 35% of the households have incomes which exceed this mean level. This raises a question whether such reporting of this mean level to a technically unsophisticated population is "fair." As an alternative, we calculate the CEV by carrying out a γ-sweep to maximize $\mathbb{E}[X_\gamma]$. Indeed, from Figure 6.7.5, we obtain $\mathbb{CEV}(X) \approx 26,919$ and our analysis further indicates that there is a 75% chance of exceeding the CEV. At first glance, upon seeing this 75% probability, there is a temptation to conclude that CEV analysis produces an overly conservative result for this example. That is, if we view exceeding the expected value more like a 50-50 proposition, then one might argue that CEV theory produces a result which is unreasonably low; i.e., a technically unsophisticated "worker" is being provided with a prospect which is too grim.

However, when we enhance the analysis to include Conservative Semi-Variance considerations, we claim that for this application, it becomes more difficult to dismiss the CEV out of hand. That is, taking one-sidedness income risk into account, we enhance the CEV number with the *Conservative Semi-Deviation* to obtain $D(X) = \sqrt{\mathbb{CSV}(X)} \approx 7,539$. When comparing $D(X)$ with the classical standard deviation $\sigma = \sqrt{\text{var}(X)} \approx 73,103$, our much smaller number points to a trade-off between risk aversion and the estimate of expected income. Analogous to the editing phase in Prospect Theory [130], the appropriateness of application will depend on the extent to which one emphasizes downside risk.

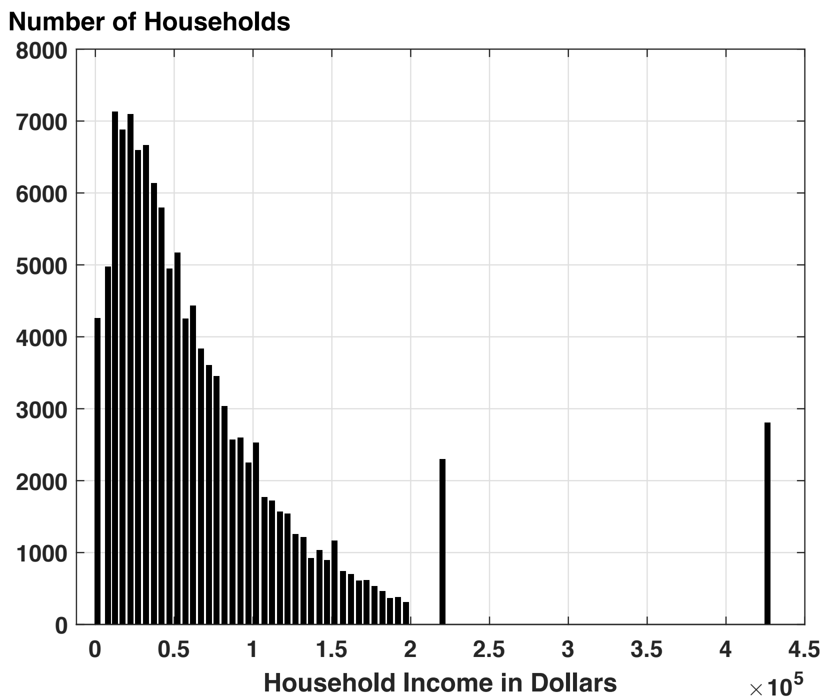

Figure 6.7.4: Histogram of US Household Income for 2011

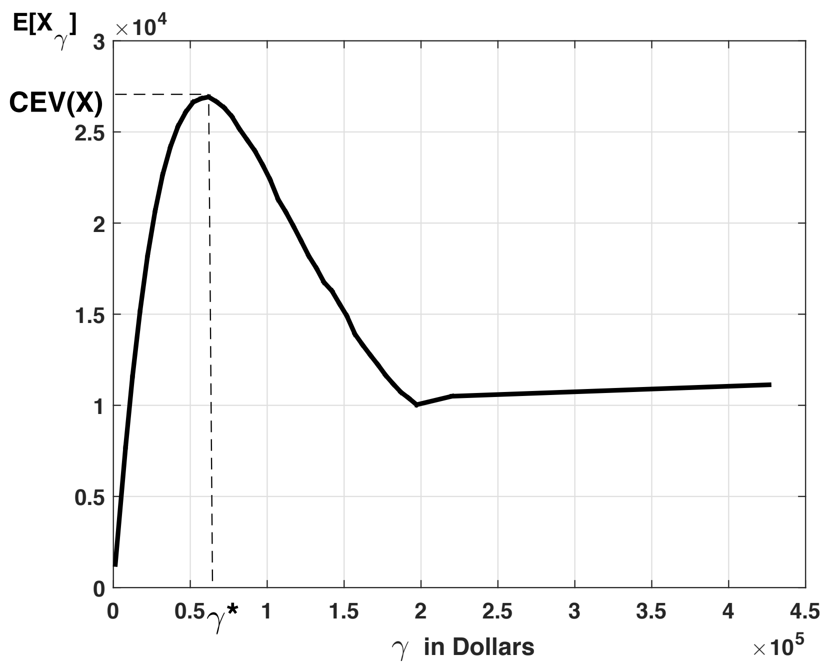

Figure 6.7.5: $\mathbb{E}[X_\gamma]$ Versus γ for the Household Income Problem

6.8 Conclusion and Further Research

In this chapter, we considered a random variable X for which high skew may be present and for which large positive values are desirable and argue that the classical expected value may be an unduly optimistic indicator of performance in that it can overstate what one might expect for the realization of X. It was also argued that this issue of over-optimism in the expected value is particularly important when a mission-critical random variable is involved. In addition, we considered the case when distrust in the underlying probability distribution may also be a concern as epitomized by our earthquake example in Section 6.2.3. We highlighted mission-critical applications with the need for a high degree of conservatism. To address such situations we defined the (CEV,CSV) reward-risk pair and developed a number of theoretical properties, interpretations and illustrative examples involving both theoretical distributions and data-generated histograms. This new theory appears to be mathematically rich and potentially useful in a number of applications. Given the analysis to date, many new research directions suggest themselves: One possible direction would be to modify the CEV and CSV definitions to accommodate the case of random variable X with unbounded leftmost support point; i.e., $\alpha_X = -\infty$.

A second new research direction begins with the fact that the CEV and CSV are defined without recourse to utility theory. Our view is that this new pair, analogous to mean and variance, may be useful in application where "conservatism" is desired and risk versus return must be carefully balanced. With regard to these considerations, one possible application might be appropriately called *Conservative Portfolio Selection*. In the spirit of [15] and [99], given n stocks with associated weights w_i being the fraction of wealth invested in each stock, a conservative version of the celebrated Markowitz problem [15], using semi-variance along the line of Harlow [91] could involve a CSV minimization with a constraint that the CEV exceed some prescribed target level.

Chapter 7

Directions for Future Work

At a high level, the focal point of this thesis has been a new line of research aimed at development of a theoretical framework to study the performance of a class of stock-trading strategies which are based on technical analysis. In direct contrast to many papers in the finance literature which are aimed at studying the profitability of such strategies via statistical analysis of backtests using historical price data, see [53] for an excellent survey and extensive bibliography, our approach here is based on theory. The strategies which we consider here involve linear feedback and can be viewed as a form of trend following; e.g., see [26, 81–83].

One interesting possible future path in this research area is the study of *discrete-time* feedback-based trading strategies. The controller with delay, described in Chapter 5, is an example of such trading strategy. We have dedicated a significant part of this chapter, Section 7.1, to introduce a new discrete-time feedback-based trading strategy which is *triggered by moving average crossing*. The moving average crossing is well known in finance; e.g., see [24, 54, 56, 57, 73, 153] and is often used time the market. In contrast, our new controller is motivated by the desire to enhance the performance of such moving average crossing algorithm via inclusion of feedback control. We formulate and study the performance via simulation using historical price data. The theoretical aspect of the research is briefly outlined and left for future work.

In this chapter, we also describe some other open research problems which are motivated by our work to date. In this regard, in Section 7.2, we discuss the problem of optimizing the parameters involved in the controllers under consideration. In Section 7.3, we briefly motivate and propose three new feedback controllers. Since the results provided in this thesis has been developed for the case of trading a single stock, in Section 7.4, we describe a research path to extend the analysis to the case of trading a portfolio of stocks. Finally, the focus of Section 7.5 is further work on Conservative Expected Value (CEV) and Conservative Semi-Variance (CSV) in finance.

7.1 Moving Average Crossing Problems

In this section, we describe a new direction of research involving a discrete-time linear feedback controller which is *triggered by moving average crossings* with the goal being to use feedback to improve upon the well-known results in finance; e.g., see [24, 54, 56, 57, 73, 153]. As a first step in this research direction, we carry out some backtesting on historical price data. The theoretical analysis is relegated to future research.

The description of our new controller begins with a classical algorithm summarized as follows: At each time instant k, the current price $p(k)$ is compared to an n-day moving average price

$$p_{av}(k) = \frac{1}{n} \sum_{i=k-n+1}^{k} p(i)$$

where n is selected by the trader based on considerations such as the volatility of the stock and its perceived drift. Now, assuming for simplicity that only long positions are allowed, if the price $p(k)$ crosses this moving average from below, a buy signal is generated. This suggests that an upward price trend is forming and dictates going long in the stock. Subsequently, if a crossing of the moving average occurs from above, the upward trend is deemed to be in doubt and a sell signal is generated. After selling the stock, the trader "lies in wait" holding cash until the next crossing from below occurs. This entry-exit process continues up to some terminal time $k = N$. Once the first crossing from below occurs, the investment for the trader who is either "fully invested" or "fully out" is given by

$$I(k) = \frac{1}{2} V(k) \left(\text{sign}\big(p(k) - p_{av}(k)\big) + 1 \right)$$

where $V(k)$ is the account value at stage k.

7.1.1 Illustration of Crossings: Before we proceed to our new feedback-based version of the strategy, we illustrate the ideas above using adjusted daily closing prices for Xerox (Ticker: XRX) stock. The price data is shown in Figure 7.1.1 and $n = 200$ is used for the moving average. The data covers the five-year time period, January 1, 2005 until December 31, 2009 which includes both the major bull market high in 2007 and the period of the "crash" in 2008 and 2009.

Calculating the moving average over the given time period, as seen in Figure 7.1.1, we obtain fourteen buy signals which dictate entering the trade, and thirteen sell signals which force exiting the trade. Using the investment rule given above leads to the account value shown in Figure 7.1.2. It is clear that the use of moving average timing keeps the trader out of the trade during the crash of $2008 - 2009$; i.e., during the period associated with $750 \leq k \leq 1180$.

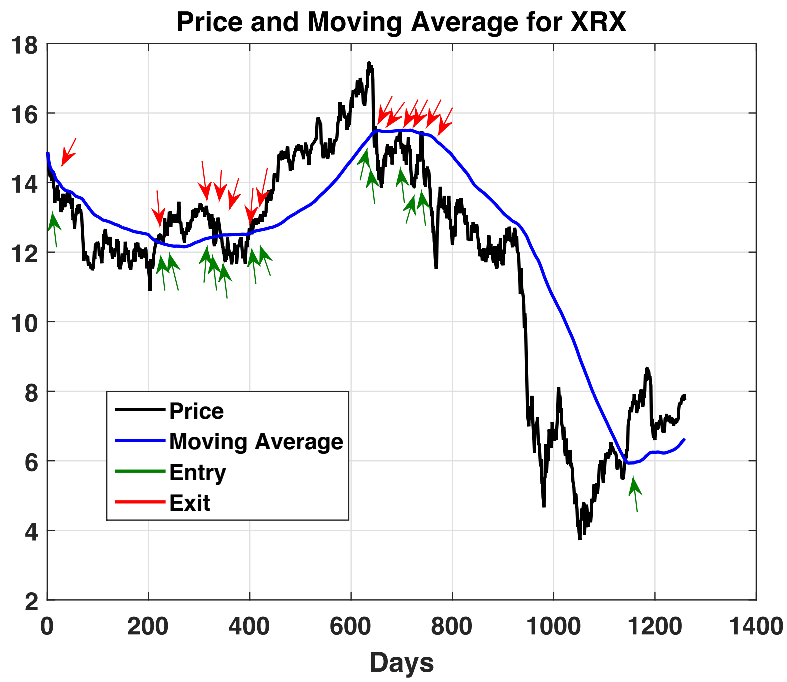

Figure 7.1.1: Price and 200-Day Moving Average for XRX

7.1.2 Inclusion of Linear Feedback: With the definition of classical moving average crossing in place, in this section, we now introduce a new trading rule which is triggered by moving average crossings but also includes a linear feedback controller aimed at enhancement of performance. For this new trading rule, the investment level at step k is given by

$$I(k) = \frac{1}{2}\left(I_0 + Kg(k)\right)\big(\text{sign}\big(p(k) - p_{av}(k)\big) + 1\big)$$

where I_0 is the initial investment, $K \geq 0$ is the feedback gain and $g(k)$ is trading gain-loss function. It is important to note that the factor corresponding to the moving average crossing determines the entries and exits; i.e., the market-timing. In contrast, the factor corresponding to linear feedback determines modification of the investment based on the performance $g(k)$.

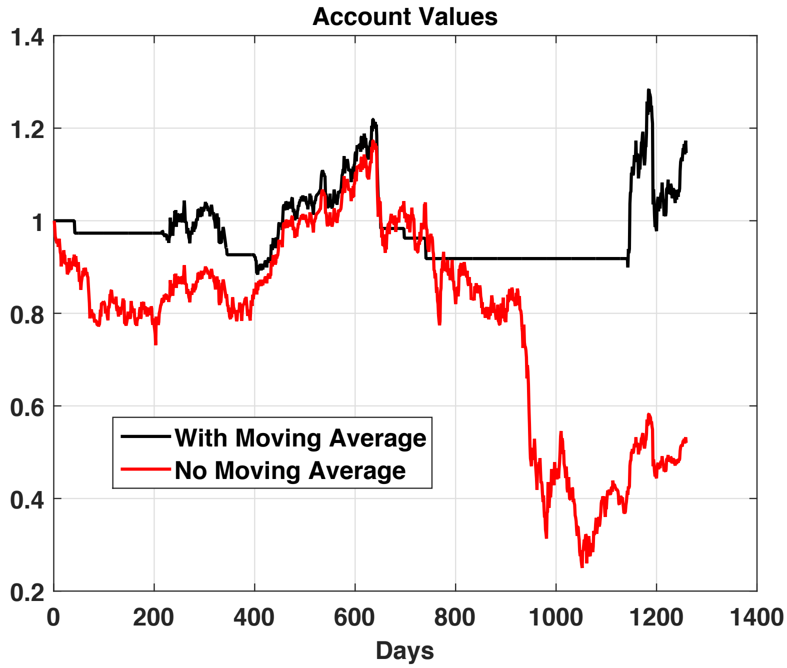

Figure 7.1.2: Account Values: With and Without Moving Average Timing

The dynamics of the account value for the fully-invested case is given by $V(k) = I_0 + g(k)$. This implies that the classical moving average crossing strategy described in the previous section is a special case of the controller introduced here with $K = 1$.

We note that this investment rule can result in the trade size exceeding the account value. To this end, a full-blown analysis of the ideas below could be expanded to include margin interest considerations for cases when $I(k) > V(k)$. In the following subsection, we refer to this Linear Feedback controller with Moving Average Timing as *LF-MAT* for short.

7.1.3 Backtesting the LF-MAT Controller: To backtest the LF-MAT controller, we consider an S&P portfolio which is now described: Among the 500 stocks in the current S&P 500 index, neglecting the possibility of "survivorship bias," we consider a portfolio consisting of 480 of them which were present in the index during the five-year time period January 1, 2005 until December 31, 2009. The list of these stocks can be found in [151]. In the simulation to follow, using adjusted closing prices, we apply LF-MAT controller to each of these stocks. We consider initial account value $V(0) = 1$ and an initial investment $I_0 = 1/480$ in dollars in each stock. Similar to our earlier approach in Subsection 5.1.1, the risk-free interest and margin interest rates are taken to be zero. Then, taking $g_i(k)$ to be the cumulative gain or loss at stage k for stock i, the total account value $V(k)$, is obtained by summing the contributions of each of the stocks. That is,

$$V(k) = V_0 + \sum_{i=1}^{480} g_i(k).$$

Taking a feedback gain of $K = 2$ in each LF-MAT controller, we compare our results with the "fully invested" moving average case which corresponds to $K = 1$ and the so-called "pure" linear feedback case with $K = 2$ and no moving average considerations.

In Figure 7.1.3, the resulting account values are shown. As seen in the figure, during the crash of 2008-2009, i.e., $850 \leq k \leq 1100$, the use of pure linear feedback leads to maximum percentage drawdown of about 63%. In comparison, the LF-MAT strategy results in a smaller drawdown of 32% and the moving average strategy leads to 22% drawdown. This larger drawdown for the pure linear feedback strategy is explained by the use of feedback gain $K = 2$ which implies taking more risk than $K = 1$ by investing a larger percentage of the account value.

For all these strategies, the plot of the leverage $L(k) \doteq I(k)/V(k)$, is shown in Figure 7.1.4. In the figure, we see that the pure linear feedback controller is the most aggressive in terms of leverage. Arguably, this is to be expected because this strategy is always in the market. In contrast, the leverage for both the moving average and LF-MAT strategies goes down to almost zero during the crash; i.e., in the period $900 \leq k \leq 1050$. During the bull markets covering $200 \leq k \leq 800$ and $1100 \leq k \leq 1250$, the use of linear feedback results in the use of more leverage as larger profits are captured. The inclusion of the moving average can "save" the trader during the crash by halting the trade while the inclusion of linear feedback is used to add to the degree of aggressiveness when bull markets are present.

This simulation suggests that there are many possibilities for further research on the use of moving averages in conjunction with feedback. These include expansion of the formulation above to include short selling and development of a theory, perhaps adaptive in nature, to select the window length n. With smaller window length, there would typically be more buy and sell signals generated and associated risk-return studies could readily be conducted to determine what trade-offs are worthwhile.

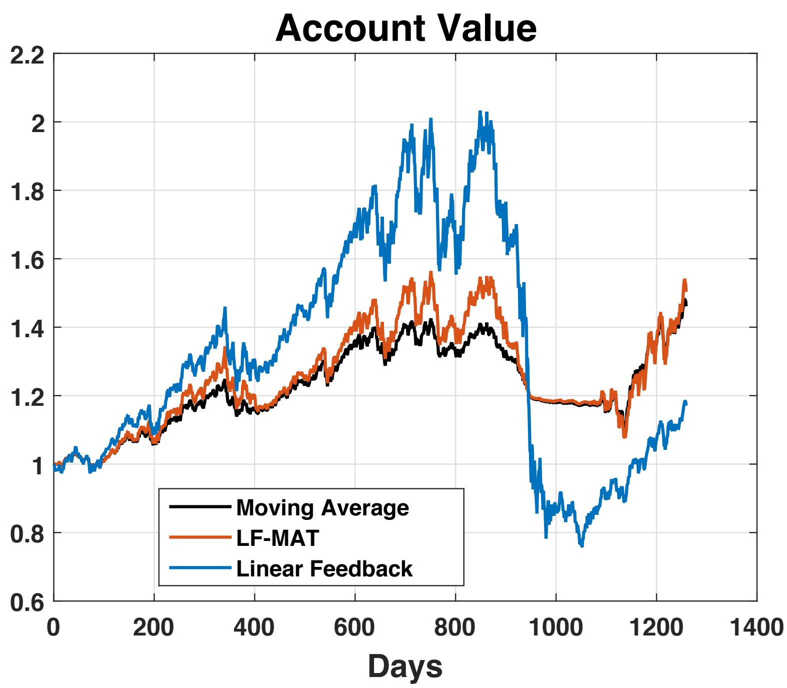

Figure 7.1.3: Account Values: Effect of Moving Average and Linear Feedback

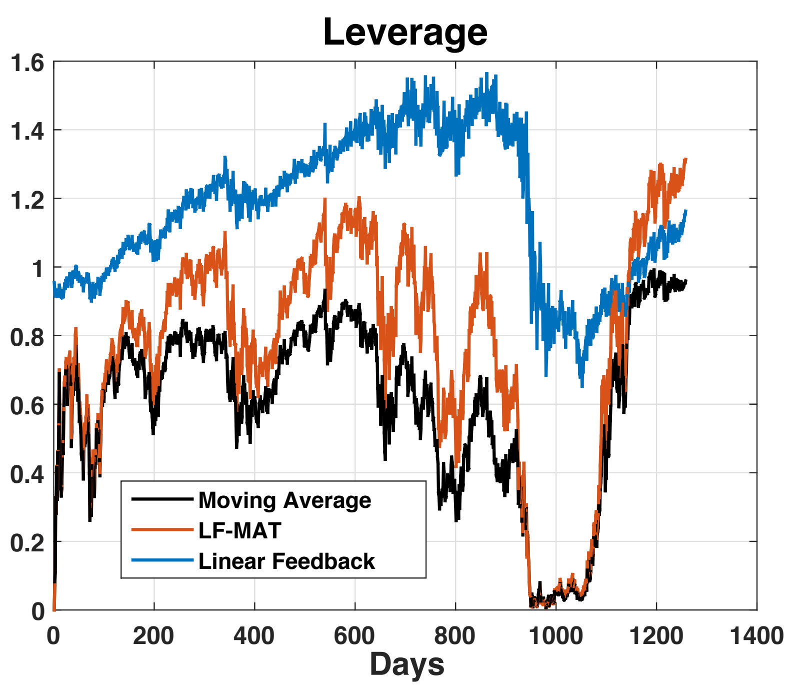

Figure 7.1.4: Leverage: Effect of Moving Average and Linear Feedback

7.2 Choice of Controller Parameters

In this thesis, we studied different trading rules motivated by various feedback control considerations and proved that the Robust Positive Expectation Property holds. Since this property does not depend on what positive feedback gains are used, many open research problems can be considered regarding the choice of these gains. For example, in Section 1.1, how should one pick the gain K? Similarly, in using PI controller with exponentially weighting scheme, see Chapter 4, there are open problems of interest regarding the choice of the parameters K_P, K_I and the discounting factor γ. Finally, for the controller with delay in Chapter 5, a possible optimization presents itself with respect to the values of feedback gain K and delay amount m.

With regard to the considerations above, there are many research directions to pursue. One obvious direction involves use of a "training data set" to optimize the parameters; i.e., we simulate the trade on historical price data using different values of parameters to obtain the best performing values.

To accomplish the above using training data, one can consider an optimization problem which uses no price model and is driven by the data. That is, the training data is directly used to find the best performing controller values which is in direct contrast with the *model-based* approach in which training data is initially used to find the "best-fit" price model and then the model is used to determine the controller values. For example, a classical mean-variance utility function can be maximized; e.g., see [15]. However, since the skewness effect studied in Chapter 2, suggests that reliance on such mean-variance based measures of performance can provide a distorted picture of the prospects for success, one possibility would be to use a different risk-return pair which includes information about higher order moments of $g(t)$. Alternatively, utility functions which capture higher-order moments, especially the cubic utility function, have been discussed in portfolio optimization literature such as [100, 102, 103] and can be extended to our framework with appropriate risk aversion coefficients. Finally, another possibility would be to include drawdown considerations in the optimization problems above. This can be important based on our discussion in Chapter 3 where we argued that the use of a large feedback gain K can lead to a highly-skewed probability distribution for gain-loss function. In turn, this can lead to a large drawdown in wealth.

7.3 Some New Possibilities for Feedback Control

In this section, we briefly describe some new possibilities for feedback controllers. We begin with a new linear negative-K rule and then discuss a class of *nonlinear* feedback controllers.

7.3.1 Use of Negative Feedback Gain K: Thus far, all of the results provided in this thesis rely on the assumption $K \geq 0$ in the trading rule. Consideration of negative K-values is motivated by the SLS analysis in Chapter 2: Despite the positive expected value for the gain-loss function, due to the large skewness, the probability of loss can be so large as to make the trade undesirable. Our finding is that SLS with feedback gain $K > 0$ is best suited for a market with a trend; i.e., a consistent upward or downward price movement. The "negative K" ideas to follow are aimed at sideways markets. With no specific direction for the price movement, the probability of loss using the SLS controller with positive feedback gain $K > 0$ is high.

Sideways Markets: There are various ways to describe sideways markets using classical price models. For example, one can consider a Geometric Brownian Motion with the drift $\mu = 0$ or a Vasicek process; e.g., see [154]. An important research direction involves development of new trading strategies which provide higher probability of winning in sideways markets while preserve SLS-type performance in a trending market. Our initial approach to development of such strategies leads us to consider the analysis of trading systems with negative feedback gain $K < 0$. Interestingly, the analysis leads to results which are quite different from the $K > 0$ case.

To begin, with positive initial investment $I_0 > 0$ and feedback gain $K < 0$, it is easy to show that the gain-loss function resulting from the use of SLS trading rule is an odd function of K. That is, denoting the dependence on K by writing $g(t, K)$; at time t, we have $g(t, -K) = -g(t, K)$. Hence, in the case of sideways markets where positive-gain SLS with $K > 0$ incurs a loss, negative-K version of SLS makes profit. Furthermore, we have

$$P\big(g(t,K) > \gamma\big) = P\big(g(t,-K) < -\gamma\big); \qquad \mathbb{E}[g(t,K)] = -\mathbb{E}[g(t,-K)].$$

Considering all of the above, a trader who believes the market is moving sideways may opt to give up the Robust Positive Expectation Property for positive-gain SLS to enjoy a large probability of winning using the negative-K version of the controller. That is, in such cases, the properties above suggest that use of negative feedback gain is more likely to lead to a profit.

7.3.2 Incorporating Volume Information: One of the important quantities which is used in technical analysis is the *trading volume*; e.g., see [23]. This corresponds to the number of shares which are traded during a pre-specified period of time. When the volume is low, traders typically interpret this to mean that market reaction is “tepid” and stock price changes are less “trustworthy.” Conversely, when volume is high, this is interpreted to mean traders view a price change as more significant. One possible path for future research involves use of this volume information in the controller.[1] For example, denoting the volume by v, suppose a smooth function of volume, $\phi(v)$, is used to modulate the SLS investment rule; i.e., with

$$I_L(t) = \phi(v)\big(I_0 + Kg_L(t)\big), \quad I_S(t) = \phi(v)\big(-I_0 - Kg_S(t)\big)$$

the dynamics at time t become

$$\frac{dg_L}{dt} = \phi(v)\big(I_0 + Kg_L(t)\big)\rho(t), \quad \frac{dg_S}{dt} = \phi(v)\big(-I_0 - Kg_S(t)\big)\rho(t).$$

and, for smooth prices, this can lead to “arbitrage” theorems similar to results given in [3].

7.3.3 Alternative Nonlinear Feedback Rules: Another possibility of future work is to consider classes of more complicated feedback controllers; e.g. controllers depending nonlinearly on measured variables. In this regard, for the case of feedback on $g(t)$, the focal point of this thesis, we now allow the investment to be a nonlinear function of g; i.e., $I = f(g)$. Associated with this controller, we define the *degree of aggressiveness* along trajectories to be

$$a(t) \doteq \frac{\partial f}{\partial g}\Big|_{g=g(t)}.$$

For a feedback-based strategy, $a(t)$ quantifies the degree of aggressiveness of an investment rule based on changes in the gain-loss function. At time $t \geq 0$, it is the amount of investment level change per dollar of profit. This concept is closely related to the evaluation of gain and losses as described in Prospect Theory, e.g., see [130]. For the special case of linear feedback with $f(g) = I_0 + Kg$, the degree of aggressiveness reduces to $a(t) = K$. That is, we increase or decrease the investment by K dollars for every dollar of profit or loss.

[1]The content of this subsection is based on the communication between Professor B. Ross Barmish, Professor James A. Primbs and me and is a part of an ongoing collaboration.

One possibility for future work is to consider more complicated nonlinear investment rules with degree of aggressiveness coefficients which are not constant with respect to $g(t)$. For example consider the investment rule

$$f(g) = \sqrt{I_0^2 + 2g}.$$

For this controller the initial investment is I_0 and the degree of aggressiveness coefficient is

$$a(t) = \frac{\partial f}{\partial g}\Big|_{g=g(t)} = \frac{1}{\sqrt{I_0^2 + 2g(t)}}.$$

The two investment rules, the linear feedback controller with feedback gain $K = 1$, and controller with square-root, are shown in Figure 7.3.1 as function of trading gain-loss $g(t)$. While both of these controllers have equal initial investment $I_0 = 1$, and initial degree of aggressiveness of $a(0) = 1$, the controller with square-root is more risk averse; that is, in case of a loss, $g < 0$, it dictates a lower investment; i.e., taking out money faster, and in case of a win; that is $g > 0$, it pumps in less money. Including this sort of considerations in developing different controllers would be of interest.

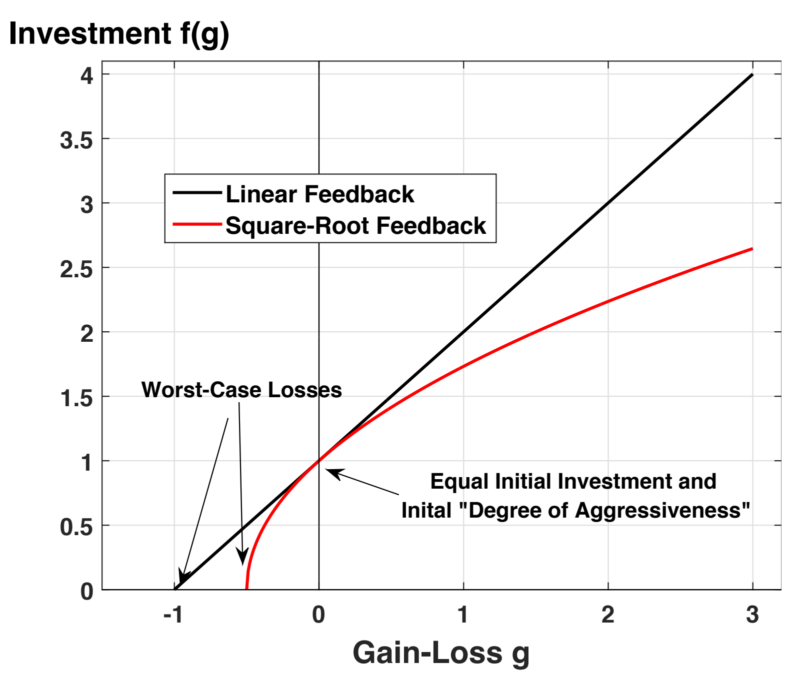

Figure 7.3.1: Linear Feedback Controller Versus Square-Root Controller

7.4 Extension of Analysis to Portfolio

An important future research direction involves extending the analysis in this thesis from trading a single stock to a portfolio of stocks. There are a variety of different ways to introduce performance-driven trading rules in rebalancing a portfolio. An interesting research direction involves combining the Markowitz portfolio optimization methods with our model-free feedback based-scheme as described below.

Indeed, we assume that each element w_i in vector W indicates the fraction of investment in stock i, for $i = 1, 2, \cdots, n$. Then, the portfolio optimization under Markowitz framework is given by

$$\begin{aligned} \underset{W}{\text{minimize}} \quad & W^T \Sigma W \\ \text{subject to} \quad & \mathbb{E}\big(R_P(W)\big) = R_P^*; \quad \sum_{i=1}^{n} w_i = 1 \end{aligned}$$

where Σ is the covariance matrix for the returns of the stocks and R_P^* is a user-defined, pre-specified target expected return; see [101] for details. To set up this optimization problem, historical data is typically used to estimate $\mathbb{E}\big(R_P(W)\big)$ and Σ. The goal is to obtain a portfolio which gives the targeted expected return with minimum variance.

Taking off from the Markowitz formulation, one idea for future research involves incorporating the nice trend-following properties of the SLS controller with the Markowitz analysis above. In this regard, one interesting research path would be to study the following two-step strategy: (i) Given the historical prices for the stocks, implement the Markowtiz portfolio optimization to obtain the optimum weights W^*. (ii) Once W^* is obtained, treat the resulting portfolio as a single stock and trade it via SLS controller. That is, the dynamics for the long component is given by

$$\begin{aligned} I_L(k) &= I_0 + K g_L(k); \\ g_L(k+1) &= g_L(k) + \rho(k) I_L(k) \end{aligned}$$

where

$$\rho(k) \doteq \sum_{i=1}^{n} w_i^* \rho_i(k).$$

and $\rho_i(k)$ is the return of stock i at stage k. The short component of SLS is obtained by turning I_0 and K to $-I_0$ and $-K$ respectively in the formulation above.

It is important to note that this formulation differs from the *model-free* approach pursued throughout this thesis. The proposed strategy above involves a mean-variance model in the first phase. By this strategy we "feedback-enhance" the Markowitz method which is essentially aimed at robustification of performance.

7.5 Further Research on CEV and CSV

The focus of Chapter 6 was the development of a new reward-risk pair: The *Conservative Expected Value (CEV)* and the *Conservative Semi-Variance (CSV)*. The development of this pair was an "offshoot" of our research with the motivation coming from evaluation of highly-skewed probability distribution for trading gain-loss function $g(t)$ based on feedback control. The properties we developed for the CEV and CSV suggest that this new theory is mathematically rich. One possible direction for future work involves modifying the definitions of CEV and CSV in order to accommodate the case when the random variable X has unbounded leftmost support point, $\alpha_X = -\infty$; i.e., the worst case is unbounded. Another possibility for extension of these definitions involves consideration of vector random variables.

New research directions also include the following: First, it would be of interest to extend the Markov-Chebyshev inequality improvements of Section 6.6.11 to other concentration-type inequalities such as the Chernoff bound. A second possibility involves formulation and solution of a new problem which might appropriately be called *Conservative Portfolio Selection*. In the spirit of [15] and [99], if we consider n stocks with associated weights w_i being the fraction of wealth invested in each stock, then a conservative version of the celebrated Markowitz problem [15], using semi-variance along the line of Harlow [91] could involve some sort of CSV minimization with constraints that the CEV exceed some prescribed level.

LIST OF REFERENCES

[1] B. R. Barmish and J. A. Primbs, "On a New Paradigm for Stock Trading Via a Model-Free Feedback Controller," *IEEE Transactions on Automatic Control*, in press, 2015.

[2] B. R. Barmish, "On Trading of Equities: A Robust Control Paradigm," in *IFAC World Congress*, pp. 1621–1626, 2008.

[3] B. R. Barmish, "On Performance Limits of Feedback Control-Based Stock Trading Strategies," in *American Control Conference*, pp. 3874–3879, 2011.

[4] B. R. Barmish and J. A. Primbs, "On Arbitrage Possibilities Via Linear Feedback in an Idealized Brownian Motion Stock Market," in *IEEE Conference on Decision and Control*, pp. 2889–2894, 2011.

[5] B. R. Barmish and J. A. Primbs, "On Market-Neutral Stock Trading Arbitrage via Linear Feedback," in *American Control Conference*, pp. 3693–3698, 2012.

[6] B. R. Barmish, J. A. Primbs, S. Malekpour, and S. Warnick, "On the Basics for Simulation of Feedback-Based Stock Trading Strategies : An Invited Tutorial Session," in *IEEE Conference on Decision and Control*, pp. 7181–7186, 2013.

[7] S. Malekpour and B. R. Barmish, "How Useful are Mean-Variance Considerations in Stock Trading via Feedback Control?," in *IEEE Conference on Decision and Control*, pp. 2110–2115, 2012.

[8] S. Malekpour and B. R. Barmish, "A Drawdown Formula for Stock Trading Via Linear Feedback in a Market Governed by Brownian Motion," in *European Control Conference*, pp. 87–92, 2013.

[9] S. Malekpour, J. A. Primbs, and B. R. Barmish, "On Stock Trading Using a PI Controller in an Idealized Market : The Robust Positive Expectation Property," in *IEEE Conference on Decision and Control*, pp. 1210–1216, 2013.

[10] S. Malekpour and B. R. Barmish, "On the Use of Model-Free Linear Feedback to Trade Stock : Two Open Research Problems," in *52nd Annual Allerton Conference*, pp. 343–349, 2014.

[11] M. H. Baumann, "Effects of Linear Feedback Trading in an Interactive Market Model." Working Paper, 2014.

[12] J. A. Primbs and B. R. Barmish, "ACC 2012 Tutorial Session: An Introduction to Hedged-Like Stock Trading from a Control Theoretic Point of View," in *American Control Conference*, pp. 4496–4497, 2012.

[13] G. C. Calafiore and B. Monastero, "Data-Driven Asset Allocation with Guaranteed Shortfall Probability," in *American Control Conference*, pp. 3687–3692, 2012.

[14] I. Khademi and K. Zhou, "How to Invest in Uncertain Markets," in *Proceedings of the 34th Chinese Control Conference*, pp. 8617–8622, 2015.

[15] H. M. Markowitz, "Portfolio Selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[16] S. Achelis, *Technical Analysis from A to Z*. McGraw-Hill, 2 ed., 2000.

[17] C. D. Kirkpatrick and J. R. Dahlquist, *Technical Analysis: The Complete Resource for Financial Market Technicians*. FT Press, 2007.

[18] A. W. Lo and J. Hasanhodzic, *The Evolution of Technical Analysis: Financial Prediction from Babylonian Tablets to Bloomberg Terminals*. Bloomberg Press, 2010.

[19] S. Neftci, "Naive Trading Rules in Financial Markets and Wiener-Kolmogorov Prediction Theory: A Study of "Technical Analysis"," *The Journal of Business*, vol. 64, no. 4, pp. 549–571, 1991.

[20] J. L. Treynor and R. Ferguson, "In Defense of Technical Analysis," *The Journal of Finance*, vol. 40, no. 3, pp. 757–773, 1985.

[21] A. Lo, H. Mamaysky, and J. Wang, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *The Journal of Finance*, vol. 55, no. 4, pp. 1705–1765, 2000.

[22] D. Brown and R. Jennings, "On Technical Analysis," *The Review of Financial Studies*, vol. 2, no. 4, pp. 527–551, 1989.

[23] L. Blume, D. Easley, and M. O'hara, "Market Statistics and Technical Analysis: The Role of Volume," *The Journal of Finance*, vol. 49, no. 1, pp. 153–181, 1994.

[24] Y. Zhu and G. Zhou, "Technical Analysis: An Asset Allocation Perspective on The Use of Moving Averages," *The Journal of Financial Economics*, vol. 92, no. 3, pp. 519–544, 2009.

[25] H. Zhang and Q. Zhang, "Trading a Mean-Reverting Asset: Buy Low and Sell High," *Automatica*, vol. 44, no. 6, pp. 1511–1518, 2008.

[26] M. Dai, Q. Zhang, and Q. Zhu, "Trend-Following Trading Under a Regime Switching Model," *SIAM Journal on Financial Mathematics*, vol. 1, no. 65, pp. 780–810, 2010.

[27] G. C. Calafiore, "Multi-Period Portfolio Optimization with Linear Control Policies," *Automatica*, vol. 44, no. 10, pp. 2463–2473, 2008.

[28] S. Iwarere and B. R. Barmish, "A Confidence Interval Triggering Method for Stock Trading via Feedback Control," in *American Control Conference*, pp. 6910–6916, 2010.

[29] S. Mudchanatongsuk, J. A. Primbs, and W. Wong, "Optimal Pairs Trading: A Stochastic Control Approach," in *American Control Conference*, pp. 1035–1039, 2008.

[30] J. A. Primbs and C. H. Sung, "A Stochastic Receding Horizon Control Approach to Constrained Index Tracking," *Asia-Pacific Financial Markets*, vol. 15, no. 1, pp. 3–24, 2008.

[31] J. A. Primbs, "LQR and Receding Horizon Approaches to Multi-Dimensional Option Hedging under Transaction Costs," in *American Control Conference*, pp. 6891–6896, 2010.

[32] A. Lim and X. Zhou, "Mean-Variance Portfolio Selection with Random Parameters in a Complete Market," *The Journal of Mathematics of Operations Research*, vol. 27, no. 1, pp. 101–120, 2002.

[33] A. Lim, "Mean-Variance Hedging When There are Jumps," *SIAM Journal on Control and Optimization*, vol. 44, no. 5, pp. 1893–1922, 2005.

[34] R. Merton, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *The Journal of Economic Theory*, vol. 3, pp. 373–413, 1970.

[35] P. Samuelson, "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, vol. 41, no. 9, pp. 41–49, 1965.

[36] A. Bemporad, T. Gabbriellini, L. Puglia, and L. Bellucci, "Scenario-Based Stochastic Model Predictive Control for Dynamic Option Hedging," in *IEEE Conference on Decision and Control*, pp. 6089–6094, 2010.

[37] G. C. Calafiore, "An Affine Control Method for Optimal Dynamic Asset Allocation with Transaction Costs," *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2254–2274, 2009.

[38] G. Calafiore, "Ambiguous Risk Measures and Optimal Robust Portfolios," *SIAM Journal on Optimization*, vol. 18, no. 3, pp. 853–877, 2007.

[39] G. C. Calafiore and B. Monastero, "Triggering Long-Short Trades on Indexes," *International Journal of Trade, Economics and Finance*, vol. 1, no. 3, pp. 289–296, 2010.

[40] V. V. Dombrovskii, D. V. Dombrovskii, and E. A. Lyashenko, "Predictive Control of Random-Parameter Systems with Multiplicative Noise. Application to Investment Portfolio Optimization," *Automation and Remote Control*, vol. 66, no. 4, pp. 583–595, 2005.

[41] V. V. Dombrovskii, D. V. Dombrovskii, and E. A. Lyashenko, "Model Predictive Control of Systems with Random Dependent Parameters under Constraints and its Application to the Investment Portfolio Optimization," *Automation and Remote Control*, vol. 67, no. 12, pp. 1927–1939, 2006.

[42] F. Herzog, "Strategic Portfolio Management for Long-Term Investments : An Optimal Control Approach," *PhD dissertation, ETH Zurich, Zurich, Switzerland*, 2005.

[43] F. Herzog, S. Keel, G. Dondi, L. M. Schumann, and H. P. Geering, "Model Predictive Control for Portfolio Selection," in *American Control Conference*, pp. 1252–1259, 2006.

[44] P. J. Meindl and J. A. Primbs, “Dynamic Hedging of Single and Multi-Dimensional Options with Transaction Costs: A Generalized Utility Maximization Approach,” *Quantitative Finance*, vol. 8, no. 3, pp. 299–312, 2008.

[45] M. B. Rudoy and C. E. Rohrs, “A Dynamic Programming Approach to Two-Stage Mean-Variance Portfolio Selection in Cointegrated Vector Autoregressive Systems,” in *IEEE Conference on Decision and Control*, no. 1, pp. 4280–4285, 2008.

[46] Q. Zhang, “Stock Trading: An Optimal Selling Rule,” *SIAM Journal on Control and Optimization*, vol. 40, no. 1, pp. 64–87, 2001.

[47] L. E. Ghaoui, M. Oks, and F. Oustry, “Worst-Case Value-At-Risk and Robust Portfolio Optimization: A Conic Programming Approach,” *Operations Research*, vol. 51, no. 4, pp. 543–556, 2003.

[48] P. Meindl and J. A. Primbs, “Dynamic Hedging with Stochastic Volatility Using Receding Horizon Control,” in *Financial Engineering Applications*, pp. 142–147, 2004.

[49] P. Meindl, “Portfolio Optimization and Dynamic Hedging with Receding Horizon Control, Stochastic Programming and Monte Carlo Simulation,” *PhD dissertation, Stanford University*, 2006.

[50] T. Duncan, B. Pasik Duncan, and L. Stettner, “Growth Optimal Portfolio Selection Under Proportional Transaction Costs with Obligatory Diversification,” *Applied Mathematics and Optimization*, vol. 63, no. 1, pp. 107–132, 2010.

[51] T. Duncan, B. Pasik-Duncan, and L. Stettner, “Parameter Continuity of The Ergodic Cost for a Growth Optimal Portfolio With Proportional Transaction Costs,” in *IEEE Conference on Decision and Control*, pp. 4275–4279, 2008.

[52] J. Primbs, “Dynamic Hedging of Basket Options Under Proportional Transaction Costs Using Receding Horizon Control,” *International Journal of Control*, vol. 82, no. 10, pp. 1841–1855, 2009.

[53] C. H. Park and S. H. Irwin, “What Do We Know About the Profitability of Technical Analysis?,” *Journal of Economic Surveys*, vol. 21, pp. 786–826, Sept. 2007.

[54] W. Brock, J. Lakonishok, and B. LeBaron, “Simple Technical Trading Rules and the Stochastic Properties of Stock Returns,” *The Journal of Finance*, vol. 47, no. 5, pp. 1731–1764, 1992.

[55] Y. Han, K. Yang, and G. Zhou, “A New Anomaly : The Cross-Sectional Profitability of Technical Analysis,” *Available at SSRN 1656460*, 2011.

[56] T. Kilgallen, “Testing the Simple Moving Average across Commodities, Global Stock Indices, and Currencies,” *Jour. Wealth Management*, vol. 15, no. 1, pp. 82–100, 2012.

[57] P. Glabadanidis, “Market Timing with Moving Averages,” *Available at SSRN 2127483*, no. 8, 2012.

[58] E. Fama and M. Blume, "Filter Rules and Stock Market Trading," *The Journal of Business*, vol. 39, no. 1, pp. 226–241, 1966.

[59] A. Cowles, "Can Stock Market Forecasters Forecast ?," *Econometrica: The Journal of the Econometric Society*, vol. 1, no. 3, pp. 309–324, 1933.

[60] J. C. Van Horne and G. G. C. Parker, "The Random-Walk Theory: An Empirical Test," *Financial Analysts Journal*, vol. 23, no. 6, pp. 87–92, 1967.

[61] J. C. Van Horne and G. G. C. Parker, "Technical Trading Rules: A Comment," *Financial Analysts Journal*, vol. 24, no. 4, pp. 128–132, 1968.

[62] M. C. Jensen and G. A. Benington, "Random Walks and Technical Theories: Some Additional Evidence," *The Journal of Finance*, vol. 25, no. 2, pp. 469–482, 1970.

[63] B. G. Malkiel, *A Random Walk Down Wall Street*. W.W. Norton & Company, INC, 2011.

[64] E. Fama, "The Behavior of Stock-Market Prices," *The Journal of Business*, vol. 38, no. 1, pp. 34–105, 1965.

[65] E. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[66] E. Fama, "Random Walks in Stock Market Prices," *Financial Analysts Journal*, vol. 51, no. 1, pp. 75–80, 1995.

[67] A. Lo and A. C. MacKinlay, "Stock Market Prices Do Not Follow Random Walks : Evidence From a Simple Specification Test," *Rev. Financial Studies*, vol. 1, no. 1, pp. 41–66, 1988.

[68] A. Lo and A. C. MacKinlay, *A Non-Random Walk Down Wall Street*. Princeton University Press, 2011.

[69] E. W. Tabell and A. W. Tabell, "The Case For Technical Analysis," *Financial Analysts Journal*, vol. 20, no. 2, pp. 67–76, 1964.

[70] S. W. Pruitt and R. E. White, "The CRISMA Trading System: Who Says Technical Analysis Can't Beat The Market?," *Jour. Portfolio Management*, vol. 14, no. 3, pp. 55–58, 1988.

[71] H. Bessembinder and K. Chan, "The Profitability of Technical Trading Rules in The Asian Stock Markets," *Pacific-Basin Finance Journal*, vol. 3, pp. 257–284, 1995.

[72] R. Hudson, M. Dempsey, and K. Keasey, "A Note on The Weak Form Efficiency of Capital Markets: The Application of Simple Technical Trading Rules to UK Stock Prices-1935 to 1994," *The Journal of Banking & Finance*, vol. 20, pp. 1121–1132, 1996.

[73] P. J. Detry and P. Grégoire, "Other Evidences of The Predictive Power of Technical Analysis: The Moving Averages Rules on European Indexes," in *Lugano Meetings*, pp. 1–25, 2001.

[74] B. R. Marshall and R. H. Cahan, "Is Technical Analysis Profitable on a Stock Market Which Has Characteristics That Suggest It May Be Inefficient?," *Research in International Business and Finance*, vol. 19, no. 3, pp. 384–398, 2005.

[75] M. P. Taylor and H. Allen, "The Use of Technical Analysis in the Foreign Exchange Market," *Journal of International Money and Finance*, vol. 11, no. 3, pp. 304–314, 1992.

[76] L. Menkhoff, "The Use of Technical Analysis by Fund Managers: International Evidence," *Journal of Banking and Finance*, vol. 34, no. 11, pp. 2573–2586, 2010.

[77] A. Lo and A. MacKinlay, "Data-Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial studies*, vol. 3, no. 3, pp. 431–467, 1990.

[78] R. Sullivan, A. Timmermann, and H. White, "Data-Snooping, Technical Trading Rule Performance, and the Bootstrap," *The Journal of Finance*, vol. 54, no. 5, pp. 1647–1691, 1999.

[79] M. Ready, "Profits From Technical Trading Rules," *Financial Management*, vol. 31, no. 3, pp. 43–61, 2002.

[80] M. Qi and Y. Wu, "Technical Trading-Rule Profitability, Data Snooping, and Reality Check: Evidence From the Foreign Exchange Market," *The Journal of Money, Credit and Banking*, vol. 38, no. 8, pp. 2135–2158, 2006.

[81] M. Covel, *Trend Following: Learn to Make Millions in Up or Down Markets*. FT Press, 2005.

[82] D. Darius, A. Ilhan, and J. Mulvey, "Trend-Following Hedge Funds and Multi-Period Asset Allocation," *Quantitative Finance*, vol. 2, no. 5, pp. 354–361, 2002.

[83] M. Davidsson, "Trend Following and Expected Return," *Economics Management and Financial Markets*, vol. 6, no. 2, pp. 39–56, 2011.

[84] R. C. Merton, *Continuous-Time Finance.* Wiley-Blackwell, 1992.

[85] M. Osborne, "Brownian Motion in The Stock Market," *Operations Research*, vol. 7, no. 2, pp. 145–173, 1959.

[86] G. Shafer and V. Vovk, *Probability and Finance, It's Only a Game*. Wiley & Sons, 2001.

[87] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge University Press, 2006.

[88] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2010.

[89] F. Black and M. Scholes, "The Pricing of Options and Corporate Liabilities," *The Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[90] F. A. Sortino and L. N. Price, "Performance Measurement in a Downside Risk Framework," *The Journal of Investing*, vol. 3, no. 3, pp. 59–64, 1994.

[91] W. Harlow, "Asset Allocation in a Downside-Risk Framework," *Financial Analysts Journal*, vol. 47, no. 5, pp. 28–40, 1991.

[92] M. Magdon-Ismail, A. F. Atiya, A. Pratap, and Y. S. Abu-Mostafa, "On the Maximum Drawdown of a Brownian Motion," *Jour. of Applied Probability*, vol. 41, no. 1, pp. 147–161, 2004.

[93] S. Grossman and Z. Zhou, "Optimal Investment Strategies for Controlling Drawdown," *Mathematical Finance*, vol. 3, no. 241-276, 1993.

[94] B. T. Hayes, "Maximum Drawdowns of Hedge Funds with Serial Correlation," *The Journal of Alternative Investments*, vol. 8, no. 4, pp. 26–38, 2006.

[95] M. Magdon-Ismail, "An Analysis of the Maximum Drawdown Risk Measure," *Risk Magazine*, 2004.

[96] R. Groeneveld and G. Meeden, "Measuring Skewness and Kurtosis," *The Statistician*, vol. 33, no. 4, pp. 391–399, 1984.

[97] C. Huang and R. Litzenberger, *Foundations of Financial Economics*, vol. 4. New York: North-Holland, 1988.

[98] W. Sharpe, "Mutual Fund Performance," *The Journal of Business*, vol. 39, no. 1, pp. 119–138, 1966.

[99] P. Samuelson, "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moment," *The Review of Economic Studies*, vol. 37, no. 4, pp. 537–542, 1970.

[100] C. R. Harvey, J. C. Liechty, M. W. Liechty, and P. Müller, "Portfolio Selection with Higher Moments," *Quantitative Finance*, vol. 10, no. 5, pp. 469–485, 2010.

[101] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. Wiley-Blackwell, 1991.

[102] G. Hanoch and H. Levy, "Efficient Portfolio Selection with Quadratic and Cubic Utility," *The Journal of Business*, vol. 43, no. 2, pp. 181–189, 1970.

[103] J. Ingersoll, "Multidimensional Security Pricing," *The Journal of Financial and Quantitative Analysis*, vol. 10, no. 5, pp. 785–798, 1975.

[104] P. Jorion, *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill Education, 2006.

[105] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Coherent Measures of Risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[106] K. Horcher, *Essentials of Financial Risk Management*. John Wiley and Sons, 2005.

[107] J. Batten and T. Fetherston, "Financial Risk and Finanacial Risk Management," *Research in International Business and Finance Series*, vol. 16, 2002.

[108] Z. Bodie, A. Kane, and A. Marcus, *Investments*. McGraw-Hill, 2010.

[109] J. Kelly JR., "A New Interpretation of Information Rate," *Information Theory, IRE Transactions on*, vol. 2, no. 3, pp. 185–189, 1956.

[110] "Q-function, http://www.cs.rpi.edu// magdon/data/Qfunctions.html."

[111] S. G. Krantz, *Handbook of Complex Variables*. Birkhäuser, 1999.

[112] J. Willems and J. Willems, "Feedback Stabilizability for Stochastic Systems with State and Control Dependent Noise," *Automatica*, vol. 12, no. 1, pp. 277–283, 1976.

[113] L. El Ghaoui, "State-feedback Control of Systems with Multiplicative Noise via Linear Matrix Inequalities," *Systems & Control Letters*, vol. 24, pp. 223–228, 1995.

[114] C. Bohn and D. P. Atherton, "An Analysis Package Comparing PID Anti-Windup Strategies," *IEEE Control Systems Magazine*, vol. 15, pp. 34–40, 1995.

[115] S. Iwarere and B. R. Barmish, "On Stock Trading Over a Lattice via Linear Feedback," in *The International Federation of Automatic Control*, pp. 7799–7804, 2014.

[116] S. Malekpour and B. R. Barmish, "The Conservative Expected Value: A New Measure with Motivation from Stock Trading via Feedback," in *The International Federation of Automatic Control*, no. 1, pp. 8719–8724, 2014.

[117] S. Malekpour and B. R. Barmish, "When the Expected Value is not Expected: A Conservative Approach," *Under Journal Review*, 2015.

[118] V. Le Sourd, "Performance Measurement for Traditional Investment: Literature Survey," *Financial Analysts Journal*, vol. 58, no. 4, pp. 36–52, 2007.

[119] N. T. Taleb, *The Black Swan: The Impact of the Highly Improbable*. Penguin, 2008.

[120] P. Embrechts, S. I. Resnick, and G. Samorodnitsky, "Extreme Value Theory as a Risk Management Tool," *North American Actuarial Journal*, vol. 3, no. 2, pp. 30–41, 1999.

[121] S. N. Neftci, "Value at Risk Calculations, Extreme Events, and Tail Estimation," *The Journal of Derivatives*, vol. 7, no. 3, pp. 23–37, 2000.

[122] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer Science & Media, 2009.

[123] J. M. Mendel, "Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, 1991.

[124] C. M. Harris and D. M. Wolpert, "Signal-Dependent Noise Determines Motor Planning.," *Nature*, vol. 394, no. 6695, pp. 780–784, 1998.

[125] H. Urkowitz, "Energy Detection of Unknown Deterministic Signals," *Proceedings of the IEEE*, vol. 55, no. 4, 1967.

[126] V. Jacobson, "Congestion Avoidance and Control," *ACM SIGCOMM Computer Communication Review*, vol. 18, no. 4, pp. 314–329, 1988.

[127] S. Hwang and S. E. Satchell, "Modelling Emerging Market Risk Premia Using Higher Moments," *International Journal of Finance and Economics*, vol. 4, no. 4, pp. 271–296, 1999.

[128] L. P. Hansen and T. J. Sargent, *Robustness*. Princeton University Press, 2008.

[129] D. Bernoulli, "Exposition of a New Theory on the Measurement of Risk," *Econometrica: Journal of the Econometric Society*, vol. 22, no. 1, pp. 23–36, 1954.

[130] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica: The Journal of the Econometric Society*, vol. 47, no. 2, pp. 263–291, 1979.

[131] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[132] A. Martin-Löf, "A Limit Theorem Which Clarifies the 'Petersburg Paradox'," *Journal of Applied Probability*, vol. 22, no. 3, pp. 634–643, 1985.

[133] M. Allais, "Le Comportement de l'homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'école Américaine," *Econometrica: The Journal of the Econometric Society*, vol. 21, no. 4, pp. 503–546, 1953.

[134] D. Ellsberg, "Risk, Ambiguity, and The Savage Axioms," *The Quarterly Journal of Economics*, vol. 75, no. 4, pp. 643–669, 1961.

[135] M. Rausand and A. Hoyland, *System Reilability Theory Models and Statistical Methods*. John Wiley & Sons, 2004.

[136] W. T. Ziemba, "The Symmetric Downside-Risk Sharpe Ratio," *The Journal of Portfolio Management*, vol. 32, no. 1, pp. 108–122, 2005.

[137] H. S. Marmer and F. L. Ng, "Mean-Semivariance Analysis of Option-Based Strategies: A Total Asset Mix Perspective," *Financial Analysts Journal*, vol. 49, pp. 47–54, May 1993.

[138] P. C. Fishburn, "Mean-Risk Analysis with Risk Associated with Below-Target Returns," *The American Economic Review*, vol. 67, no. 2, pp. 116–126, 1977.

[139] R. T. Rockafellar, S. P. Uryasev, and M. Zabarankin, "Deviation Measures in Risk Analysis and Optimization," tech. rep., 2002.

[140] T. Yegulalp and J. Kuo, "Statistical Prediction of The Occurrence of Maximum Magnitude Earthquakes," *International Journal of Rock Mechanics and Mining Sciences* & *Geomechanics*, vol. 11, no. 10, pp. 393–414, 1974.

[141] E. H. Field and WGCEP Members of the 2014, "UCERF3: A New Earthquake Forecast for Californias Complex Fault System," *Bulletin of the Seismological Society of America*, vol. 104, pp. 1122–1180, 2014.

[142] A. Kijko, "Estimation of The Maximum Earthquake Magnitude, m_{max}," *Pure and Applied Geophysics*, vol. 161, no. 8, pp. 1655–1681, 2004.

[143] T. C. Hanks and H. Kanamori, "A Moment Magnitude Scale," *Journal of Geophysical Research*, vol. 84, no. B5, pp. 2348–2350, 1979.

[144] T. Lay, Y. Fujii, E. Geist, K. Koketsu, J. Rubinstein, T. Sagiya, and M. Simons, "Introduction to the Special Issue on the 2011 Tohoku Earthquake and Tsunami," *Bulletin of the Seismological Society of America*, vol. 103, no. 2B, pp. 1165–1170, 2013.

[145] F. Tajima, J. Mori, and B. L. N. Kennett, "A Review of The 2011 Tohoku-Oki Earthquake (Mw 9.0): Large-Scale Rupture Across Heterogeneous Plate Coupling," *Tectonophysics*, vol. 586, no. 2013, pp. 15–34, 2013.

[146] A. Stork, J. Verdon, and J. M. Kendall, "The Robustness of Seismic Moment and Magnitudes Estimated Using Spectral Analysis," *Geophysical Prospecting*, vol. 62, no. 4, pp. 862–878, 2014.

[147] "National Centers For Environmental Information, Data Available to Download at: http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=35&d=35."

[148] J. W. Osborne, "Notes on the Use of Data Transformations," *Practical Assessment, Research & Evaluation*, vol. 8, no. 6, pp. 1–6, 2002.

[149] J. W. Osborne and A. Overbay, "The Power of Outliers (and Why Researchers Should Always Check for Them)," *Practical Assessment, Research* & *Evaluation*, vol. 9, no. 6, pp. 1–12, 2004.

[150] T. Micceri, "The Unicorn, the Normal Curve, and Other Improbable Creatures," *Psychological Bulletin*, vol. 105, no. 1, pp. 156–166, 1989.

[151] "S&P 500, Data Available at: http://us.spindices.com/indices/equity/sp-500."

[152] "US Census Bureau; Income, Poverty, and Health Insurance Coverage in the United States: Year 2011, Data Available to Download at: http://www.census.gov/hhes/www/cpstables/032012/hhinc/toc.htm."

[153] F. E. James, "Monthly Moving Averages–An Effective Investment Tool?," *Journal of Financial and Quantitative Analysis*, vol. 3, no. 3, pp. 315–326, 1968.

[154] O. Vasicek, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, vol. 5, pp. 177–188, 1977.