Applications of Stochastic Control Theory in The Trading of Stocks and Futures

Applications of Stochastic Control Theory in The Trading of Stocks and Futures

BY

RAPHAEL YAN, M.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS & STATISTICS AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

PHD OF MATHEMATICS

PhD of Mathematics (2020) McMaster University

(Mathematics & Statistics) Hamilton, Ontario, Canada

TITLE: Applications of Stochastic Control Theory in The Trading of Stocks and Futures

AUTHOR: Raphael Yan

M.Sc., (Applied Mathematics)

University of Western Ontario, London, Ontario, Canada

SUPERVISOR: Dr. Traian Pirvu

NUMBER OF PAGES: ix, 89

Abstract

In this thesis, we apply stochastic control methodology to analyze the trading of stocks and futures dynamically. The trader's quest for profits is formulated as an optimal control problem with finite horizon, where the objective is to maximize the expected utility of wealth at the end of the horizon, and the optimized trading strategy is given by the optimal control. Based on various well-established stochastic models of these financial securities, we derive the stochastic differential equations that describe the price dynamics in each model, formulate the utility maximization problems, analyze the associated Hamilton-Jacobi-Bellman (HJB) equations, and solve for the trading strategies in closed form. Numerical examples based on securities traded in the US markets are presented for all models.

Specifically, we investigate pairs trading with cointegrated stock pairs under the Duan-Pliska model, and with volatility index futures where the spot index is modeled as a Central Tendency Ornstein-Uhlenbeck process. We also study optimal trading of commodity futures under the two-factor Schwartz model, and under a more general n-factor Cortazar and Naranjo model. Given the closed-form expressions for the optimal strategies, the value functions, and the wealth processes, we see directly the dependence of the optimal positions on different model parameters, and therefore we can quantify the impact of varying parameter and coefficient values. Qualitatively, we will see, in line with intuition, that the optimal positions in general decrease in magnitude as the volatilities in the underlying factors increase. In all cases we find that the magnitudes of the optimal positions are inversely proportional to the degree of risk aversion, as expected.

In the pairs trading cases where two stocks or two futures contracts are traded, the optimal positions are of opposite signs, corresponding to a long and a short position, where the quantities are given explicitly by the closed-form formulae. Based on parameters calibrated from VIX futures historical data, we find that traders should take bigger positions in the long end of the futures curve. In the WTI oil futures trading example, we see that the optimal positions are insensitive to time to maturity. We also find that the certainty equivalent for trading two contracts simultaneously is significantly greater than that derived from trading only a single contract, regardless of the maturity of either of the single contract. Analogous result holds for the more general n-factor model as well.

Acknowledgements

The writing of this thesis has been a *long* journey. It would not have been possible without Tim Leung. I would like to thank him first of all for agreeing to supervise me during the writing of the papers for this thesis. His guidance is always insightful and helpful. His energy and enthusiasm inspire me and it is always fun to talk to him. I would like to especially thank him for his timely and thoughtful responses to all my questions. With the completion of this thesis, I expect to continue working with him on future projects.

Thanks to Traian Pirvu who kindly agreed to be my supervisor in the McMaster Math department, read my work in detail, guided me through the preparation of this thesis, and provided valuable suggestions and prompt feedback on my revisions. Sincere thanks to Matheus Grasselli for his time spent in supporting my re-admission and in discussing research problems, to Tom Hurd for being on my committee, and to Hans Boden who continued to believe in me. Thanks to Agnès Tourin for co-authoring with me, and for taking the time to meet and discuss. Thanks to all of my family members and relatives who always showed support, and I thank my wife Alice for all her jokes, instead of complaints, on the amount of time I spent on my thesis. Here it is, it is finished.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Optimal trading of financial securities has long been a central area of finance, in practice as well as in research. With the availability of a wide array of assets, such as stocks, bonds, currency, futures, and more recently, cryptocurrencies, participants in the financial markets trade these securities in order to maximize their economic benefits. Regardless of the asset class, once chosen, the three key questions are: 1. whether to buy or sell, 2. how many shares or how many contracts, and 3. when to initiate and subsequently unwind the trade, that is, the timing of the entry and exit. This thesis aims to answer these questions by taking a quantitative approach. While the mathematical techniques and the financial models considered in this thesis are well-established, the novelty of this work is the derivation of the optimal trading strategies in closed form using tools from stochastic control theory.

To the extent that these well-documented models capture the realistic dynamics of these securities, the strategies given by the closed-form formulae are optimal with respect to the trader's risk preference. The main contribution of this thesis is the formulation of a quantitative approach to trading that can be applied in practice, which is illustrated by concrete examples throughout.

The relatively basic, and non-mathematical, approaches to trading based on technical indicators and charting have been in existence since the beginning of the establishment of stock markets [47]. While their effectiveness were examined by studies such as [7, 89], technical analysis nonetheless forms the basis of trading strategies for a certain segment of market participants. On the other hand, academic research in the finance departments traditionally has considered trading in the context of investing and asset allocation, and the theory developed therein is the foundation for quantitative hedge funds whose goal is to generate superior risk-adjusted return, the so-called *alpha* [20]. Another major strand of studies in trading that originates from computer science departments focuses on generation of trading signals through machine learning techniques such as neural networks [46]. Thanks to the availability of market data and computing power, nonlinear statistical models have been developed in order to exploit patterns in time series data. While the main critique of the nonlinear models upon which these black-box trading models are based is the lack of transparency in the relationship of the model parameters, the allure of the ability of artificial intelligence to uncover patterns in large data sets and to generate profitable trading signals continues to attract further research.

In contrast, the approach taken in this thesis originates from the area of mathematical finance, which was essentially started by the Black-Scholes option pricing formula. Since their publication, no-arbitrage models of various underlying securities have been extensively developed, mostly for the purpose for pricing options and other derivatives. Among many others, [67] is a basic introduction to this field, while [23] focuses on interest rate models. [35] and [106] survey jump diffusion models and describe methods to price and hedge derivatives.

Given this wide array of financial models, this thesis attempts to exploit these models, not for pricing derivatives, but for trading the securities. Robert Merton, in addition to his significant contribution to option pricing, applied stochastic control theory to finance, and determined explicitly the consumption and investment rules for an agent with a hyperbolic absolute risk aversion (HARA) utility function, under basic Brownian motion models [92]. In this thesis, we follow the same approach: we formulate optimization problems of finite horizons, and assume the trader fixes a pre-determined future date as the terminal time for optimization, and assume the trader measures economic benefits through a utility function, parametrized by the trader's risk aversion coefficient. In practice, while the finite horizon assumption is a natural one for futures trading, namely, the horizon should be the maturity of the earliest-to-expire contract, for stocks trading, it should be the time at which the investor must liquidate her positions. For traders who cannot maintain overnight positions, the daily market closing time is the correct specification of their horizon; while for long-term retail investors, it is reasonable to set the horizon to the date of retirement. Infinite horizon is appropriate for traders who have no limitations on when their positions have to be liquidated; however this results in a slightly different control problem with other technical difficulties, which is outside of the scope of this thesis.

For more detailed introductions to Merton's problem and solution, and stochastic control theory in general, we refer to [45, 57, 96], and to the lecture notes [105] which provide a clear introduction, and to [58] for a detailed and technical overview of the theory. When the market is frictionless, and consists only of a bond paying a fixed interest rate and a stock driven by geometric Brownian motion, Merton showed that the optimal strategy is to keep a constant proportion of wealth in the stock. To achieve this constant proportion requires constant rebalancing, which is similar to delta-hedging the options within the Black-Scholes framework.

Since Merton's work, there has been a large literature incorporating more realistic features and refining the basic model along different angles. [41] was one of the first papers that considered portfolio optimization in the presence of transaction costs. [117] provided numerical scheme for an investment and consumption problem with transaction costs. Many other features, for examples capital tax gains as considered in [74], default risks as in [69], counter-party risks as in [100], or fixed-income portfolios as in [122], have been explored. [98] summarized theoretical results when the assets are modeled by jump-diffusions. Recently, among the large number of papers that investigated limit order books, the same methodologies are again applied extensively; see the seminal papers [9] and [64], and a recent paper by [1] and the references therein which uses a mean-reverting process to model the underlying price.

These studies illustrate the wide range of realistic financial problems in which stochastic control theory is applicable. In this thesis, we contribute by adding to the literature a novel application of the theory, namely, finding optimal strategies to dynamically trade financial securities. Briefly, the steps consist of 1. deriving the stochastic differential equations (SDEs) that describe the dynamics of the securities under consideration; 2. formulating the value function, which is the maximized expected utility of terminal time $t = T$ wealth, as a function of time t, of amount of wealth at time t, and of the levels of the state variables at time t; 3. obtaining a candidate solution by solving the associated Hamilton-Jacobi-Bellman (HJB) equation; and finally, 4. checking the conditions to *verify* that the candidate solution is equal to the value function.

Exponential utility functions are assumed throughout due to their tractability: it is well-known that the factoring out of the wealth variable reduces the dimension of the equation by one. Another consequence of the separation of wealth variable is the independence of the trading strategy on the level of wealth. This feature effectively addresses the question of the relationship between risk aversion and margin requirements, for both stocks and futures, within our framework. Margin is a good-faith deposit, or an amount of capital, that the trader needs to post to initiate a trade. To avoid margin calls, we would argue that posting a larger margin, that is assuming a larger initial wealth, would suffice, which will not affect the choice of the risk aversion coefficient γ.

Without loss of generality, we can assume that any margin requirement change takes place prior to time $t = 0$. If the margin requirement is increased, the trader would only need to increase the initial wealth endowment, without adjusting γ, since γ only enters our control problems through the objective function, that is, the maximization of terminal wealth at time T. Whatever happens before time zero should not affect the choice of γ. If the trader is more directly concerned about margin calls, for example if the goal is to minimize the probability of violating the margin requirement (over the entire optimization horizon, or, only at time T), then a different objective function should be formulated, which will require totally different solution methodology, and is outside the scope of this thesis. Hence, within our models, increasing margin requirement increases initial wealth w. Since our solutions lead to optimal strategies that are independent of wealth, margin changes will not affect our results.

We now preview the coming chapters and highlight the main contributions. In Chapter 2, we apply these steps to the Duan-Pliska [44] model where the stocks are cointegrated. Cointegration is a statistical concept that quantifies the tendency of different time series to move in tandem. In this chapter, the focus is on the time series of different, but related (for example, in the same sector) stock prices. The idea of *pairs trading* in the stock market has been well-known, and well-applied, for a long time, where the objective, loosely speaking in qualitative terms, is to execute trades when a pair of related stocks diverge from their usual relationship, in the hope that the mean reverting process that drives the drifts in both stocks will revert back to its equilibrium level. The novelty of this chapter, as published in [116], is the application of stochastic control theory in the pairs trading context, which results in a closed-form formula that precisely prescribes the exact positions to take, dynamically as a function of time and stock prices. We then furthermore add correlation into the stocks' dynamics in a later section, which complicates the formula but makes the model more realistic.

We continue the study of optimal trading in Chapter 3, and apply the same methodology in volatility index futures, when the spot index is modeled as a Central Tendency Ornstein-Uhlenbeck process (CTOU) [91]. In contrast to stocks, futures contracts on the same underlying, but of different maturities, typically exhibit a pattern, commonly referred to as the term structure. For example, for futures on a volatility index such as VIX, the contracts of longer maturities typically have lower volatility, as compared to those of shorter maturities, and the decrease is monotonic. Furthermore, since the VIX index itself is based on the implied volatilities of S&P options [31], as the stock market cycles through bull and bear markets, the VIX index exhibits mean-reversion, a property from which the VIX futures will obviously inherit. For trading strategies to be effective, they clearly have to be based on models such as CTOU that captures these various features, and we contribute to the literature by presenting the optimal trading strategy in a closed-form formula for trading volatility index futures.

After studying futures on a volatility index, we next consider commodity futures under the well-known two-factor Schwartz model [107] in Chapter 4. In contrast to models for volatility indices, models for the spot prices of commodities typically include convenience yield, which reflects the value of direct access minus the cost of carry, and can be interpreted as the *dividend yield* for holding the physical commodity. It is a benefit to the holder of the spot commodity, but not to the owner of the futures contract; this benefit is absent, for example, in the VIX futures case where the underlying is not a physical, or even tradable, asset. Furthermore, mean-reversion in the spot commodity price processes is not relevant under the Schwartz model. In light of these differences, we again apply the same steps, and the main contribution of this study is a closed-form formula for the optimal strategy under the Schwartz model.

All of the previous chapters are based on stochastic models driven by two factors. The natural next step is to consider a more general n-factor model. They are by definition more complex than their two factor counterparts, and the complication is necessary since these fit the futures term structures better, as described in [37]. We therefore study optimal trading under the more general n-factor Cortazar and Naranjo model in Chapter 5. The main contributions from extending the closed-form trading strategies to multi-factor models is two-fold: we will show how the CTOU model, and Schwartz model, are nested under this general model when $n = 2$, and we present a tractable formula for the trading strategy, under a sufficiently realistic model, that fits the observed term structures.

In all chapters, we proved that the candidate solution is equal the value function, based on a standard verification theorem presented in Appendix A. Essentially, the theorem requires a uniform integrability condition to hold, and will be proved in the various chapters, under different models. In the cointegrated and correlated stocks pairs trading case in Chapter 2, we derived the sufficient conditions, in the form of inequalities, for uniform integrability to hold. We will see that these inequalities are independent of the risk aversion parameter, which obviates any restriction on the arbitrarily chosen parameter. On the other hand, the uniform integrability condition holds unconditionally for the futures trading cases in Chapter 3-5.

In general, we find that the optimal positions decrease in magnitude as the volatilities in the underlying factors increase, which is in agreement with a risk-averse trader's intuition. In all cases, we find that the magnitudes of the optimal positions are inversely proportional to the degree of risk aversion. Using the calibrated, hence realistic, model parameters, for all the 2-factor models, the positions are shown to have opposite signs, corresponding to one long and one short position.

For the futures trading applications, we find that the certainty equivalent for trading two contracts simultaneously is significantly greater than that derived from trading only a single contract, regardless of the maturity of either of the single contract. Similar result extends to the more general n-factor model. Furthermore, it will be shown, in general, for the futures trading models, the wealth processes controlled by the optimal strategies are submartingales with a positive drift.

In Chapter 3, based on parameters calibrated from VIX futures historical data, we find that traders should take bigger positions in the long end of the futures curve, in line with the monotonic decrease in the volatility term structure. It is an open question whether this can actually be verified in practice. This investigation would first entail collection of firm or individual-specific open interests and volume data on VIX futures; the emergence of *dark pools* certainly complicates the information content of publicly available trading data. Second of all, data for which the individual only trades single maturities would be filtered out. Moreover, even with such data, taking bigger positions on maturities with lower volatilities requires the subject being risk-averse and rational (and therefore follows optimal strategies specified by the solution of the control problem in this thesis), which is itself questionable and has attracted many studies in the area of behavioral finance and economics. We will leave empirical investigations along this line to future studies.

In Chapter 4, based on parameters calibrated from WTI crude oil futures data, we see that the optimal positions change little with respect to maturities; in other words they are insensitive to time to maturity. The optimal strategies are independent of the optimization horizon for all futures trading cases, but not for the cointegrated stock case.

In the models with underlying hidden processes (spot process coupled with the stochastic convenience yield in Schwartz model, stochastic equilibrium in CTOU model, and all the factors in the n-factor model), it is important to note that the optimal holdings are independent of the hidden processes, thereby eliminating the need to estimate the state variables. Moreover, as noted the optimal holdings are also independent of the optimization horizon, which is chosen arbitrarily. These findings greatly facilitate the practical implementation of the strategies, as illustrated by the numerical examples given in the corresponding chapters.

While closed-form solutions to control problems are relatively rare, a class of solvable problems, in addition to Merton's, is the so-called *linear regulator* or *linear quadratic* problems. They are often given as textbook examples (see for example in pg.233 in Chapter 11 of [96], or pg.165 of Chapter 6 of [57], or Chapter 6 in [126]), which feature a linear quadratic objective function, and result in optimal controls and value functions being characterized by two coupled backward stochastic differential equations (BSDE) [13]. However, we are the first to solve a portfolio optimization problem for a wide, yet specialized, class of futures models. This class of futures models is characterized by their time-dependent drifts and volatilities for the log futures price processes which are Gaussian, and this class includes the futures price processes from models such as CTOU, Schwartz and Cortazar-Naranjo as special cases, to be considered in this thesis. This class of futures models are readily estimated using Kalman filtering. The optimal trading strategy for futures of this class is readily obtained in closed form; moreover, no restrictions on the values of parameters and the horizon is necessary. The optimized wealth processes for this class of futures models are submartingales. Based on parameters estimated from actual data, the certainty equivalents from trading more maturities are higher than that from trading fewer in this class of futures models.

1.1 Student's Contribution

For the paper with Agnès Tourin [116] on pairs trading cointegrated stocks, I formulated the control problem and derived the HJB equation. I coded a numerical scheme based on finite-difference in Matlab. Agnès observed that the log substitutions led to a closed-form solution, which obviated the need for numerical calculations. She derived the closed-form solution and stated the verification theorem in the original paper, but the details of the proof was not published. However, the proof in this thesis was completely derived by myself, as suggested by Traian. For the paper, I furthermore collected the data and completed the numerical example.

After the successful publication of [116], I saw the need for a different model for the futures market, due to the presence of term structure. I formulated the problem based on Schwartz's model and derived the closed-form solution, and proceeded to work on the paper, eventually published as [82], with Tim Leung. I originally tested the solution in both commodity and volatility index futures markets. However, Tim suggested that these are two different markets so I should treat them separately. He guided me through my writing and my presentation of the results, such as what graphs to plot, how the optimal wealth process should be derived, and the implications of the optimal strategies.

For the paper on VIX Futures Trading under CTOU with Tim Leung [81], Tim pointed me to the CTOU model, which is mean-reverting so it is a better model for volatility indices. Under his guidance, I applied the same methodology and completed the whole paper by myself.

For the last chapter on the generalization to a n-factor model, Tim pointed me to the work of [37]. I examined the work of [87] and [65], derived the equations in matrix form, and the solution. Traian suggested that I should investigate trading $m < n$ contracts under a n factor model. I completed the chapter myself.

Chapter 2

Pairs Trading Cointegrated Equities

In this chapter, as published in [116], we develop an optimal stochastic control model for analyzing dynamic pairs trading strategies, where the asset price processes are cointegrated. Here we consider a portfolio composed of a risk-free asset and two cointegrated and correlated stocks. As in the classic Merton's problem, the goal is to determine the trading policies that maximize the expected utility of terminal wealth.

The statistical concept of cointegration was initiated by Engle and Granger, as described in [53], who considered security price processes which are not themselves mean-reverting when considered in isolation, but a linear combination of them are. From a statistical perspective, the goal is to select a combination of securities that exhibits mean-reversion, which would have been more difficult or impossible by considering a single security alone. Empirically, the persistence of cointegration in the US stock market has been studied in [32]. The related statistical methodologies, such as detection of cointegration, and parameter estimation, are the tools behind pairs trading.

Pairs trading involves two, usually related, securities, and the trader executes trades upon detection of deviations from the pairs' equilibrium level. By taking opposite positions in the securities, the trader hopes to maintain market-neutrality so that in aggregate the portfolio is immune to systemic market shocks, and to generate profits regardless of the general market trend. While [59], or more recently [24], among many others, studied the empirical profitability in the finance area, mathematical formulation of the problem have been given in [50], [95], and [101]. A survey of the literature on statistical approaches to pairs trading is given in [72], and a non-technical introduction is given in [120].

Our work was inspired by [95] who applied the optimal stochastic control approach to a simplified model of optimal pairs trading. In their model, they only allow positions that are short one stock and long the other, in equal dollar amounts. This is not realistic since in any stock market, the trader is not restricted to buy or sell equal amounts of different stocks: any participant can buy or sell any quantity (up to the margin requirement, of course). We use a similar cointegration model but we relax the above constraint to mirror reality more closely, and allow strategies with arbitrary amounts in each stock. Mathematically, this results in expanding the set of admissible strategies. The model for the cointegrated stocks is taken from [44] who obtain it as the diffusion limit of a discrete time cointegration model. While their focus is to value options on cointegrated assets, our goal is to compute optimal allocations directly on the cointegrated assets. Consequently, we work with the historical probability measure, not a risk-neutral measure. The parameters in our cointegration model can be estimated as in [44], by using a two-step [53] method coupled with a Dicker–Fuller test. It is worth noting that, alternately, a filtering method could be used as in [50].

For the exponential utility function and for a zero risk-free interest rate, we are able to reduce the problem to a one-dimensional linear parabolic partial differential equation (PDE); we compute explicitly the optimal the value function which turns out to be smooth. As in the standard Merton problem, the amounts invested in each stock are inversely proportional to the risk aversion coefficient. From a practitioner's point of view, the investor must pick a particular value for the risk aversion coefficient, since this quality is investor specific, and the quantification of which, given by the coefficient, is different based on the individual's risk appetite. For instance, cutting in half the risk aversion level is equivalent to investing twice as much capital in each stock, as can be seen in the explicit formula given later. Besides, at the optimum, the amounts invested in each stock do not depend on the values of the individual stocks, but rather on a mean reverting process which is a linear combination of the stock prices.

There is a similar paper by [18] who consider a Merton investment problem with a mean reverting asset price. They derive an explicit solution for the value function and the optimal trading strategies by reducing the model to a one-dimensional linear parabolic PDE. Although our application is different than theirs, our calculations and proofs are similar.

Since the publication of [116], which focused on two stocks, the generalization to the trading of a larger number of stocks was carried out in [87] and [29], and more recently in [65]. Based on similar frameworks as ours, [65] generalized the continuous-time control problem to higher dimensions, and furthermore solved the problem in a sequence of optimally chosen trading times. [29] and [87] provided closed-form formula for optimal holdings of a basket of securities when they are cointegrated. While [29] focused on stocks in Nasdaq, the authors in [87] focused on Bitcoins, compared their strategy to different ones derived outside of control theory, and conducted out-of-sample tests. [56] studied the trading of cryptocurrency further, and formulated optimal trading rules based on machine-learning predictions.

There has been many other development on the application of stochastic control theory to pairs and basket trading of multiple cointegrated securities. In [28], the authors formulated a control problem for stock market participants seeking to liquidate a large basket of assets, derived and solved the HJB equation, and took *price impact* into account. [30] examined a mean-variance portfolio optimization problem where the assets are cointegrated, and the cointegration parameters vary in different regimes. [94] applied ideas of pairs trading and stochastic control methodology in *operational management* to determine relative operational performance between firms. [52] studied Lévy-driven Ornstein-Uhlenbeck processes with regime-switching in the context of high-frequency pairs trading – their complicated model captured many realistic aspects in the market, but their strategy is based on a regime classification algorithm, outside the theory of stochastic control.

2.1 Formulation of the optimal stochastic control model

We essentially use for the co-integrated stocks the model derived in [44] as the diffusion limit of a discrete-time model, in the case when there are only two assets. We fix a time horizon $T > 0$. We denote by $S_t^{(1)}$ and $S_t^{(2)}$ the co-integrated stock prices for $t \in [0, T]$. $S_t^{(1)}$, $S_t^{(2)}$ satisfy the stochastic differential equations

$$\begin{aligned} d\log S_t^{(1)} &= \left(\mu_1 - \frac{\sigma_1^2}{2} + \delta z_t\right) dt + \sigma_1 dB_t^{(1)} \\ d\log S_t^{(2)} &= \left(\mu_2 - \frac{\sigma_2^2}{2}\right) dt + \sigma_2 dB_t^{(2)} \end{aligned}$$

where $\left(B_t^{(1)}, B_t^{(2)}\right)$ is a 2-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, the underlying filtration is $\mathcal{F}_t = \sigma((B_s^{(1)}, B_s^{(2)}) : 0 \leq s \leq t)$ and the co-integrating vector z_t is defined by

$$z_t = a + \log S_t^{(1)} + \beta \log S_t^{(2)}. \tag{2.1}$$

The main differences with [44] are the following: first of all, we work under the historical probability measure instead of the risk-neutral probability measure; secondly, for the sake of simplicity, the co-integration term z_t appears only in the drift of the first stock instead of affecting the drifts of the two stocks as in [44]. However, our results can be easily generalized to the fully symmetric case. Thirdly, we omit the deterministic linear trend in our definition of z_t. This only affects the estimation of the coefficients in the dynamics of z_t and our results are readily applicable in the case when a linear deterministic trend is present.

Furthermore, as in [44], we can see that z_t is mean-reverting

$$\begin{aligned} dz_t &= \left(\mu_1 - \frac{\sigma_1^2}{2} + \delta z_t\right) dt + \sigma_1 dB_t^{(1)} + \beta\left(\mu_2 - \frac{\sigma_2^2}{2}\right) dt + \beta\sigma_2 dB_t^{(2)} \\ &= \left(\mu_1 - \frac{\sigma_1^2}{2} + \beta\mu_2 - \beta\frac{\sigma_2^2}{2} + \delta z_t\right) dt + \sigma_1 dB_t^{(1)} + \beta\sigma_2 dB_t^{(2)} \\ &= \alpha(\eta - z_t)dt + \sigma_\beta dB_t \end{aligned}$$

where $\alpha = -\delta$ is the speed of mean reversion, $\sigma_\beta = \sqrt{\sigma_1^2 + \beta^2\sigma_2^2}$,

$$B_t = \frac{\sigma_1}{\sigma_\beta} B_t^{(1)} + \beta\frac{\sigma_2}{\sigma_\beta} B_t^{(2)}$$

is a Brownian motion process adapted to $\mathcal{F}_t$, and

$$\eta = -\frac{1}{\delta}\left(\mu_1 - \frac{\sigma_1^2}{2} + \beta\left(\mu_2 - \frac{\sigma_2^2}{2}\right)\right)$$

is the equilibrium level.

We also assume that there is a risk-free asset such as a money market account. In this chapter, we set the risk-free interest rate to 0 because this allows us to derive a closed-form solution by factoring out the wealth variable. This can be done by a change of numeraire.

Next, we introduce the variable W_s representing the value of the investor's portfolio at time s. Note that we do not require the wealth W_s to be nonnegative. The investor starts at time t with an initial wealth $W_t = w$, in the money market account. Then she invests, at every time $s \in [t, T]$, in both the risk-free money market account and the two stocks. We denote by $\pi_s^{(1)}, \pi_s^{(2)}$ the number of shares held respectively in the first and second stocks at time s. We only consider self-financing strategies and hence, the evolution of the wealth variable is given by

$$dW_s = \pi_s^{(1)} dS_s^{(1)} + \pi_s^{(2)} dS_s^{(2)}.$$

The value of the portfolio W_s satisfies the SDE

$$dW_s = \pi_s^{(1)}(\mu_1 + \delta z_s) S_s^{(1)} ds + \pi_s^{(2)} \mu_2 S_s^{(2)} ds + \pi_s^{(1)} \sigma_1 S_s^{(1)} dB_s^{(1)} + \pi_s^{(2)} \sigma_2 S_s^{(2)} dB_s^{(2)}.$$

Finally, the dynamics of the state variables $W_s, S_s^{(1)}, S_s^{(2)}$ are given by the following controlled system of SDE

$$\begin{aligned}
dW_s &= \pi_s^{(1)}(\mu_1+\delta z_s) S_s^{(1)} ds + \pi_s^{(2)} \mu_2 S_s^{(2)} ds + \pi_s^{(1)} \sigma_1 S_s^{(1)} dB_s^{(1)} + \pi_s^{(2)} \sigma_2 S_s^{(2)} dB_s^{(2)}, && (2.2)\\
dS_s^{(1)} &= (\mu_1+\delta z_s)\, S_s^{(1)} dt + \sigma_1 S_s^{(1)} dB_s^{(1)}, && (2.3)\\
dS_s^{(2)} &= \mu_2 S_s^{(2)} ds + \sigma_2 S_s^{(2)} dB_s^{(2)}, && (2.4)\\
W_t &= w, S_t^{(1)} = s_1, S_t^{(2)} = s_2, && (2.5)
\end{aligned}$$

where z_t is defined in (2.1).

A pair of controls $(\pi^{(1)}, \pi^{(2)})$ is said to be admissible if $\pi^{(1)}, \pi^{(2)}$ are real-valued, progressively measurable, are such that, (2.2),(2.3),(2.4),(2.5) define a unique solution $(W_s, S_s^{(1)}, S_s^{(2)})$ for every time $s \in [0, T]$ and $(\pi^{(1)}, \pi^{(2)}, S^{(1)}, S^{(2)})$ satisfy the integrability condition

$$\mathbb{E}_t \int_t^T (\pi_s^{(1)} S_s^{(1)})^2 + (\pi_s^{(2)} S_s^{(2)})^2 ds < +\infty. \qquad (2.6)$$

where $\mathbb{E}_t$ denote the expectation operator, conditional on the filtration $\mathcal{F}_t$ at time t.

We denote the set of admissible controls at the initial time of investment t, by $\mathcal{A}_t$. Next, we define the value function $u(t, w, s_1, s_2)$ of the following backward dynamic optimization problem: the investor seeks an admissible strategy $(\pi_s^{(1)}, \pi_s^{(2)})$ for every $s \in [t, T)$, that maximizes the utility he derives from wealth at time T, that is,

$$V(t, w, s_1, s_2) = \sup_{(\pi^1, \pi^2) \in \mathcal{A}_t} \mathbb{E}_t[U(W_T)]. \qquad (2.7)$$

Furthermore, in this chapter, we only treat the case of the exponential utility function, i.e.

$$U(w) = -e^{-\gamma w},$$

where $\gamma > 0$ denotes the constant risk aversion coefficient.

One could alternately use the power function $U(w) = \frac{1}{\gamma} c^{\gamma}$ as in [18]. However, for this model, we are unable to carry out the calculations explicitly and derive a solution in closed form. As we will see in the next section, the exponential utility function allows us to factor out the wealth variable and reduce the problem to a two-dimensional PDE. We simply cannot make the same ansatz for the power function.

2.2 The Hamilton-Jacobi-Bellman equation and its solution

Since we do not know a priori the regularity of the value function of the stochastic control problem, we only proceed formally with the hope of obtaining a smooth candidate solution, denoted by $u(t, w, s_1, s_2)$, for the stochastic control problem. We will verify later that our candidate solution coincides with the solution of the stochastic control problem, which is the value function. We use the following short hand notations for partial derivatives

$$u_t = \frac{\partial u}{\partial t}, \quad u_w = \frac{\partial u}{\partial w}, \quad u_{ww} = \frac{\partial^2 u}{\partial w^2},$$

$$u_1 = \frac{\partial u}{\partial s_1}, \quad u_{11} = \frac{\partial^2 u}{\partial s_1^2}, \quad u_2 = \frac{\partial u}{\partial s_2}, \quad u_{22} = \frac{\partial^2 u}{\partial s_2^2},$$

$$u_{w1} = \frac{\partial^2 u}{\partial w \partial s_1}, \quad u_{w2} = \frac{\partial^2 u}{\partial w \partial s_2}, \quad u_{12} = \frac{\partial^2 u}{\partial s_1 \partial s_2}.$$

We expect the candidate solution defined above to satisfy the following HJB partial differential equation:

$$\begin{aligned} u_t + \sup_{\pi_1,\pi_2} \Big[& (\pi_1(\mu_1 + \delta z)s_1 + \pi_2 \mu_2 s_2) u_w + (\mu_1 + \delta z) s_1 u_1 \\ & + \mu_2 s_2 u_2 + \pi_1 \sigma_1^2 s_1^2 u_{w1} + \pi_2 \sigma_2^2 s_2^2 u_{w2} \\ & + \frac{1}{2}(\pi_1^2 \sigma_1^2 s_1^2 + \pi_2^2 \sigma_2^2 s_2^2) u_{ww} + \frac{1}{2}\sigma_1^2 s_1^2 u_{11} + \frac{1}{2}\sigma_2^2 s_2^2 u_{22} \Big] = 0, \end{aligned} \tag{2.8}$$

for all $0 \leq t < T, w, 0 \leq s_1, 0 \leq s_2$, together with the final condition

$$u(T, w, s_1, s_2) \quad = \quad U(w) = -\exp(-\gamma w). \tag{2.9}$$

We introduce the linear operator $\mathcal{L}^{(\pi_1,\pi_2)}$ that will be needed in order to apply the verification theorem in Appendix A:

$$\begin{aligned} \mathcal{L}^{(\pi_1,\pi_2)} u(t, w, s_1, s_2) = & (\pi_1(\mu_1 + \delta z)s_1 + \pi_2 \mu_2 s_2) u_w \\ & + (\mu_1 + \delta z) s_1 u_1 + \mu_2 s_2 u_2 + \pi_1 \sigma_1^2 s_1^2 u_{w1} + \pi_2 \sigma_2^2 s_2^2 u_{w2} \\ & + \frac{1}{2}(\pi_1^2 \sigma_1^2 s_1^2 + \pi_2^2 \sigma_2^2 s_2^2) u_{ww} + \frac{1}{2}\sigma_1^2 s_1^2 u_{11} + \frac{1}{2}\sigma_2^2 s_2^2 u_{22}. \end{aligned}$$

The HJB equation can be written compactly as

$$u_t + \sup_{\pi_1,\pi_2} \mathcal{L}^{(\pi_1,\pi_2)} u = 0.$$

We then proceed by making two classic changes of variable: we first apply the standard logarithmic transformation and secondly, we reduce the number of dimensions in the HJB equation by factoring out the wealth variable. To this end, we let $s_1 = e^x, s_2 = e^y$ and $u(t, w, s_1, s_2) = -e^{-\gamma w} g(t, x, y)$. Then g solves the transformed HJB equation

$$\begin{aligned} -g_t \quad + \quad & \sup_{\pi_1,\pi_2} \Big[(\pi_1(\mu_1 + \delta z)s_1 + \pi_2 \mu_2 s_2)\gamma g - (\mu_1 + \delta z) g_x \\ - \quad & \mu_2 g_y + \pi_1 \sigma_1^2 s_1 \gamma g_x + \pi_2 \sigma_2^2 s_2 \gamma g_y \\ - \quad & \frac{1}{2}(\pi_1)^2 \sigma_1^2 s_1^2 \gamma^2 g - \frac{1}{2}(\pi_2)^2 \sigma_2^2 s_2^2 \gamma^2 g \\ - \quad & \frac{1}{2}\sigma_1^2 (g_{xx} - g_x) - \frac{1}{2}\sigma_2^2 (g_{yy} - g_y) \Big] = 0, \end{aligned} \tag{2.10}$$

subject to

$$g(T, x, y) = 1. \tag{2.11}$$

We notice that the maximization over π_1, π_2 in (2.10) is achieved at

$$\pi_1^* = \frac{(\mu_1 + \delta z)g + \sigma_1^2 g_x}{\sigma_1^2 s_1 \gamma g}, \tag{2.12}$$

$$\pi_2^* = \frac{\mu_2 g + \sigma_2^2 g_y}{\sigma_2^2 s_2 \gamma g}. \tag{2.13}$$

Substituting the formulae (2.12), (2.13) into the PDE (2.10), we obtain

$$\begin{aligned} g_t &= \frac{1}{2}\left(\frac{(\mu_1+\delta z)^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right) g + \frac{\sigma_1^2}{2} g_x + \frac{\sigma_2^2}{2} g_y \\ &\quad + \frac{1}{2}\left(\frac{\sigma_1^2 g_x^2}{g} + \frac{\sigma_2^2 g_y^2}{g}\right) - \frac{1}{2}\left(\sigma_1^2 g_{xx} + \sigma_2^2 g_{yy}\right). \end{aligned} \tag{2.14}$$

Note that we can recover the HJB equation for the standard Merton problem (see [92]) in 2 dimensions by setting $\delta = 0$ in our model; in this well-known case, the value function is independent of the stocks and the dollar amounts invested in each stock are constant. We recall below the closed-form formulae for the value function and the dollar amounts corresponding to the particular case $\delta = 0$.

$$\begin{aligned} g(t,x,y) &= \exp\left(-\frac{\mu_1^2}{2\sigma_1^2}(T-t) - \frac{\mu_2^2}{2\sigma_2^2}(T-t)\right), \\ \pi_1^*(t,x,y)s_1 &= \frac{\mu_1}{\sigma_1^2\gamma}, \\ \pi_2^*(t,x,y)s_2 &= \frac{\mu_2}{\sigma_2^2\gamma}. \end{aligned}$$

In the co-integrated case, the solution is no longer independent of the stocks. It essentially depends on the value of the co-integration process z, rather than on the individual stocks. We can therefore reduce (2.14) to a one dimensional equation in the variable

$$X = \mu_1 + \delta z = \mu_1 + \delta(a + x + \beta y).$$

Furthermore, we combine this change of variable with a logarithmic transformation of the value function, in order to get rid of the nonlinearity. Indeed, simple calculations show that the function $\Phi(t,X) = -\log(g(t,x,y))$ solves the linear parabolic PDE

$$\Phi_t = -\frac{1}{2}\left(\frac{X^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right) + \frac{1}{2}(\sigma_1^2 + \beta\sigma_2^2)(\delta\Phi_X) - \frac{1}{2}(\sigma_1^2 + \beta^2\sigma_2^2)\left(\delta^2\Phi_{XX}\right) \tag{2.15}$$

for any real number X and time $0 \le t < T$ and satisfies the terminal condition

$$\Phi(T,X) = 0. \tag{2.16}$$

It is easy to see that $\Phi(t,X) = a(t)X^2 + b(t)X + c(t)$ is an explicit solution of the linear PDE (2.15),(2.16), where the coefficients a, b, c are given by

$$a(t) = \frac{1}{2}\frac{(T-t)}{\sigma_1^2}, \tag{2.17}$$

$$b(t) = -\frac{1}{4}\frac{(\sigma_1^2 + \beta\sigma_2^2)\delta}{\sigma_1^2}(T-t)^2, \tag{2.18}$$

$$\begin{aligned} c(t) &= \frac{1}{2}\frac{\mu_2^2}{\sigma_2^2}(T-t) + \frac{1}{4}\frac{(\sigma_1^2 + \beta^2\sigma_2^2)\delta^2}{\sigma_1^2}(T-t)^2 + \\ &\quad \frac{1}{24}\frac{(\sigma_1^2 + \beta\sigma_2^2)^2\delta^2}{\sigma_1^2}(T-t)^3. \end{aligned} \tag{2.19}$$

We also compute the optimal policies

$$
\begin{aligned}
\pi_1^* s_1 &= \frac{X}{\sigma_1^2 \gamma} + \frac{\delta(-2a(t)X - b(t))}{\gamma}, \\
\pi_2^* s_2 &= \frac{\mu_2}{\gamma \sigma_2^2} + \frac{\delta\beta(-2a(t)X - b(t)}{\gamma}.
\end{aligned}
$$

After substituting for $a(t), b(t)$ their expressions in (2.17),(2.18), we obtain the dollar amounts invested in the two stocks as

$$
\pi_1^* s_1 = \frac{(\mu_1 + \delta z)}{\gamma \sigma_1^2} - \delta \frac{(\mu_1 + \delta z)}{\gamma \sigma_1^2}(T-t) + \frac{1}{4}\frac{\delta^2(\sigma_1^2 + \beta\sigma_2^2)}{\gamma\sigma_1^2}(T-t)^2, \tag{2.20}
$$

$$
\pi_2^* s_2 = \frac{\mu_2}{\gamma \sigma_2^2} - \beta\delta \frac{(\mu_1 + \delta z)}{\gamma \sigma_1^2}(T-t) + \frac{1}{4}\beta\frac{\delta^2(\sigma_1^2 + \beta\sigma_2^2)}{\gamma\sigma_1^2}(T-t)^2. \tag{2.21}
$$

2.3 Verification Theorem

We must verify that the smooth candidate solution we derived in the previous section is indeed the value function of the stochastic control problem. This can be achieved by proving a *verification* result, which connects the HJB equation to the optimal control problem. The statement of the theorem and proof is presented in Appendix A, Theorem 7. The next step then consists in checking that the assumptions of the verification theorem are satisfied and we show that this is indeed the case, under certain conditions on the parameters in the model.

The main assumption to verify is that the uniform integrability condition holds. The proof of this fact follows along the lines of [18]. More precisely, we apply Lemma 4.3 in [18] to prove the exponential integrability of the square of an Ornstein-Uhlenbeck process, and use it to derive some sufficient conditions on the parameters of our model, under which the uniform integrability condition holds. First, we recall the lemma in [18], adapted to the parameters here.

Lemma 1 (Benth-Karlsen, Lemma 4.3 in [18])**.** *If λ is a constant such that*

$$
\lambda < \frac{|\delta|}{2\sigma_\beta^2 (T-t)},
$$

Then

$$
\mathbb{E}_t\left[\exp\left\{\lambda \int_t^T z_u^2 du\right\}\right] < \infty.
$$

We then state our main result. It provides an explicit solution for the control problem, under some conditions on the parameters.

Theorem 1. *For some $\epsilon > 0$, if*

$$
2(1+\epsilon)\left(\left|\frac{2\delta\mu_1}{\sigma_1^2}\right| + \left|\frac{\delta^2\left(2\mu_1 + \beta\mu_2\right)}{\sigma_1^2}\right| T + \left|\frac{\delta^3\left(\sigma_1^2 + \beta\sigma_2^2\right)}{4\sigma_1^2}\right| T^2\right) \leq \frac{|\delta|}{2\sigma_\beta^2 T}, \tag{2.22}
$$

then the value function of the optimal stochastic problem is given by

$$
V(t, w, s_1, s_2) = -\exp(-\gamma w)\exp(-a(t)X^2 - b(t)X - c(t)),
$$

where

$$
X = \mu_1 + \delta z = \mu_1 + \delta(a + \log s_1 + \beta \log s_2),
$$

and the coefficients $a(\cdot), b(\cdot), c(\cdot)$ are given by (2.17),(2.18),(2.19). Furthermore, the optimal control pair is given by (2.20),(2.21).

The proof of Theorem 1 uses exactly the same arguments as in [18], and we are presenting it here.

Proof. The optimized wealth process W_t^*, that is, the wealth process following the optimal controls, satisfies the SDE

$$\begin{aligned}dW_t^* &= \pi_t^{(1)*}(\mu_1+\delta z_t)S_t^{(1)}dt + \pi_t^{(2)*}\mu_2 S_t^{(2)}dt + \pi_t^{(1)*}\sigma_1 S_t^{(1)}dB_t^{(1)} + \pi_t^{(2)*}\sigma_2 S_t^{(2)}dB_t^{(2)} \\ &= (a_0(t)+a_1(t)z_t+a_2(t)z_t^2)dt + (\sigma_{10}(t)+\sigma_{11}(t)z_t)dB_t^{(1)} + (\sigma_{20}(t)+\sigma_{21}(t)z_t)dB_t^{(2)}, \qquad (2.23)\end{aligned}$$

where

$$a_0(t) = \frac{\mu_2^2\sigma_1^2+\mu_1^2\sigma_2^2}{\gamma\sigma_1^2\sigma_2^2} - \frac{\delta\mu_1^2+\beta\delta\mu_1\mu_2}{\gamma\sigma_1^2}(T-t) \qquad (2.24)$$

$$+\frac{\delta^2\mu_1\left(\sigma_1^2+\beta\sigma_2^2\right)+\beta\delta^2\mu_2\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\sigma_1^2}(T-t)^2, \qquad (2.25)$$

$$a_1(t) = \frac{2\delta\mu_1}{\gamma\sigma_1^2} - \frac{\delta^2\left(2\mu_1+\beta\mu_2\right)}{\gamma\sigma_1^2}(T-t) + \frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\sigma_1^2}(T-t)^2, \qquad (2.26)$$

$$a_2(t) = \frac{\delta^2(1-\delta(T-t))}{\gamma\sigma_1^2}$$

$$\sigma_{10}(t) = \frac{\mu_1}{\gamma\sigma_1} - \frac{\delta\mu_1}{\gamma\sigma_1}(T-t) + \frac{\delta^2\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\sigma_1}(T-t)^2,$$

$$\sigma_{11}(t) = \frac{\delta(1-\delta(T-t))}{\gamma\sigma_1},$$

$$\sigma_{20}(t) = \frac{\mu_2}{\gamma\sigma_2} - \frac{\beta\delta\mu_1}{\gamma\sigma_1}(T-t) + \frac{\beta\delta^2\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\sigma_1}(T-t)^2,$$

$$\sigma_{21}(t) = \frac{\delta^2(T-t)\sigma_2\beta}{\gamma\sigma_1^2}.$$

Let $\epsilon>0$ be an arbitrary positive number. From the verification theorem in Appendix A, it suffices to show that $|u(\tau,W_\tau^*,X_\tau)|^{1+\epsilon}<\infty$ uniformly with respect to any stopping time $t\le\tau\le T$. We have by Cauchy-Schwarz inequality

$$\begin{aligned}\mathbb{E}_t[|u(\tau,W_\tau^*,X_\tau)|^{1+\epsilon}] =& \mathbb{E}_t[\exp(-\gamma(1+\epsilon)W_\tau^*)\exp(-(1+\epsilon)(a(\tau)X_\tau^2+b(\tau)X_\tau+c(\tau)))] \\ \le& \mathbb{E}_t[\exp(-2\gamma(1+\epsilon)W_\tau^*)]^{1/2}\times \qquad (2.27)\\ & \mathbb{E}_t[\exp(-2(1+\epsilon)(a(\tau)X_\tau^2+b(\tau)X_\tau+c(\tau)))]^{1/2}. \qquad (2.28)\end{aligned}$$

For the first expectation in (2.27), we can assume without loss of generality $W_0^*=0$, so

$$\begin{aligned}\mathbb{E}_t[\exp(-2\gamma(1+\epsilon)W_\tau^*)] =& \mathbb{E}_t\Bigg[\exp\Bigg(-2(1+\epsilon)\gamma\int_t^\tau(a_0(u)+a_1(u)z_u+a_2(u)z_u^2)du \\ & -2(1+\epsilon)\gamma\int_t^\tau(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)} - 2(1+\epsilon)\gamma\int_t^\tau(\sigma_{20}(u)+\sigma_{21}(u)z_u)dB_u^{(2)}\Bigg)\Bigg] \\ \le& \mathbb{E}_t\left[\exp\left(-4(1+\epsilon)\gamma\int_t^\tau(a_0(u)+a_1(u)z_u+a_2(u)z_u^2)du\right)\right]^{1/2}\times \qquad (2.29)\\ & \mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^\tau(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}\right)\right]^{1/4}\times \qquad (2.30)\\ & \mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^\tau(\sigma_{20}(u)+\sigma_{21}(u)z_u)dB_u^{(2)}\right)\right]^{1/4}. \qquad (2.31)\end{aligned}$$

by two applications of Cauchy-Schwarz inequality.

First we look at the expectations in (2.30). By Doob's martingale inequality [103] we have

$$\mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^\tau(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}\right)\right]^{1/4}$$
$$\leq 2\sup_{t\leq s\leq T}\mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}\right)\right]^{1/4}. \tag{2.32}$$

We will show this expectation is finite. We will add and subtract $32(1+\epsilon)^2\gamma^2\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du$ in the exponential

$$\mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}\right)\right]^{1/4}$$
$$=\mathbb{E}_t\Bigg[\exp\Bigg(-\int_t^s 8(1+\epsilon)\gamma(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}+\frac{1}{2}\int_t^s(8(1+\epsilon)\gamma)^2(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du$$
$$-32(1+\epsilon)^2\gamma^2\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du\Bigg)\Bigg]^{1/4}. \tag{2.33}$$

Then we split up the above by Cauchy-Schwarz again

$$\leq\mathbb{E}_t\Bigg[\exp\Bigg(-2\int_t^s 8(1+\epsilon)\gamma(\sigma_{10}(u)+\sigma_{11}(u)z_u)dB_u^{(1)}+\int_t^s(8(1+\epsilon)\gamma)^2(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du\Bigg)\Bigg]^{1/8} \tag{2.34}$$
$$\times\,\mathbb{E}_t\left[\exp\left(-64(1+\epsilon)^2\gamma^2\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du\right)\right]^{1/8}. \tag{2.35}$$

Using the martingale property of the first expectation in (2.34) which is finite, we are left with the second expectation in (2.35)

$$\mathbb{E}_t\left[\exp\left(-64(1+\epsilon)^2\gamma^2\int_t^s(\sigma_{10}(u)+\sigma_{11}(u)z_u)^2du\right)\right]^{1/8},$$

but this is clearly bounded by 1.

Similarly, for the expectation in (2.31), we use the same arguments to arrive at

$$\mathbb{E}_t\left[\exp\left(-8(1+\epsilon)\gamma\int_t^\tau(\sigma_{20}(u)+\sigma_{21}(u)z_u)dB_u^{(2)}\right)\right]^{1/4}$$
$$\leq\mathbb{E}_t\left[\exp\left(-64(1+\epsilon)^2\gamma^2\int_t^s(\sigma_{20}(u)+\sigma_{21}(u)z_u)^2du\right)\right]^{1/8}, \tag{2.36}$$

which is also bounded by 1.

Now we look at the expectation in (2.29), which is

$$\mathbb{E}_t\left[\exp\left(-4(1+\epsilon)\gamma\int_t^\tau(a_0(u)+a_1(u)z_u+a_2(u)z_u^2)du\right)\right]^{1/2}.$$

Since $a_0(u)$ is a quadratic function of $u\in[t,T]$, it has a minimum in $[t,T]$. Also, since $a_2(u)$ is positive, we only need to find conditions under which

$$\mathbb{E}_t\left[\exp\left(-4(1+\epsilon)\gamma\int_t^\tau a_1(u)z_udu\right)\right]\equiv\mathbb{E}_t\left[\exp\left(-4(1+\epsilon)\int_t^\tau\tilde{a}_1(u)z_udu\right)\right] \tag{2.37}$$

is finite. We define $\tilde{a}_1(u)\equiv\gamma a_1(u)$, and note that it is independent of γ.

We can see that

$$\tilde{a}_1(u) = \frac{2\delta\mu_1}{\sigma_1^2} - \frac{\delta^2\left(2\mu_1+\beta\mu_2\right)}{\sigma_1^2}(T-u) + \frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\sigma_1^2}(T-u)^2$$

is in general not monotone, not concave or convex, and can be positive or negative in $t \le u \le T$. However, the maximum is bounded by

$$|\tilde{a}_1(u)| \le \left|\frac{2\delta\mu_1}{\sigma_1^2}\right| + \left|\frac{\delta^2\left(2\mu_1+\beta\mu_2\right)}{\sigma_1^2}\right|T + \left|\frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\sigma_1^2}\right|T^2 \equiv \tilde{a}_1^*, \tag{2.38}$$

which we denote by $\tilde{a}_1^*$. We will now find a bound for (2.37). We have

$$\begin{aligned}
\mathbb{E}_t\left[\exp\left(-4(1+\epsilon)\int_t^\tau \tilde{a}_1(u)z_u du\right)\right] &\le \mathbb{E}_t\left[\exp\left(4(1+\epsilon)\left|\int_t^\tau \tilde{a}_1(u)z_u du\right|\right)\right] \\
&\le \mathbb{E}_t\left[\exp\left(4(1+\epsilon)\int_t^\tau |\tilde{a}_1(u)z_u| du\right)\right] \\
&\le \mathbb{E}_t\left[\exp\left(4(1+\epsilon)\tilde{a}_1^*\int_t^\tau |z_u| du\right)\right] \\
&\le \mathbb{E}_t\left[\exp\left(4(1+\epsilon)\tilde{a}_1^*\int_t^T |z_u| du\right)\right] \\
&\le \mathbb{E}_t\left[\exp\left(4(1+\epsilon)\tilde{a}_1^*\int_t^T \left(\frac{1}{2}+\frac{z_u^2}{2}\right) du\right)\right] \\
&= c\mathbb{E}_t\left[\exp\left(2(1+\epsilon)\tilde{a}_1^*\int_t^T z_u^2 du\right)\right].
\end{aligned}$$

We can now apply Lemma 4.3 from [18] and see that a sufficient condition for the finiteness of (2.37) is

$$2(1+\epsilon)\left(\left|\frac{2\delta\mu_1}{\sigma_1^2}\right| + \left|\frac{\delta^2\left(2\mu_1+\beta\mu_2\right)}{\sigma_1^2}\right|T + \left|\frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\sigma_1^2}\right|T^2\right) \le \frac{|\delta|}{2\sigma_\beta^2(T-t)}. \tag{2.39}$$

For the second expectation

$$\mathbb{E}_t[\exp(-2(1+\epsilon)(a(\tau)X_\tau^2 + b(\tau)X_\tau + c(\tau)))]^{1/2}$$

in (2.28), since $a(t)$ and $c(t)$ are positive for all $0 \le t \le T$, we only have to look at

$$\mathbb{E}_t[\exp(-2(1+\epsilon)b(\tau)X_\tau)]^{1/2} = \mathbb{E}_t[\exp(-2(1+\epsilon)b(\tau)(\mu_1+\delta z_\tau))]^{1/2}. \tag{2.40}$$

Recall that

$$z_t = e^{-\alpha t}z_0 + \eta(1-e^{-\alpha t}) + \sigma_\beta e^{-\alpha t}\int_0^t e^{\alpha s}dB_s$$

which is normally distributed with

$$z_t \sim \mathcal{N}\left(e^{-\alpha t}z_0 + \eta(1-e^{-\alpha t}), \frac{1-e^{-2\alpha t}}{2\alpha}\right).$$

Now since the Ito integral $\int_0^t e^{\alpha s}dB_s \equiv M_t$ is a martingale, and since by Jensen's inequality e^{kM_t} is a submartingale for any $k \in \mathbb{R}$, and in particular it is a positive submartingale, we can apply again Doob's martingale inequality [103] in Eq (2.40). The function $b(t)$

$$b(t) = -\frac{1}{4}\frac{(\sigma_1^2+\beta\sigma_2^2)\delta}{\sigma_1^2}(T-t)^2$$

is zero at $t = T$ and has a maximum bounded by $b(0)$. By defining

$$k \equiv -2(1+\epsilon)b(0)\delta = \frac{(1+\epsilon)}{2}\frac{(\sigma_1^2+\beta\sigma_2^2)\delta^2}{\sigma_1^2}T^2,$$

and picking some $\xi > 0$, we will estimate the expectation in (2.40) as

$$\mathbb{E}_t[\exp(kM_\tau)] = \left\| \exp\left(\frac{k}{1+\xi}M_\tau\right)\right\|_{1+\xi}^{1+\xi} \tag{2.41}$$

$$\leq \left\| \sup_{t\leq s\leq T} \exp\left(\frac{k}{1+\xi}M_s\right)\right\|_{1+\xi}^{1+\xi} \tag{2.42}$$

$$\leq \left(1+\frac{1}{\xi}\right)^{1+\xi} \sup_{t\leq s\leq T} \mathbb{E}_t[\exp(kM_s)] \tag{2.43}$$

$$\leq c \tag{2.44}$$

where c is a positive constant independent of the stopping time τ. $\square$

We can see that the condition (2.22) on the parameters translates into a limit on the time-horizon. Beyond a certain time horizon, we cannot guarantee that the solution we computed is the unique solution of the control problem. However, this condition is only a sufficient condition, which means the violation of the condition *does not necessarily* invalidate the solution.

2.4 A more general model with correlations

It is also possible to incorporate correlations between the stocks into the above model and we present briefly this extension in this section. The dynamics will read in this case

$$\begin{aligned} dW_t &= \pi_1 dS_t^{(1)} + \pi_2 dS_t^{(1)} \\ dS_t^{(1)} &= (\mu_1 + \delta z_t)\, S_t^{(1)} dt + \sigma_1 S_t^{(1)} dB_t^{(1)} \\ dS_t^{(2)} &= \mu_2 S_t^{(2)} dt + \sigma_2 S_t^{(2)} (\rho dB_t^{(1)} + \sqrt{1-\rho^2} dB_t^{(2)}) \end{aligned}$$

where $-1 \leq \rho \leq 1$ denotes the correlation coefficient between $B_t^{(1)}$ and $B_t^{(2)}$, and the co-integrating vector z_t is still defined by

$$z_t = a + \log S_t^{(1)} + \beta \log S_t^{(2)}.$$

Substituting, we find that the wealth satisfies the SDE

$$\begin{aligned} dW_t &= \pi_1(\mu_1+\delta z_t) S_t^{(1)} dt + \pi_2\mu_2 S_t^{(2)} dt + \pi_1\sigma_1 S_t^{(1)} dB_t^{(1)} \\ &\quad + \pi_2\sigma_2 S_t^{(2)} (\rho dB_t^{(1)} + \sqrt{1-\rho^2} dB_t^{(2)}). \end{aligned}$$

In this case, z_t satisfies the SDE

$$\begin{aligned} dz_t &= \left(\mu_1 - \frac{\sigma_1^2}{2} + \delta z_t\right) dt + \sigma_1 dB_t^{(1)} + \beta\left(\mu_2 - \frac{\sigma_2^2}{2}\right) dt + \beta\sigma_2 dB_t^{(2)} \\ &= \left(\mu_1 - \frac{\sigma_1^2}{2} + \beta\mu_2 - \beta\frac{\sigma_2^2}{2} + \delta z_t\right) dt + \sigma_1 dB_t^{(1)} \\ &\quad + \beta\sigma_2\left(\rho dB_t^{(1)} + \sqrt{1-\rho^2} dB_t^{(2)}\right) \\ &= \alpha(\eta - z_t) dt + \sigma_\beta dB_t \end{aligned}$$

where $\alpha = -\delta$ is the speed of mean reversion, $\sigma_\beta = \sqrt{\sigma_1^2 + \beta^2\sigma_2^2 + 2\beta\sigma_1\sigma_2\rho}$,

$$B_t = \frac{\sigma_1 + \beta\sigma_2\rho}{\sigma_\beta} B_t^{(1)} + \beta\frac{\sigma_2\sqrt{1-\rho^2}}{\sigma_\beta} B_t^{(2)}$$

is a Brownian motion process adapted to $\mathcal{F}_t$, and

$$\eta = -\frac{1}{\delta}\left(\mu_1 - \frac{\sigma_1^2}{2} + \beta\left(\mu_2 - \frac{\sigma_2^2}{2}\right)\right)$$

is the equilibrium level. The value function of this stochastic control problem is defined, as earlier, in (2.7).

Next, we expect the candidate solution $u(t, w, s_1, s_2)$ to solve the HJB equation

$$\begin{aligned} u_t &+ \sup_{\pi_1,\pi_2}\Big[(\pi_1(\mu_1+\delta z)s_1 + \pi_2\mu_2 s_2)u_w + (\mu_1+\delta z)s_1 u_1 + \mu_2 s_2 u_2 \\ &+ \pi_1\sigma_1^2 s_1^2 u_{w1} + \pi_2\rho\sigma_1\sigma_2 s_1 s_2 u_{w1} + \pi_2\sigma_2^2 s_2^2 u_{w2} \\ &+ \pi_1\rho\sigma_1\sigma_2 s_1 s_2 u_{w2} + \frac{1}{2}(\pi_1^2\sigma_1^2 s_1^2 + \pi_2^2\sigma_2^2 s_2^2 + \rho\pi_1\pi_2\sigma_1\sigma_2 s_1 s_2)u_{ww} \\ &+ \frac{1}{2}\sigma_1^2 s_1^2 u_{11} + \frac{1}{2}\sigma_2^2 s_2^2 u_{22} + \rho\sigma_1\sigma_2 s_1 s_2 u_{12}\Big] = 0 \end{aligned}$$

The function $g(t, x, y)$ satisfies the HJB equation

$$\begin{aligned} -g_t &+ \sup_{\pi_1,\pi_2}\Big[(\pi_1(\mu_1+\delta z)s_1 + \pi_2\mu_2 s_2)\gamma g - (\mu_1+\delta z)g_x - \mu_2 g_y \\ &+ \pi_1\sigma_1^2 s_1\gamma g_x + \pi_2\gamma\rho\sigma_1\sigma_2 s_2 g_x + \pi_2\sigma_2^2 s_2\gamma g_y + \pi_1\gamma\rho\sigma_1\sigma_2 s_1 g_y \\ &- \frac{1}{2}(\pi_1)^2\sigma_1^2 s_1^2\gamma^2 g - \frac{1}{2}(\pi_2)^2\sigma_2^2 s_2^2\gamma^2 g - \gamma^2\pi_1\pi_2\rho\sigma_1\sigma_2 s_1 s_2 g \\ &- \frac{1}{2}\sigma_1^2(g_{xx} - g_x) - \frac{1}{2}\sigma_2^2(g_{yy} - g_y) - \rho\sigma_1\sigma_2 g_{xy}\Big] = 0, \end{aligned} \tag{2.45}$$

subject to

$$g(T, x, y) = 1. \tag{2.46}$$

The optimal controls are

$$\pi_1^* = \frac{(\mu_1+\delta z)}{\gamma(1-\rho^2)\sigma_1^2 s_1} + \frac{g_x}{\gamma g s_1} - \rho\frac{\mu_2}{\gamma(1-\rho^2)\sigma_1\sigma_2 s_1}, \tag{2.47}$$

$$\pi_2^* = \frac{\mu_2}{\gamma(1-\rho^2)\sigma_2^2 s_2} + \frac{g_y}{\gamma g s_2} - \rho\frac{(\mu_1+\delta z)}{\gamma(1-\rho^2)\sigma_1\sigma_2 s_2}. \tag{2.48}$$

After substituting the controls into the HJB equation, we obtain the PDE

$$\begin{aligned} g_t = &\left(\frac{1}{2}\frac{(\mu_1+\delta z)^2}{(1-\rho^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho^2)\sigma_2^2} - \frac{\rho\mu_2(\mu_1+\delta z)}{(1-\rho^2)\sigma_1\sigma_2}\right)g \\ &+ \frac{\sigma_1^2}{2}g_x + \frac{\sigma_2^2}{2}g_y + \frac{1}{2}\sigma_1^2\frac{g_x^2}{g} + \frac{1}{2}\sigma_2^2\frac{g_y^2}{g} + \rho\sigma_1\sigma_2\frac{g_x g_y}{g} \\ &- \frac{1}{2}\sigma_2^2\sigma_1^2 g_{xx} - \frac{1}{2}\sigma_2^2 g_{yy} - \rho\sigma_1\sigma_2 g_{xy}. \end{aligned}$$

Replacing the variables (x, y) by the single variable $X = \mu_1 + \delta z$ and by using the exponential change of variable $g = \exp(-\Phi)$, we reduce the problem to the linear parabolic PDE

$$\begin{aligned} \Phi_t = &-\frac{1}{1-\rho^2}\left(\frac{1}{2}\left(\frac{X^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right) - \rho\frac{\mu_2 X}{\sigma_1\sigma_2}\right) + \frac{1}{2}(\sigma_1^2 + \beta\sigma_2^2)(\delta\Phi_X) \\ &-\frac{1}{2}(\sigma_1^2 + \beta^2\sigma_2^2 + 2\sigma_1\sigma_2\beta\rho)\left(\delta^2\Phi_{XX}\right) \end{aligned} \tag{2.49}$$

for any real number X and time $0 \leq t < T$ and is subject to the terminal condition

$$\Phi(T,X) \quad = \quad 0. \tag{2.50}$$

The above linear PDE has the explicit solution $\phi(t,X) = a(t)X^2 + b(t)X + c(t)$ where

$$a(t) = \frac{1}{2}\frac{(T-t)}{(1-\rho^2)\sigma_1^2}, \tag{2.51}$$

$$b(t) = -\frac{1}{4}\frac{(T-t)^2}{(1-\rho^2)\sigma_1^2}(\sigma_1^2+\beta\sigma_2^2)\delta - \frac{\rho}{(1-\rho^2)}\frac{\mu_2(T-t)}{\sigma_1\sigma_2}, \tag{2.52}$$

$$\begin{aligned} c(t) &= \frac{1}{2}\frac{(T-t)\mu_2^2}{(1-\rho^2)\sigma_2^2} + \frac{1}{4}\frac{(\sigma_1^2+\beta^2\sigma_2^2+2\sigma_1\sigma_2\beta\rho)\delta^2}{(1-\rho^2)\sigma_1^2}(T-t)^2 \\ &+\frac{1}{4}\frac{\delta\rho}{(1-\rho^2)}\frac{\mu_2(\sigma_1^2+\beta\sigma_2^2)(T-t)^2}{\sigma_1\sigma_2} \\ &+\frac{1}{24}\frac{(\sigma_1^2+\beta\sigma_2^2)^2\delta^2}{(1-\rho^2)\sigma_1^2}(T-t)^3. \end{aligned} \tag{2.53}$$

Finally, we substitute the values of g, g_x and g_y into the formulae (2.47),(2.48), in order to obtain the optimal policies

$$\begin{aligned} \pi_1^* s_1 &= \frac{\mu_1+\delta z}{\gamma(1-\rho^2)\sigma_1^2} + \frac{\delta(-2a(t)(\mu_1+\delta z)-b(t))}{\gamma} - \frac{\rho\mu_2}{\gamma(1-\rho^2)\sigma_1\sigma_2}, \\ \pi_2^* s_2 &= \frac{\mu_2}{\gamma(1-\rho^2)\sigma_2^2} + \frac{\delta\beta(-2a(t)(\mu_1+\delta z)-b(t))}{\gamma} - \frac{\rho(\mu_1+\delta z)}{\gamma(1-\rho^2)\sigma_1\sigma_2}. \end{aligned}$$

Written out explicitly,

$$\begin{aligned} \pi_1^* s_1 &= \frac{\mu_1+\delta z}{\gamma(1-\rho^2)\sigma_1^2} - \frac{\rho\mu_2}{\gamma(1-\rho^2)\sigma_1\sigma_2} \\ &+ \frac{\delta}{\gamma(1-\rho^2)}\left(-\frac{(\mu_1+\delta z)}{\sigma_1^2} + \frac{\rho\mu_2}{\sigma_1\sigma_2}\right)(T-t) \\ &+ \frac{1}{4}\frac{\delta^2(\sigma_1^2+\beta\sigma_2^2)}{\gamma(1-\rho^2)\sigma_1^2}(T-t)^2, \end{aligned} \tag{2.54}$$

$$\begin{aligned} \pi_2^* s_2 &= \frac{\mu_2}{\gamma(1-\rho^2)\sigma_2^2} - \frac{\rho(\mu_1+\delta z)}{\gamma(1-\rho^2)\sigma_1\sigma_2} \\ &+ \frac{\delta\beta}{\gamma(1-\rho^2)}\left(-\frac{(\mu_1+\delta z)}{\sigma_1^2} + \frac{\rho\mu_2}{\sigma_1\sigma_2}\right)(T-t) \\ &+ \frac{1}{4}\frac{\delta^2\beta(\sigma_1^2+\beta\sigma_2^2)}{\gamma(1-\rho^2)\sigma_1^2}(T-t)^2. \end{aligned} \tag{2.55}$$

We summarize our findings in the following Theorem.

Theorem 2. *For some $\epsilon > 0$, if*

$$2(1+\epsilon)\left(\left|\frac{2\delta\left(\mu_1\sigma_2-\rho\mu_2\sigma_1\right)}{\left(1-\rho^2\right)\sigma_1^2\sigma_2}\right| + \left|\frac{\delta^2\left(\rho\mu_2\sigma_1-2\mu_1\sigma_2-\beta\mu_2\sigma_2\right)}{\left(1-\rho^2\right)\sigma_1^2\sigma_2}\right|T + \left|\frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\left(1-\rho^2\right)\sigma_1^2}\right|T^2\right) \leq \frac{|\delta|}{2\sigma_\beta^2 T} \tag{2.56}$$

the value function of the optimal stochastic problem with correlation coefficient ρ is given by

$$u(t,w,s_1,s_2) = -\exp(-\gamma w)\exp(-a(t)X^2 - b(t)X - c(t)),$$

where

$$X = \mu_1 + \delta z = \mu_1 + \delta(a + \log s_1 + \beta \log s_2),$$

and the coefficients $a(\cdot), b(\cdot), c(\cdot)$ are given above by (2.51), (2.52) and (2.53). Furthermore, the optimal control pair is given by (2.54),(2.55).

Proof. The optimized wealth process W_t^*, that is, the wealth process following the optimal controls, satisfies the SDE

$$\begin{aligned}
dW_t^* =&\pi_t^{(1)*}(\mu_1+\delta z_t)S_t^{(1)}dt+\pi_t^{(2)*}\mu_2 S_t^{(2)}dt+\pi_t^{(1)*}\sigma_1 S_t^{(1)}dB_t^{(1)}+\pi_t^{(2)*}\sigma_2 S_t^{(2)}(\rho dB_t^{(1)}+\sqrt{1-\rho^2}dB_t^{(2)})\\
=&\pi_t^{(1)*}(\mu_1+\delta z_t)S_t^{(1)}dt+\pi_t^{(2)*}\mu_2 S_t^{(2)}dt+\left(\pi_t^{(1)*}\sigma_1 S_t^{(1)}+\rho\pi_t^{(2)*}\sigma_2 S_t^{(2)}\right)dB_t^{(1)}+\sqrt{1-\rho^2}\pi_t^{(2)*}\sigma_2 S_t^{(2)}dB_t^{(2)}\\
=&(a_0(t)+a_1(t)z_t+a_2(t)z_t^2)dt+(\sigma_{10}(t)+\sigma_{11}(t)z_t)dB_t^{(1)}+(\sigma_{20}(t)+\sigma_{21}(t)z_t)dB_t^{(2)},
\end{aligned}$$

where

$$\begin{aligned}
a_0(t)=&\frac{\mu_1^2}{\gamma\left(1-\rho^2\right)\sigma_1^2}+\frac{\mu_2^2}{\gamma\left(1-\rho^2\right)\sigma_2^2}-\frac{2\rho\mu_1\mu_2}{\gamma\left(1-\rho^2\right)\sigma_1\sigma_2}-\left(\frac{\delta(\mu_1^2+\beta\mu_1\mu_2)}{\gamma\left(1-\rho^2\right)\sigma_1^2}-\frac{\delta\rho(\mu_1\mu_2+\beta\mu_2^2)}{\gamma\left(1-\rho^2\right)\sigma_1\sigma_2}\right)(T-t)\\
&+\frac{\left(\delta^2\mu_1+\beta\delta^2\mu_2\right)\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\left(1-\rho^2\right)\sigma_1^2}(T-t)^2,
\end{aligned}$$

$$a_1(t)=\frac{2\delta\left(\mu_1\sigma_2-\rho\mu_2\sigma_1\right)}{\gamma\left(1-\rho^2\right)\sigma_1^2\sigma_2}+\frac{\delta^2\left(\rho\mu_2\sigma_1-2\mu_1\sigma_2-\beta\mu_2\sigma_2\right)}{\gamma\left(1-\rho^2\right)\sigma_1^2\sigma_2}(T-t)+\frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\gamma\left(1-\rho^2\right)\sigma_1^2}(T-t)^2,$$

$$a_2(t)=\delta^2\frac{1-\delta(T-t)}{\gamma(1-\rho^2)\sigma_1^2},$$

$$\sigma_{11}(t)=\frac{\delta}{\gamma\sigma_1}-\delta^2\frac{\sigma_1\sigma_2+\beta\rho\sigma_2^2}{\gamma\left(1-\rho^2\right)\sigma_1^2\sigma_2}(T-t),$$

$$\begin{aligned}
\sigma_{10}(t)=&\frac{\mu_1}{\gamma\sigma_1}+\delta^2\frac{\sigma_1^3\sigma_2+\beta\rho\sigma_1^2\sigma_2^2+\beta\sigma_1\sigma_2^3+\beta^2\rho\sigma_2^4}{4\gamma\left(1-\rho^2\right)\sigma_1^2\sigma_2}(T-t)^2\\
&-\delta\frac{\mu_1\sigma_1\sigma_2+\beta\rho\mu_1\sigma_2^2-\rho\mu_2\sigma_1\left(\sigma_1+\beta\rho\sigma_2\right)}{\gamma\left(1-\rho^2\right)\sigma_1^2\sigma_2}(T-t),
\end{aligned}$$

$$\sigma_{20}(t)=\frac{\mu_2\sigma_1^2-\rho\mu_1\sigma_1\sigma_2}{\gamma\sqrt{1-\rho^2}\sigma_1^2\sigma_2}+\frac{\beta\delta(\rho\mu_2\sigma_1\sigma_2-\mu_1\sigma_2^2)}{\gamma\sqrt{1-\rho^2}\sigma_1^2\sigma_2}(T-t)+\frac{\delta^2(\beta\sigma_1^2\sigma_2^2+\beta^2\sigma_2^4)}{4\gamma\sqrt{1-\rho^2}\sigma_1^2\sigma_2}(T-t)^2,$$

$$\sigma_{21}(t)=-\frac{\delta\rho}{\gamma\sqrt{1-\rho^2}\sigma_1}-\frac{\beta\delta^2\sigma_2}{\gamma\sqrt{1-\rho^2}\sigma_1^2}(T-t).$$

We proceed using the same arguments as in the uncorrelated case, and note that again $a_2(t)$ is positive, and that again we can define $\tilde{a}_1(t)\equiv\gamma a_1(t)$, which has γ canceled out. Again, the maximum in $\tilde{a}_1(t)$ can be bounded by

$$\left|\frac{2\delta\left(\mu_1\sigma_2-\rho\mu_2\sigma_1\right)}{\left(1-\rho^2\right)\sigma_1^2\sigma_2}\right|+\left|\frac{\delta^2\left(\rho\mu_2\sigma_1-2\mu_1\sigma_2-\beta\mu_2\sigma_2\right)}{\left(1-\rho^2\right)\sigma_1^2\sigma_2}\right|T+\left|\frac{\delta^3\left(\sigma_1^2+\beta\sigma_2^2\right)}{4\left(1-\rho^2\right)\sigma_1^2}\right|T^2,$$

from which the condition (2.56) holds after applying Benth-Karlsen Lemma 4.3. □

2.5 Example

We provide an example to illustrate our results. We wish to emphasize that we are not conducting a comprehensive study here, on whether stock market indices are co-integrated or not; this exercise is merely to illustrate our equations with a concrete real-life example. More precisely, we did not include a detection component in our algorithm. We simply browsed a number of arbitrary data sets and we picked one among those. We collected on October 17, 2011, minute-by-minute data, on two stocks traded on the New York Stock Exchange, Goldman Sachs Group, Incorporated, with ticker symbols GS, and JPMorgan Chase and Company, with ticker symbol JPM. This gives us a 2 dimensional time series with 390 data points. As in [44], we follow the standard two-step Engle-Granger methodology coupled with a Dickey-Fuller test [53], to test for co-integration and estimate the parameters in our co-integration model. We also performed in addition a Phillips-Ouliaris [102] test whose outcome further confirmed that our time series were co-integrated. We also refer to the book by Enders [51] for a general presentation of co-integration tests. We first run a regression as in equation (7) in [44], but without a time trend variable, which seems reasonable for an intra-day data set. Also following [44], we provide the Ordinary Least Square (OLS) standard errors for the regression results. For the parameters in the cointegrating vector z_t, which is a Ornstein-Uhlenbeck process, we use the Maximum Likelihood Estimators (MLE) and display asymptotic standard errors as given in [115]. For the normally distributed (log) stock price increments after removing the effect of the cointegrating vector in the drift terms, we again provide Maximum Likelihood Estimates, and the standard errors are based on the Fisher information matrix as described in [62]. For the standard error of the correlation parameter, a formula is given in [8]. Running the so-called Augmented-Dickey-Fuller (up to 16 lags) test suggests that $\log(S^{(1)})$ and $\log(S^{(2)})$ are co-integrated at the 5% level. We perform, in addition, the Phillips-Ouliaris [102] test. The variance-ratio test statistic suggests that after detrending with constant and linear trend, the data are co-integrated at the 5% level. We obtain the following annualized parameters in Table 2.1, with standard errors displayed in parenthesis underneath:

a	β	δ	η	σ_β
-3.105148	-0.5363258	-3561.194	$9.298768e-04$	0.3081191
(0.10374)	(0.03816)	(1366.112)	(0.001375323)	(0.009598445)
μ_1	μ_2	ρ	σ_1	σ_2
4.662472	2.623764	0.61181479	0.3828025	0.5076713
(1909.959)	(2532.981)	(0.0009814703)	(0.02748362)	(0.03644868)

Table 2.1: Parameters for the cointegration model with correlation

For this set of parameters, we estimate the maximal time horizon T satisfying the sufficient condition (2.56) to be 0.75 day, assuming 6.5 hours trading per day, and 252 trading days per year. For the purpose of illustrating, we choose to go over the bound of 0.75 day, or about 300 minutes, and plot our stock prices and the optimal policies for a whole trading day. We show the stock prices $S^{(1)}$, $S^{(2)}$, as well as the optimal policies π_1^*, π_2^* in figure 2.1. We then present in figure 2.2, the ratio $|\pi_1^*/\pi_2^*|$ and the cumulative profit and loss function.

As expected, in a pairs trading setting, the controls are opposite in sign. For this data set and for a risk tolerance $\gamma = 0.1$, a significant profit is instantly realized. The profit then fluctuates throughout the day but remains strongly positive and by the end of the day, it is approximately \$1,348. Of course, this figure does not take into account the cost of borrowing or transaction costs which are both assumed to be 0 in this model. We also notice that the positions, which are large during the first half of the day, are both progressively unwound in the second half, ending close to 0 by the end of the trading day.

We now examine certainty equivalent, as a function of time t and wealth w at t, which is denoted by $C(t,w)$ in the following. It is the inverse of the utility function; for the cointegration model it is

$$C(t,w) = w + \frac{1}{\gamma}\left(a(t)X^2 + b(t)X + c(t)\right).$$

We will assume $w = 0$ and will be plotting $C(t,0)$. Furthermore, we examine certainty equivalent in the absence of cointegration, that is, when $\delta = 0$, as in the classical Merton case. The certainty equivalent in this case is denoted by $\tilde{C}(t,w)$, and is given by

$$\tilde{C}(t,w) = w + \frac{1}{\gamma}\left(\frac{\tilde{\mu}_1^2\tilde{\sigma}_2^2 + \tilde{\mu}_2^2\tilde{\sigma}_1^2 - 2\tilde{\rho}\tilde{\mu}_1\tilde{\mu}_2\tilde{\sigma}_1\tilde{\sigma}_2}{2(1-\tilde{\rho}^2)\tilde{\sigma}_1^2\tilde{\sigma}_2^2}\right)(T-t),$$

Figure 2.1: Stocks and optimal policies

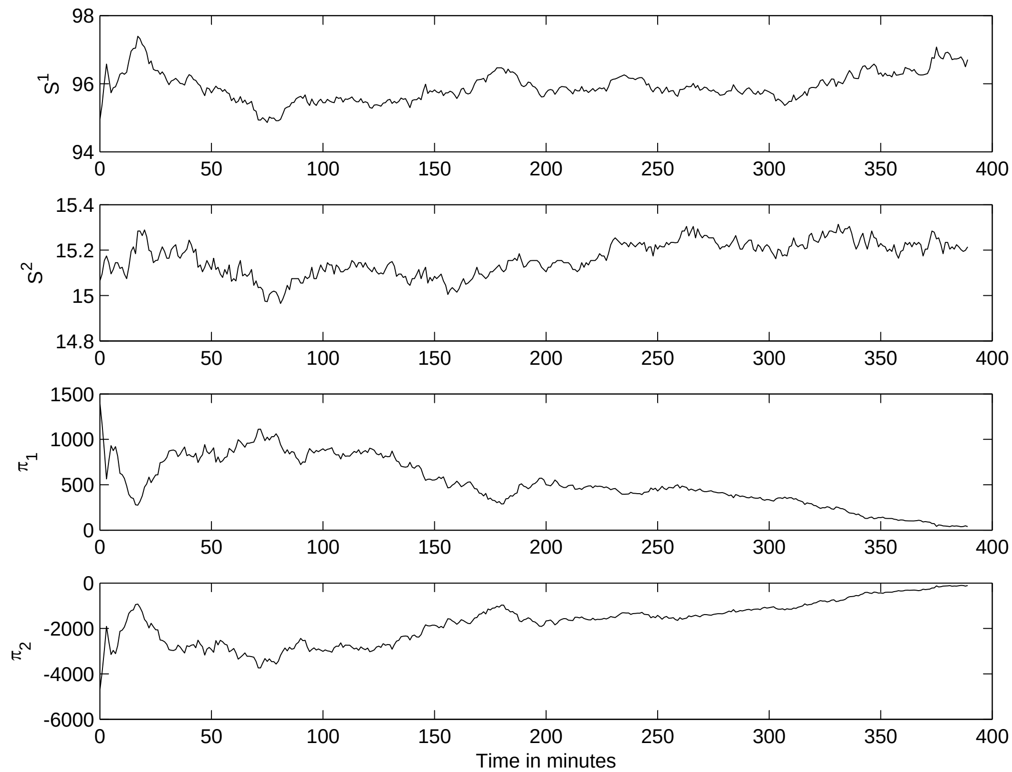
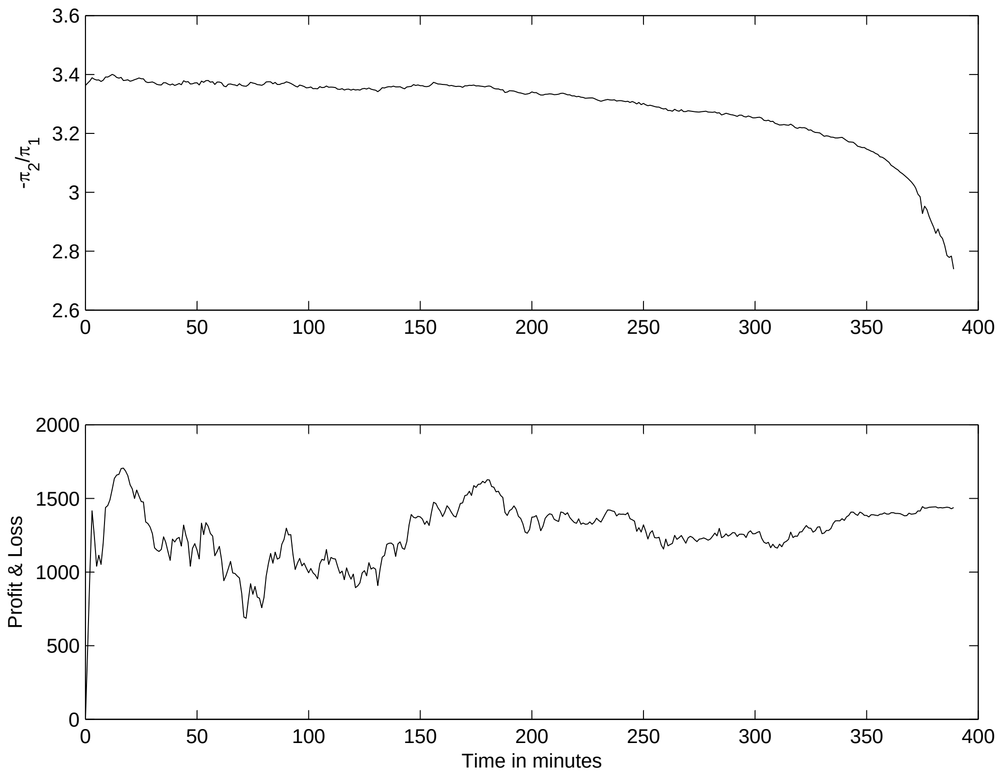


Figure 2.2: The optimal ratio and the Profit and Loss processes



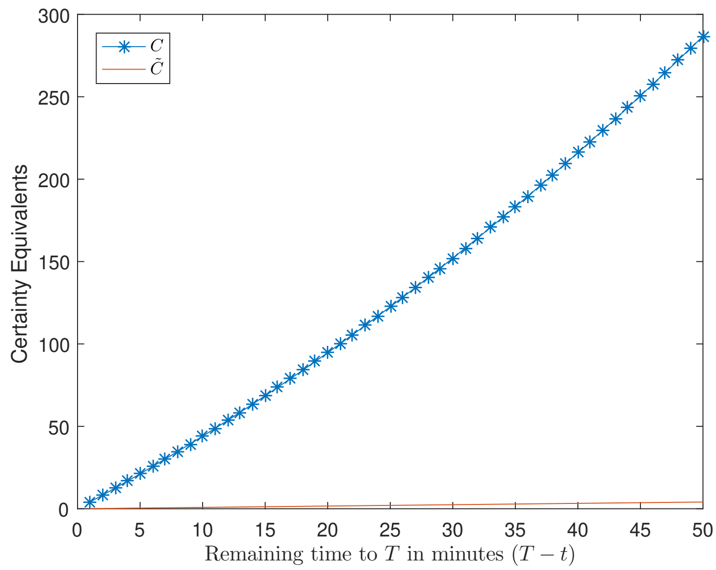
$\tilde{\mu}_1$	$\tilde{\mu}_2$	$\tilde{\rho}$	$\tilde{\sigma}_1$	$\tilde{\sigma}_2$
4.722745	2.623764	0.6130236	0.3841824	0.5076713
(1916.844)	(2532.981)	(0.001001615)	(0.02758269)	(0.03644868)

Table 2.2: Parameters for the classical Merton model when there is no cointegration

where the parameter values are re-estimated in the absence of cointegration, and are displayed in Table 2.2.

We plot the results on Fig 2.3, for the last 50 minutes in a trading day. The certainty equivalent for the cointegration case is much higher. In fact, if we set $\delta = 0$, and use the same set of parameters in the cointegration model without re-estimation, the certainty equivalent is similarly much lower than that derived from the cointegrated case. We can observe that trading in a market in the presence of cointegration yields significantly higher utility.

Figure 2.3: Comparison of Certainty Equivalents

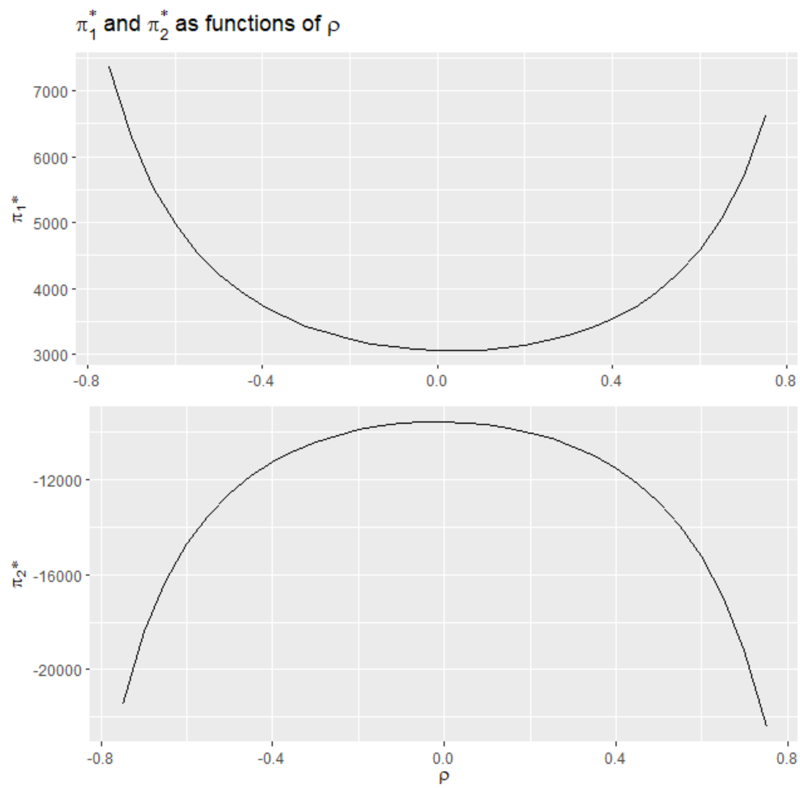


We further study the effect of correlation on the optimal strategies, with cointegration (Fig 2.4) and without (Fig 2.5). In Fig 2.4, it can be seen that the optimal position sizes are at their minimums when $\rho = 0$. The interpretation is that when the stocks are uncorrelated, there is less offsetting between the stocks; the higher risk leads to lower trade sizes. However, even with negative correlation, the optimal strategy is still to have opposite (long S_1 and short S_2) positions, when cointegration is present. When cointegration is absent ($\delta = 0$, Fig 2.5), at sufficiently negative correlation the optimal strategy is to keep both stocks in long positions. This shows the cointegration effect dominates correlation, and the model is a non-trivial generalization of that in the classic Merton's problem. However, this also shows that while the incorporation of cointegration increases certainty equivalent (Fig 2.3), it also increases the risk of model or parameter mis-specification.

2.6 Summary

In this chapter, we propose a dynamic model for pairs trading based on the theory of optimal stochastic control, and we illustrate the applicability of our method with minute-by-minute historical stock data. In our model, the two stock processes are cointegrated, correlated, and have constant volatility. We note that the asymmetry in the drifts, that is, the presence of the cointegrating vector z_t only in (2.3) but not

Figure 2.4: Effect of Correlation on π_i^*, in the presence of cointegration

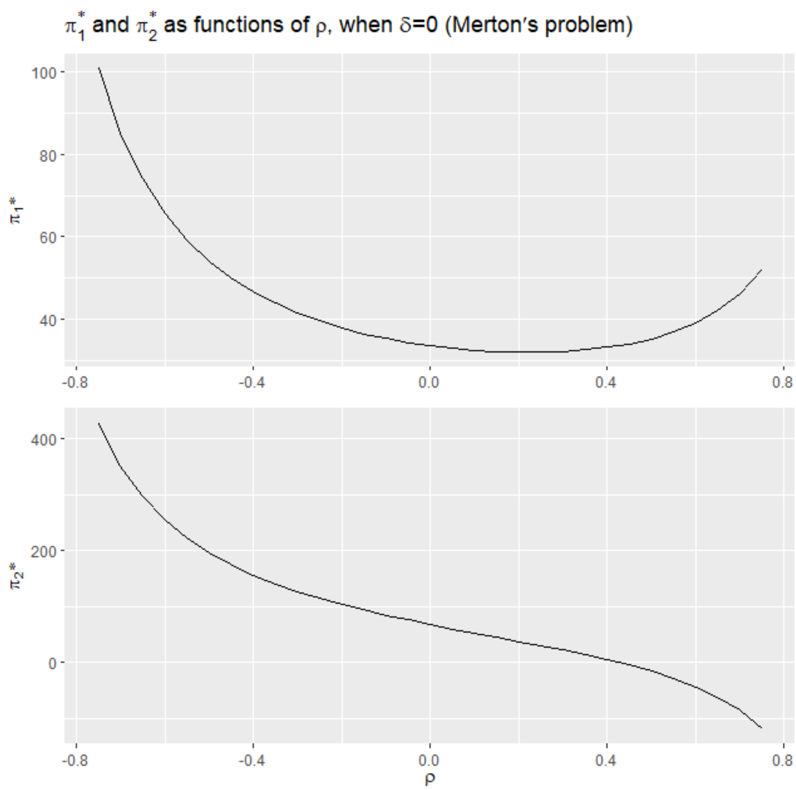


in (2.4), is for the sake of simplicity. For the generalization to a symmetric case, and moreover to a more general multi-factor model, we refer to [87]. We also assume that the risk-free interest rate is zero and we ignore the costs associated with trading. The simplicity of the present formulation enables a feasible implementation of parameter calibration and the derivation of analytical formulae for the optimal trading strategies. For the numerical example, the pair of stocks chosen from the financial sector, are positively correlated, and have a negative β in (2.1), which means that the pair tends to move together.

Based on a verification theorem in Appendix A, we deduce sufficient condition on the model parameters in the form of an inequality. It is interpreted as a bound on the optimization horizon, based on a fixed set of model parameters. Interestingly, this bound does not depend on the coefficient of risk aversion γ, which is an investor-specific quantity chosen according to the investor's risk preference. Our model generalizes the Merton's problem by including cointegration; we can see that the optimal policy reduces to that in the classical Merton portfolio optimization problem, if the cointegrating vector z were absent by setting $\delta = 0$. Another generalization results from adding correlation between the stocks, which complicates the formula and the inequality but is more realistic.

While we focus solely on the exponential utility function due its tractability, the power utility functions have been considered in for example [127] and [18]. We acknowledge that we have not addressed in this work the question of detecting two instruments whose market prices tend to evolve in tandem, although this is undoubtedly a fundamental issue. Empirically, as observed in [32], the cointegration property in US stock market are not generally persistent, meaning a good number of pairs of stocks which are tested positive for cointegration in a year might not be cointegrated in the next. See also the excellent overview by [72] for a general survey of empirical studies on pairs trading. Detecting cointegration in other asset classes, or other markets, is a critical issue for the strategies to be effective.

Figure 2.5: Effect of Correlation on π_i^*, in the absence of cointegration



Chapter 3

VIX Futures Trading under the Central Tendency Ornstein-Uhlenbeck model

While the equity market is the usual setting in which pairs trading and cointegration are considered, the futures market, composed of contracts on the same spot asset but with different maturities, is fundamentally different. Futures contracts on the same underlying asset are obviously related, and they exhibit a term structure, that is, a pattern associated with increasing maturities, which is absent in stocks, since there is no natural ending dates. This leads us to explore pairs trading using the stochastic control approach in the futures market.

In this chapter, as published in [81], we combine the ideas of pairs trading and *dynamic* portfolio, and apply them to the trading of two futures on the same underlying. We consider a two-factor mean-reverting model, where the spot price mean-reverts around its stochastic equilibrium, which is itself also mean-reverting. Since we are considering pairs trading where two futures contracts are traded, we need at least a two-factor model, since in a one-factor model, contracts of different maturities are perfectly correlated with one another. We will study a general n-factors model in the last chapter, but for now two-factor models are tractable enough to allow us to obtain closed-form expressions, as well as realistic enough to capture dynamics observed in the market. The model we work with is called the Central Tendency Ornstein-Uhlenbeck (CTOU), as studied by [91] for pricing volatility index (such as VIX) futures. See the references therein for a survey of the history of, and other approaches to, VIX modeling. There has been researches that explored cointegration in the futures market, and they focus on the existence of cointegration between spot and forward prices; [22] is a representative paper. Since VIX is based on the implied volatility of S&P options, there also has been a substantial literature that model VIX and S&P simultaneously: see [33], [12] or [99] for examples of the joint modeling of a volatility index and the underlying. Empirical examination of the VIX term structure was described in [90]. In addition, we refer to [85] and [109] for other studies of the trading of VIX futures.

We determine the optimal futures trading strategy by solving a utility maximization problem. By analyzing the associated HJB equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. Our strategies are applied to VIX futures trading, and are illustrated in a series of numerical examples. The CTOU model is also suitable for other time series where mean-reversion is present, such as interest rates or credit-related products.

A number of approaches apply machine learning and optimization algorithms to identify mean-reverting portfolios with a few assets from a larger collection of stocks [38, 128]. Typically in pairs trading the portfolio is *static* during the trading horizon. There are numerical empirical studies on the empirical performance of pairs trading and timing of trades given mean-reverting prices [71]. For portfolio optimization when the assets are mean-reverting, [18], [21], [111] or [121] are early examples. Optimal stopping/switching approach, with features like transaction costs and stop-loss exits, are incorporated by [76] and [101]. For more related studies, we refer to [77] and references therein. From an economic perspective, [124] solve for the equilibrium in a market populated with convergence traders, who are similar to traders practicing pairs trading in our context, along with noise traders and long-term investors.

On the other hand, futures has been an integral part of the global financial market and continues to grow. There are also existing studies that investigate cointegration and trading strategies in the futures market. [125] examine large-scale multiple pairs trading using a derivative-free optimization algorithm. [4] provides theoretical conditions under which the pairs trading optimization problem is market neutral. These related studies motivate us to consider a two-factor mean-reverting model for the underlying index to effectively capture the price dynamics of futures, and develop a stochastic control approach for pairs trading in the futures market.

We found that, based on parameters calibrated from VIX futures historical data, traders should take bigger positions in the long end of the futures curve. As is well-known, the volatility of longer-term VIX contracts are in general lower than the short-term contracts. Therefore, as the model calibrated to VIX data shows, it is optimal for the risk-averse trader to make larger bets on the pair trade when the contracts are far from maturity since volatility is lower in the long end as observed empirically. Again, as mentioned in the Introduction, to the extent that market participants are risk-averse and rational and therefore are allocating bigger positions as volatilities of the securities decrease, and that market data on multiple maturities can be obtained at the firm or individual levels, it might be possible to observe such allocation empirically; but this type of investigation is beyond the scope of this thesis. We also found, from the explicit formulae, that the optimal positions are inversely proportional to the degree of risk aversion, as well as to the volatility of the stochastic equilibrium process. The details of these results and other findings will be given in later sections.

3.1 Futures Price Dynamics

The two-factor mean-reverting model we consider is called the Central Tendency Ornstein Uhlenbeck (CTOU). This model has been used for pricing volatility futures (see [91]). One major feature of this model is the mean-reverting dynamics of the spot price. Specifically, the spot price tends to evolve around its stochastic equilibrium, which is also mean-reverting. The CTOU is able to capture the stylistic features of empirically observed mean reversion in volatility indices and commodity prices. Empirically, the spot price mean-reverts relatively faster than the stochastic equilibrium to its long-run mean.

Moreover, our choice is also motivated by the model's tractability. As we will see, the structure of the associated stochastic differential equations (SDEs) is very amenable to analysis and allows us to obtain closed-form solutions for the optimization problem. As noted in [91], the simplicity of this model allows for easy estimation, and it is shown that the model fits well with historical data empirically. They further note that if a jump component is added, the resulting estimates become less stable, which suggests a jump component would unnecessarily complicate model estimation and application.

The spot price is denoted by V_t. The spot's log-price mean-reverts to a stochastic equilibrium process θ_t, which in turn mean-reverts to its own *constant* equilibrium level $\bar{\theta}$. Under the risk-neutral measure $\mathbb{Q}$, the log-price process and its stochastic equilibrium follow the SDEs

$$d\log V_t = \kappa(\theta_t - \log V_t)\, dt + \sigma\, d\tilde{Z}_t^v, \tag{3.1}$$

$$d\theta_t = \bar{\kappa}(\bar{\theta} - \theta_t)\, dt + \bar{\sigma}\, d\tilde{Z}_t^\theta. \tag{3.2}$$

Here, the constants κ and $\bar{\kappa}$ represent the speeds of mean reversion for $\log V_t$ and θ_t respectively, while σ and $\bar{\sigma}$ are the respective volatilities. The model has two independent standard Brownian motions, $\tilde{Z}_t^v$ and $\tilde{Z}_t^\theta$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, with the filtration $\mathbb{F}$ generated by $(\tilde{Z}_t^v, \tilde{Z}_t^\theta)_{t\geq 0}$.

To relate the dynamics of $\log V_t$ and θ_t to the physical measure $\mathbb{P}$, we specify the market prices of risk as the constants ζ and $\bar{\zeta}$. Under the physical measure $\mathbb{P}$, we have

$$dZ_t^v = d\tilde{Z}_t^v - \zeta\, dt, \tag{3.3}$$

$$dZ_t^\theta = d\tilde{Z}_t^\theta - \bar{\zeta}\, dt. \tag{3.4}$$

Note that the conditions under which the $\mathbb{P}$ measure is identical to $\mathbb{Q}$ is: $\zeta = \bar{\zeta} = 0$.

Thus, we can recover the dynamics of V_t and θ_t under the physical measure $\mathbb{P}$ as

$$d\log V_t = \kappa\left(\theta_t + \frac{\sigma\zeta}{\kappa} - \log V_t\right) dt + \sigma\, dZ_t^v, \tag{3.5}$$

$$d\theta_t = \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta_t\right) dt + \bar{\sigma}\, dZ_t^\theta. \tag{3.6}$$

Remark 1. *The CTOU model is a variation of the concatenated SQR (CSQR)*

$$dV_t = \kappa(\theta_t - V_t)\, dt + \sigma\sqrt{V_t}\, d\tilde{Z}_t^v,$$
$$d\theta_t = \bar{\kappa}(\bar{\theta} - \theta_t)\, dt + \bar{\sigma}\sqrt{\theta_t}\, d\tilde{Z}_t^\theta,$$

which has been studied by [16], among others, and is presented in [91] as well.

We now consider futures contracts of different maturities written on the spot V. For the futures contract with maturity T_i, $i = 1, \ldots, n$, we define the price at time $t \in [0, T_i]$ by

$$F^{(i)}(t, V_t, \theta_t) = \tilde{\mathbb{E}}[V_{T_i} \mid V_t, \theta_t].$$

We will continue to work with log prices for both the spot and its futures prices. Hence, we define

$$v_t \equiv \log V_t, \tag{3.7}$$

$$f^{(i)}(t, v_t, \theta_t) \equiv \log F^{(i)}(t, V_t, \theta_t). \tag{3.8}$$

From Appendix C of [91], we obtain the explicit log futures price:

$$\begin{aligned} f_t^{(i)} &\equiv f^{(i)}(t, v_t, \theta_t) \\ &= \bar{\theta} + D(T_i - t)(\theta_t - \bar{\theta}) + e^{-\kappa(T_i - t)}(\log V_t - \theta_t) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa(T_i - t)}) \\ &+ \frac{\bar{\sigma}^2}{2}\left(\frac{\kappa}{\kappa - \bar{\kappa}}\right)^2 \left(\frac{1 - e^{-2\bar{\kappa}(T_i - t)}}{2\bar{\kappa}} + \frac{1 - e^{-2\kappa(T_i - t)}}{2\kappa} - 2\frac{1 - e^{-(\kappa + \bar{\kappa})(T_i - t)}}{\kappa + \bar{\kappa}}\right), \end{aligned}$$

where

$$D(\tau) = \frac{\kappa}{\kappa - \bar{\kappa}} e^{-\bar{\kappa}\tau} - \frac{\bar{\kappa}}{\kappa - \bar{\kappa}} e^{-\kappa\tau}.$$

Since our objective is to dynamically trade a futures portfolio under the physical measure $\mathbb{P}$, we derive the SDE for $f_t^{(i)}$ using Ito's Lemma (see Appendix 3.5.1).

$$df_t^{(i)} = m_i(t)\, dt + \sigma e^{-\kappa(T_i - t)}\, dZ_t^v + \frac{\bar{\sigma}\kappa(e^{-\bar{\kappa}(T_i - t)} - e^{-\kappa(T_i - t)})}{\kappa - \bar{\kappa}}\, dZ_t^\theta,$$

where the drift is a deterministic function time, given by

$$\begin{aligned} m_i(t) = \frac{e^{-\bar{\kappa}(T_i - t)}}{\kappa - \bar{\kappa}}\kappa\bar{\sigma}\xi + \frac{e^{-\kappa(T_i - t)}}{(\kappa - \bar{\kappa})^2}(\kappa\bar{\sigma}\bar{\kappa}\xi - \kappa^2\bar{\sigma}\xi + \kappa^2\zeta\sigma - 2\kappa\zeta\bar{\kappa}\sigma + \zeta\bar{\kappa}^2\sigma) \\ + \frac{\kappa^2\bar{\sigma}^2}{2(\kappa - \bar{\kappa})^2}\left(2e^{-(\kappa + \bar{\kappa})(T_i - t)} - e^{-2\bar{\kappa}(T_i - t)}\right) + \frac{e^{-2\kappa(T_i - t)}}{2(\kappa - \bar{\kappa})^2}(2\kappa\bar{\kappa}\sigma^2 - \kappa^2\bar{\sigma}^2 - \kappa^2\sigma^2 - \bar{\kappa}^2\sigma^2). \end{aligned} \tag{3.9}$$

In turn, we can write the futures price dynamics under $\mathbb{P}$ compactly as

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)\, dt + \sigma_{vi}(t)\, dZ_t^v + \sigma_{\theta i}(t)\, dZ_t^\theta, \tag{3.10}$$

where all three time-deterministic coefficients are defined by

$$\sigma_{vi}(t) \equiv \sigma e^{-\kappa(T_i - t)}, \tag{3.11}$$

$$\sigma_{\theta i}(t) \equiv \frac{\bar{\sigma}\kappa(e^{-\bar{\kappa}(T_i - t)} - e^{-\kappa(T_i - t)})}{\kappa - \bar{\kappa}}, \tag{3.12}$$

and

$$\mu_i(t) \equiv \frac{e^{-\bar{\kappa}(T_i - t)}\,\kappa\,\bar{\sigma}\,\bar{\zeta} - e^{-\kappa\,(T_i - t)}\,\left(\kappa\,\bar{\sigma}\,\bar{\zeta} - \kappa\,\zeta\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)}{\kappa - \bar{\kappa}} \tag{3.13}$$

$$= m_i(t) + \frac{\sigma_i(t)^2}{2}, \tag{3.14}$$

where

$$\sigma_i(t)^2 \equiv \sigma_{vi}(t)^2 + \sigma_{\theta i}(t)^2. \tag{3.15}$$

The instantaneous correlation between the two futures is defined by

$$\rho_{12}(t) \equiv \frac{\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta 1}(t)\sigma_{\theta 2}(t)}{\sigma_1(t)\sigma_2(t)}, \tag{3.16}$$

which is a deterministic function of time only, independent of the state variables.

3.2 Utility Maximization Problem

Having established the dynamics of the futures prices in the previous section, we now consider the utility maximization problem involving a pair of futures. Let T_1 and T_2 be the maturities of the two futures in our portfolio. The optimization horizon will be denoted by T. Since the futures cannot be traded past expiry, we require $T_i \geq T$ for $i = 1, 2$. Using the futures price dynamics in (3.10), we write down the SDE for the portfolio wealth process as

$$dW_t = \pi_1(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(1)} + \pi_2(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(2)}, \tag{3.17}$$

where $\pi_i(t, F_t^{(1)}, F_t^{(2)})$, $i = 1, 2$, denote the number of contracts, and positive/negative values mean a long/short position, respectively. For brevity, we may write $\pi_i \equiv \pi_i(t, F_t^{(1)}, F_t^{(2)})$.

Re-writing in matrix form in terms of the two fundamental sources of randomness (Z_t^v, Z_t^θ), we get

$$\begin{aligned} \begin{bmatrix} dW_t \\ dF_t^{(1)} \\ dF_t^{(2)} \end{bmatrix} &= \begin{bmatrix} \pi_1\mu_1(t)F_t^{(1)} + \pi_2\mu_2(t)F_t^{(2)} \\ \mu_1(t)F_t^{(1)} \\ \mu_2(t)F_t^{(2)} \end{bmatrix} dt \\ &+ \begin{bmatrix} \pi_1\sigma_{v1}(t)F_t^{(1)} + \pi_2\sigma_{v2}(t)F_t^{(2)} & \pi_1\sigma_{\theta1}(t)F_t^{(1)} + \pi_2\sigma_{\theta2}(t)F_t^{(2)} \\ \sigma_{v1}(t)F_t^{(1)} & \sigma_{\theta1}(t)F_t^{(1)} \\ \sigma_{v2}(t)F_t^{(2)} & \sigma_{\theta2}(t)F_t^{(2)} \end{bmatrix} \begin{bmatrix} dZ_t^v \\ dZ_t^\theta \end{bmatrix}. \end{aligned} \tag{3.18}$$

A pair of controls (π_1, π_2) is said to be admissible if (π_1, π_2) are real-valued, progressively measurable, and are such that the system of SDE (3.18) defines a unique solution $(W_t, F_t^{(1)}, F_t^{(2)})$ for every time $t \in [0, T]$ and $(\pi_1, \pi_2, F^{(1)}, F^{(2)})$ satisfy the admissibility condition

$$\mathbb{E}\left(\int_t^T [\pi_1(s, F_s^{(1)}, F_s^{(2)})F_s^{(1)}]^2 + [\pi_2(s, F_s^{(1)}, F_s^{(2)})F_s^{(2)}]^2 ds\right) < \infty.$$

We denote by $\mathcal{A}_t$ the set of admissible controls with an initial time of investment t. Next, we define the value function $V(t, w, F_1, F_2)$ of the following optimization problem: the investor seeks an admissible strategy (π_1, π_2) that maximizes the utility from wealth at time T, that is,

$$V(t, w, F_1, F_2) = \sup_{(\pi_1,\pi_2)\in\mathcal{A}_t} \mathbb{E}\left(U(W_T) \,|\, W_t = w, F_t^{(1)} = F_1, F_t^{(2)} = F_2\right). \tag{3.19}$$

Here we only treat the case of the exponential utility function $U(w) = -e^{-\gamma w}$ where γ denotes the constant risk aversion coefficient. Following the standard verification approach to dynamic programming [58, 105, 96], we assume the existence of a sufficiently smooth candidate solution $u(t, w, F_1, F_2)$, which will later be shown to be equal to the value function V in (3.19).

3.2.1 HJB Equation and Closed-Form Solution

To facilitate presentation, we define the partial derivatives by

$$u_t = \frac{\partial u}{\partial t}, \quad u_w = \frac{\partial u}{\partial w}, \quad u_{ww} = \frac{\partial^2 u}{\partial w^2},$$

$$u_1 = \frac{\partial u}{\partial F_1}, \quad u_{11} = \frac{\partial^2 u}{\partial F_1^2}, \quad u_2 = \frac{\partial u}{\partial F_2}, \quad u_{22} = \frac{\partial^2 u}{\partial F_2^2},$$

$$u_{w1} = \frac{\partial^2 u}{\partial w \partial F_1}, \quad u_{w2} = \frac{\partial^2 u}{\partial w \partial F_2}, \quad u_{12} = \frac{\partial^2 u}{\partial F_1 \partial F_2}.$$

We determine the value function $u(t, w, F_1, F_2)$ by solving the HJB equation

$$\begin{aligned}
u_t + \sup_{\pi_1,\pi_2} \Big(& \pi_1\mu_1(t)F_1u_w + \pi_2\mu_2(t)F_2u_w \\
& + (\pi_1\sigma_1(t)^2F_1^2 + \pi_2(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{w1} \\
& + (\pi_2\sigma_2(t)^2F_2^2 + \pi_1(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{w2} \\
& + \frac{1}{2}(\pi_1^2\sigma_1(t)^2F_1^2 + \pi_2^2\sigma_2(t)^2F_2^2)u_{ww} \\
& + (\pi_1\pi_2(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{ww}\Big) \\
& + \frac{\sigma_1(t)^2}{2}F_1^2u_{11} + \frac{\sigma_2(t)^2}{2}F_2^2u_{22} + \mu_1(t)F_1u_1 + \mu_2(t)F_2u_2 \\
& + (\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2u_{12} = 0,
\end{aligned} \tag{3.20}$$

subject to the terminal condition

$$u(T, w, F_1, F_2) = -e^{-\gamma w}.$$

We introduce the linear operator $\mathcal{L}^{(\pi_1,\pi_2)}$ that will be needed in order to apply the verification theorem in Appendix A

$$\begin{aligned}
\mathcal{L}^{(\pi_1,\pi_2)}u = & \pi_1\mu_1(t)F_1u_w + \pi_2\mu_2(t)F_2u_w \\
& + (\pi_1\sigma_1(t)^2F_1^2 + \pi_2(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{w1} \\
& + (\pi_2\sigma_2(t)^2F_2^2 + \pi_1(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{w2} \\
& + \frac{1}{2}(\pi_1^2\sigma_1(t)^2F_1^2 + \pi_2^2\sigma_2(t)^2F_2^2)u_{ww} \\
& + (\pi_1\pi_2(\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2)u_{ww} \\
& + \frac{\sigma_1(t)^2}{2}F_1^2u_{11} + \frac{\sigma_2(t)^2}{2}F_2^2u_{22} + \mu_1(t)F_1u_1 + \mu_2(t)F_2u_2 \\
& + (\sigma_{v1}(t)\sigma_{v2}(t) + \sigma_{\theta1}(t)\sigma_{\theta2}(t))F_1F_2u_{12}.
\end{aligned}$$

The HJB equation can now be written compactly as

$$u_t + \sup_{\pi_1,\pi_2} \mathcal{L}^{(\pi_1,\pi_2)}u = 0.$$

Next, we apply the transformation

$$u(t, w, F_1, F_2) = -e^{-\gamma w}G(t, f_1, f_2),$$

with $f_1 = \log F_1$ and $f_2 = \log F_2$. By direct substitution, we obtain the PDE for G:

$$\begin{aligned}
-e^{-\gamma w}G_t + \sup_{\pi_1,\pi_2} \big[& (\pi_1\mu_1F_1 + \pi_2\mu_2F_2)\gamma e^{-\gamma w}G \\
& + (\pi_1\sigma_1^2F_1^2 + \pi_2\rho_{12}\sigma_1\sigma_2F_1F_2)\gamma e^{-\gamma w}G_1/F_1 + (\pi_2\sigma_2^2F_2^2 + \pi_1\rho_{12}\sigma_1\sigma_2F_1F_2)\gamma e^{-\gamma w}G_2/F_2 \\
& + \frac{1}{2}(\pi_1^2\sigma_1^2F_1^2 + \pi_2^2\sigma_2^2F_2^2 + \rho_{12}\pi_1\pi_2\sigma_1\sigma_2F_1F_2)(-\gamma^2e^{-\gamma w}G) \big] \\
& + \mu_1F_1(-e^{-\gamma w}G_1/F_1) + \mu_2F_2(-e^{-\gamma w}G_2/F_2) \\
& + \frac{\sigma_1^2}{2}F_1^2(e^{-\gamma w}(G_1 - G_{11}))/F_1^2 + \frac{\sigma_2^2}{2}F_2^2(e^{-\gamma w}(G_2 - G_{22}))/F_2^2 \\
& + \rho_{12}\sigma_1\sigma_2F_1F_2(-e^{-\gamma w}G_{12})/F_1F_2 = 0,
\end{aligned} \tag{3.21}$$

where we have defined the partial derivatives

$$G_t = \frac{\partial G}{\partial t}, \quad G_1 = \frac{\partial G}{\partial f_1}, \quad G_2 = \frac{\partial G}{\partial f_2},$$

$$G_{11} = \frac{\partial^2 G}{\partial f_1^2}, \quad G_{22} = \frac{\partial^2 G}{\partial f_2^2}, \quad G_{12} = \frac{\partial^2 G}{\partial f_1 \partial f_2},$$

and suppressed the dependence on t, in μ_i, σ_i, σ_{vi}, $\sigma_{\theta i}$ and ρ_{12}, in order to simplify notation.

Canceling $e^{-\gamma w}$ and rearranging, we get

$$\begin{aligned}
-G_t + \sup_{\pi_1,\pi_2} \Bigg[\quad & (\pi_1\mu_1 F_1 + \pi_2\mu_2 F_2)\gamma G \\
& + (\pi_1\sigma_1^2 F_1\gamma + \pi_2\gamma\rho_{12}\sigma_1\sigma_2 F_2)G_1 + (\pi_2\sigma_2^2 F_2\gamma + \pi_1\gamma\rho_{12}\sigma_1\sigma_2 F_1)G_2 \\
& - \frac{\sigma_1^2}{2}\pi_1^2 F_1^2\gamma^2 G - \frac{\sigma_2^2}{2}\pi_2^2 F_2^2\gamma^2 G - \gamma^2\pi_1\pi_2\rho_{12}\sigma_1\sigma_2 F_1 F_2 G \quad \Bigg] \\
& - \frac{\sigma_1^2}{2}(G_{11} - G_1) - \frac{\sigma_2^2}{2}(G_{22} - G_2) - \rho_{12}\sigma_1\sigma_2 G_{12} - \mu_1 G_1 - \mu_2 G_2 = 0,
\end{aligned} \tag{3.22}$$

with the terminal condition

$$G(T, f_1, f_2) = 1.$$

Performing the optimization in (3.22), we obtain the optimal controls

$$\pi_1^*(t, F_1, F_2) = \frac{\mu_1}{\gamma(1-\rho_{12}^2)\sigma_1^2 F_1} + \frac{G_1}{\gamma G F_1} - \rho_{12}\frac{\mu_2}{\gamma(1-\rho_{12}^2)\sigma_1\sigma_2 F_1}, \tag{3.23}$$

$$\pi_2^*(t, F_1, F_2) = \frac{\mu_2}{\gamma(1-\rho_{12}^2)\sigma_2^2 F_2} + \frac{G_2}{\gamma G F_2} - \rho_{12}\frac{\mu_1}{\gamma(1-\rho_{12}^2)\sigma_1\sigma_2 F_2}. \tag{3.24}$$

Then we substitute the optimal controls as in (3.23) and (3.24) to arrive at a *nonlinear* PDE for G:

$$\begin{aligned}
G_t = & \left(\frac{1}{2}\frac{\mu_1^2}{(1-\rho_{12}^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho_{12}^2)\sigma_2^2} - \frac{\rho_{12}\mu_1\mu_2}{(1-\rho_{12}^2)\sigma_1\sigma_2}\right) G \\
& + \frac{1}{2G}\left(G_1^2\sigma_1^2 + 2G_1 G_2\rho_{12}\sigma_1\sigma_2 + G_2^2\sigma_2^2\right) \\
& - \frac{1}{2}((G_{11} - G_1)\sigma_1^2 + (G_{22} - G_2)\sigma_2^2 + 2G_{12}\rho\sigma_1\sigma_2).
\end{aligned} \tag{3.25}$$

To solve (3.25), we apply another transformation

$$G(t, f_1, f_2) = e^{-\Phi(t, f_1, f_2)} \tag{3.26}$$

to (3.25) to obtain a *linear* PDE for Φ:

$$\begin{aligned}
0 = \Phi_t + & \left(\frac{1}{2}\frac{\mu_1^2}{(1-\rho_{12}^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho_{12}^2)\sigma_2^2} - \frac{\rho_{12}\mu_1\mu_2}{(1-\rho_{12}^2)\sigma_1\sigma_2}\right) \\
& + \frac{\sigma_1^2}{2}(\Phi_{11} - \Phi_1) + \frac{\sigma_2^2}{2}(\Phi_{22} - \Phi_2) + \rho_{12}\sigma_1\sigma_2\Phi_{12},
\end{aligned} \tag{3.27}$$

subject to $\Phi(T, f_1, f_2)$=0.

We can solve this linear PDE of Φ by using the ansatz

$$\Phi(t, f_1, f_2) = a_{11}(t)f_1^2 + a_1(t)f_1 + a_{22}(t)f_2^2 + a_2(t)f_2 + a_{12}(t)f_1 f_2 + a(t)$$

to deduce that

$$a_{11}'(t) = a_{22}'(t) = a_{12}'(t) = 0, \; a_{11}(t) = a_{22}(t) = a_{12}(t) = 0,$$
$$a_1'(t) = a_2'(t) = 0, \; a_1(t) = a_2(t) = 0.$$

From this, we deduce that Φ is a function of t only, independent of f_1 and f_2, and satisfies the first-order differential equation

$$\frac{d\Phi}{dt} = -\frac{\mu_1(t)^2\sigma_2(t)^2 + \mu_2(t)^2\sigma_1(t)^2 - 2\rho_{12}(t)\mu_1(t)\mu_2(t)\sigma_1(t)\sigma_2(t)}{2(1-\rho_{12}(t)^2)\sigma_1(t)^2\sigma_2(t)^2}.$$

Solving this and applying (3.13), (3.15), and (3.16), we obtain a closed-form expression for Φ. Precisely,

$$\Phi(t) = \frac{(T-t)\left(\zeta^2 + \bar{\zeta}^2\right)}{2}. \tag{3.28}$$

Unraveling the transformations, we write the candidate solution as

$$u(t, w, F_1, F_2) = -e^{-\gamma w - \Phi(t)} \tag{3.29}$$

which will later be verified to be the value function.

Very interestingly, the value function depends on only two model parameters, namely, the market prices of risk ζ and $\bar{\zeta}$, along with the optimization horizon T. Moreover, the value function does not depend on the two futures current prices (F_1, F_2). The simplicity of the value function is unexpected, especially since there are two stochastic factors and two futures in the trading problem. This is a very useful result that shows clearly the dependence of the value function on the two model parameters (see Fig. 3.2 below). Nevertheless, it does not mean that the corresponding trading strategies are trivial. In fact, the strategies depend not only on other model parameters but also the futures prices, as we will discuss next.

3.2.2 Optimal Wealth Process

By applying (3.26) and (3.28) to (3.23) and (3.24), we obtain the optimal trading strategies (i.e. number of futures contracts)

$$\pi_1^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_1(t)F_1}\left(\frac{\mu_1(t)}{\sigma_1(t)} - \rho_{12}(t)\frac{\mu_2(t)}{\sigma_2(t)}\right), \tag{3.30}$$

$$\pi_2^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_2(t)F_2}\left(\frac{\mu_2(t)}{\sigma_2(t)} - \rho_{12}(t)\frac{\mu_1(t)}{\sigma_1(t)}\right). \tag{3.31}$$

We recall (3.13), (3.15), and (3.16), and express the optimal strategies explicitly in terms of model parameters. Precisely,

$$\pi_1^*(t, F_1, F_2) = \frac{-e^{-\bar{\kappa}(T_2-t)}\kappa\zeta\bar{\sigma} + e^{-\kappa(T_2-t)}\left(\kappa\zeta\bar{\sigma} + \kappa\bar{\zeta}\sigma - \bar{\kappa}\bar{\zeta}\sigma\right)}{e^{t(\kappa+\bar{\kappa})}\left(e^{-\bar{\kappa}T_1-\kappa T_2} - e^{-\kappa T_1 - \bar{\kappa}T_2}\right)\kappa\gamma\bar{\sigma}\sigma F_1}, \tag{3.32}$$

$$\pi_2^*(t, F_1, F_2) = \frac{e^{-\bar{\kappa}(T_1-t)}\kappa\zeta\bar{\sigma} - e^{-\kappa(T_1-t)}\left(\kappa\zeta\bar{\sigma} + \kappa\bar{\zeta}\sigma - \bar{\kappa}\bar{\zeta}\sigma\right)}{e^{t(\kappa+\bar{\kappa})}\left(e^{-\bar{\kappa}T_1-\kappa T_2} - e^{-\kappa T_1 - \bar{\kappa}T_2}\right)\kappa\gamma\bar{\sigma}\sigma F_2}. \tag{3.33}$$

Note that the optimal controls are functions of time and the futures prices, but not functions of the spot price V_t and its equilibrium level θ_t. The strategies are inversely proportional to γ, as is expected. For each $i \in \{1, 2\}$, the optimal strategy π_i^* depends only on the corresponding futures price F_i, but not the other futures price.

If we substitute the optimal controls π_1^* and π_2^* into the wealth process (3.17), we have

$$\begin{aligned} dW(t) &= \pi_1^*\, dF_t^{(1)} + \pi_2^*\, dF_t^{(2)} \\ &= \pi_1^* F_t^{(1)}\mu_1(t)\, dt + \pi_2^* F_t^{(2)}\mu_2(t)\, dt + \\ &+ \left(\pi_1^*\sigma_{v1}(t)F_t^{(1)} + \pi_2^*\sigma_{v2}(t)F_t^{(2)}\right) dZ_t^v + \left(\pi_1^*\sigma_{\theta 1}(t)F_t^{(1)} + \pi_2^*\sigma_{\theta 2}(t)F_t^{(2)}\right) dZ_t^\theta \\ &\equiv \mu_W\, dt + \sigma_W\, dZ_t^W, \end{aligned} \tag{3.34}$$

where we have defined

$$\mu_W = \pi_1^* F_t^{(1)} \mu_1(t) + \pi_2^* F_t^{(2)} \mu_2(t),$$
$$\sigma_W^2 = \left(\pi_1^* \sigma_{v1}(t) F_t^{(1)} + \pi_2^* \sigma_{v2}(t) F_t^{(2)}\right)^2 + \left(\pi_1^* \sigma_{\theta 1}(t) F_t^{(1)} + \pi_2^* \sigma_{\theta 2}(t) F_t^{(2)}\right)^2.$$

Direct computation simplifies μ_W and σ_W to

$$\mu_W = \frac{\zeta^2 + \bar{\zeta}^2}{\gamma}, \qquad \text{and} \qquad \sigma_W^2 = \frac{\zeta^2 + \bar{\zeta}^2}{\gamma^2} = \frac{\mu_W}{\gamma}. \tag{3.35}$$

Substituting (3.35) into (3.34) implies that the optimal wealth process is in fact a Brownian motion with constant drift μ_W and volatility σ_W parameters. As a result, the optimal wealth depends on the market prices of risk, ζ and $\bar{\zeta}$ (see (3.3)-(3.4)), as well as the investor's risk aversion parameter γ.

3.2.3 Verification Theorem

Based on the verification theorem in Appendix A, Theorem 7, in order to show that the candidate solution is the value function, it suffices to prove the uniform integrability of the family of random variables $\{u(\tau, W_\tau^*)\}$ for any stopping time $\tau \in [0, T]$.

Theorem 3. *The candidate solution found in* (3.29) *is equal to the value function* (3.19)*: namely,*

$$V(t, w) = u(t, w) = -\exp(-\gamma w - \Phi(t))$$

on $(t, w) \in [0, T] \times \mathbb{R}$*, where* $\Phi(t)$ *is given by* (3.28)*, and the optimal control pair is given by* (3.32) *and* (3.33).

Proof. Let $\epsilon > 0$ be an arbitrary positive number. From the verification theorem in Appendix A, it suffices to show that $\mathbb{E}(|u(\tau, W_\tau^*)|^{1+\epsilon}) < \infty$, uniformly with respect to any stopping time τ with $0 \le \tau \le T$. First, after applying the Cauchy-Schwarz inequality, we find that

$$\mathbb{E}(|u(\tau, W_\tau^*)|^{1+\epsilon}) \le \mathbb{E}(\exp(-2(1+\epsilon)\gamma W_\tau^*))^{1/2} \times \tag{3.36}$$
$$\mathbb{E}(\exp(-2(1+\epsilon)\Phi(\tau)))^{1/2}. \tag{3.37}$$

For the first expectation (3.36), clearly

$$\begin{aligned}
&\mathbb{E}[\exp(-2(1+\epsilon)\gamma W_\tau^*)] \\
=&\mathbb{E}\left[\exp\left(-2(1+\epsilon)\left(\int_0^\tau \mu_W ds + \int_0^\tau \sigma_W dZ_s^W\right)\right)\right] \\
=&\mathbb{E}\left[\exp\left(-2(1+\epsilon)\left(\mu_W \tau + \sigma_W Z_\tau^W\right)\right)\right] \\
\le& c\mathbb{E}\left[\exp\left(-2(1+\epsilon)\sigma_W Z_\tau^W\right)\right]
\end{aligned} \tag{3.38}$$

for some constant c, since μ_W is a constant and $\tau \le T$.

Now since Z_t^W is a martingale, we have that for any constant k, $\exp(kZ_t^W)$ is a submartingale by Jensen's inequality, since

$$\mathbb{E}_0(\exp(kZ_t^W)) \ge \exp(k\mathbb{E}_0(Z_t^W)) = \exp(k\mathbb{E}_0(Z_0^W)) = 1.$$

Moreover, $\exp(kZ_t^W)$ is positive. Therefore we can use Doob's martingale inequality [103]: for $\xi > 0$,

$$\begin{aligned}
\mathbb{E}[\exp(kZ_\tau^W)] &= \left\| \exp\left(\frac{k}{1+\xi} Z_\tau^W\right) \right\|_{1+\xi}^{1+\xi} \\
&\leq \left\| \sup_{0\leq t\leq T} \exp\left(\frac{k}{1+\xi} Z_t^W\right) \right\|_{1+\xi}^{1+\xi} \\
&\leq \left(1+\frac{1}{\xi}\right)^{1+\xi} \sup_{0\leq t\leq T} \mathbb{E}[\exp(kZ_t^W)] \\
&\leq c
\end{aligned}$$

where c is another positive constant, independent of the stopping time τ.

For the second expectation (3.37), recalling that

$$\Phi(t) = \frac{(T-t)\left(\zeta^2 + \bar{\zeta}^2\right)}{2},$$

we can clearly see that $\mathbb{E}(\exp(-2(1+\epsilon)\Phi(\tau)))$ is bounded as well. Hence we proved that $\{u(\tau, W_\tau^*)\}_\tau$ is uniformly integrable for any stopping time $\tau \in [0,T]$. □

3.3 Numerical Implementation

We will now further examine our results with numerical examples with simulated and empirical data. For our examples, we will set γ to be 1, and use the estimated parameters from the "full sample" in Table 4 of [91], which are displayed here in Table 3.1.

κ	$\bar{\kappa}$	$\bar{\theta}$	σ	$\bar{\sigma}$	ζ	$\bar{\zeta}$
5.827	0.300	3.019	1.037	0.446	−0.010	2.242

Table 3.1: CTOU model parameters

According to Section 3 in [91], the parameters are obtained from maximization of the so-called pseudo likelihood in state-space modeling, which is described in more details in [118]. The parameters so obtained were further tested by comparing to VIX option prices and compared on the basis of Root Mean Square Error (RMSE). As noted, one of the advantages of the CTOU model is its tractability, and in the context of estimation, the continuous time SDE for $\log V_t$ and θ_t can be easily written as a Gaussian VAR(1), for which the transition density is known in closed forms.

In Figure 3.1, we show the dependence of the optimal trading strategies, π_1^* and π_2^*, on the volatility $\bar{\sigma}$ of the stochastic equilibrium θ_t in the CTOU model. Observe that π_1^* is positive and increasing concave while π_1^* negative and decreasing convex. With the parameters given in Table 3.1, we are short the T_1-futures $F^{(1)}$ and long the T_2-futures $F^{(2)}$. When we rearrange the formulae (3.32) and (3.33) for π_1^* and π_2^*, and collect terms involving $\bar{\sigma}$, we see that for both $i = 1, 2$, the optimal strategies are of the form $A_i + B_i/\bar{\sigma}$, which means that the absolute value of the each strategy π_i^* decreases as $\bar{\sigma}$ gets large, with other variables held constant. The practical consequence is that the number of contracts held, on both the long and short sides, are decreasing as the volatility of the stochastic equilibrium increases. This is in line with a risk-averse trader's intuition, who would prefer less exposures on both legs of the paired-trade, if the volatility of the stochastic equilibrium is high. To verify this empirically, we would have to first identify the subset of traders in the whole universe of market participants who first of all are risk-averse, who practice pairs trading of futures contracts, and who model the spot price as a process with a stochastic equilibrium such as CTOU; moreover, we would also have to gain access to their trading records. This type of data might be available in self-reporting surveys; however it is beyond the scope of this thesis to investigate actual trading empirically.

Figure 3.2 illustrates how the optimal trading strategies, π_1^* and π_2^*, vary with respect to the time-to-maturity. We see the number of contracts to buy, or to sell, are both increasing as maturity increases, with π_1^* becoming more negative and π_2^* more positive. From the trader's perspective, this corresponds to taking bigger positions in the long end of the futures curve. As is well-known [2], the volatility of longer-term VIX contracts are in general lower than the short-term contracts. Therefore, under the CTOU model with parameters as calibrated to the VIX futures price history, it is optimal for the risk-averse trader to make larger bets on the pair trade when the contracts are far from maturity since volatility is lower in the long end as observed empirically.

In Figure 3.3 we compare the optimal trading strategies, π_1^* and π_2^* (for two futures) to the optimal strategy $\tilde{\pi}_i^*$ for trading a single futures. The case of dynamically trading a single futures is discussed in an appendix later. The optimal strategy is explicitly given in (3.30) and (3.31). As in Figure 3.2, we plot the strategies as functions of T_i, using same set of parameters. When trading a single contract, either with maturity T_1 or T_2, the corresponding optimal strategy, $\tilde{\pi}_1^*$ and $\tilde{\pi}_2^*$, is positive. This is in contrast to the two-contract case where the optimal strategies, π_1^* and π_2^*, are of opposite signs. This is intuitive because when two contracts are available, along with the fact that the two futures are based on the same sources of randomness, risk aversion drives the investor to reduce risks by taking long and short positions simultaneously.

A related question is: when only one contract is traded, does the investor favor the longer or shorter maturity? As we can see, $\tilde{\pi}_2^*$ is greater than $\tilde{\pi}_1^*$. This means that, given only one contract is available, the trader tends to take a larger position in the contract if it is further away from maturity. When we compare the long and short positions in the two-contract case to the single-contract case in terms of position size, we can see the optimal long-short strategy requires taking bigger positions in both contracts than either position in the single-contract case.

Using historical VIX futures data, we consider two contracts, one with maturity January 2011 and the other with maturity February 2011. We show the empirical optimal positions over the period October 2010 to December 2010. This period is chosen to correspond to the post-calibration period of the full sample in Table 4 of [91]. Over this post-crisis period, the market was relatively calm compared to the market during the crisis, with the VIX index hovering around 20. Applying our explicit formulae for the strategies, we compute π_1^*, π_2^*, and $\pi_1^* + \pi_2^*$ based on the daily settlement prices of these contracts as well as the parameters in Table 3.1. As shown in Figure 3.4, the optimal strategy π_1^* is negative throughout this period, corresponding to a short position in the front-month contract, and the opposite holds for π_2^*. However, the absolute value of position of π_2^* is larger, leading to a net positive position.

We now turn our attention to the value functions. To distinguish between the single-contract and two-contract cases, we let $\tilde{u}^{(i)}$ denote the value function in the single-contract case with the superscript (i) indicating the maturity T_i of that single contract in the portfolio. In Figure 3.5 we plot $\tilde{u}^{(1)}$, $\tilde{u}^{(2)}$ and u as functions of t, and set $w = 0$. We observe that the maximized expected utility from trading two contracts simultaneously is greater than the maximized expected utility derived from trading only a single contract regardless of the choice of maturity. In fact, the value function u is larger than the sum of the two value functions $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$. This makes sense since the single-contract case can be viewed as two-contracts case but with one strategy constrained at zero. Effectively, the single-contract case is restricting the admissible set $\mathcal{A}_t$, thus reducing the maximum expected utility. Our result confirms the intuition that more choices of trading instruments are preferable to fewer.

Next, we consider the *certainty equivalent* for the trading opportunity in the two futures with wealth w at time t. Recall that the value function is in exponential form

$$u(t, w, F_1, F_2) = -e^{-\gamma w - \Phi(t)}.$$

We define the certainty equivalent by taking the inverse of the exponential utility function. Precisely, we have

$$C(t, w) = w + \frac{\Phi(t)}{\gamma}.$$

As we can see, the certainty equivalent is the sum of the investor's wealth w and the positive value $\Phi(t)/\gamma$. The latter is inversely proportional to the risk aversion parameter γ. Like the value function, the certainty equivalent does not depend on the current futures prices (F_1, F_2) but it does depend on the model parameters that drive the futures prices.

We now evaluate the behavior of C at time $t = 0$ and with zero initial wealth $W_0 = 0$. In other words, we will examine the following quantity:

$$C_0 = \frac{\Phi(0)}{\gamma},$$

and its sensitivity with respect to as we have plotted in Figure 3.6. In Figure 3.6 we plot the certainty equivalent against the price of risk. From (3.28) it is clear that C_0 is quadratic in ζ and $\bar{\zeta}$ under the CTOU model, and tends to infinity as the prices of risk increase.

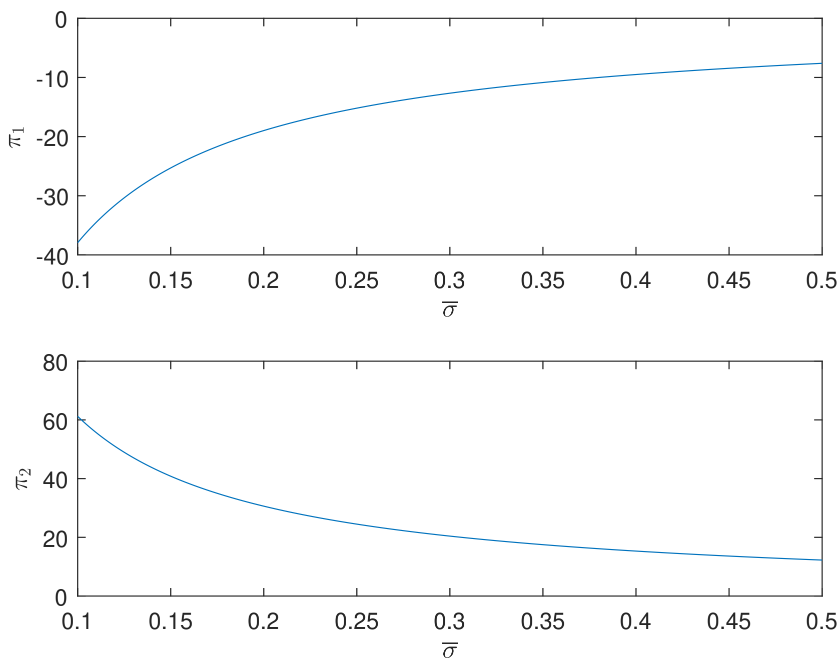


Figure 3.1: Optimal controls π_1^* and π_2^* as a function of $\bar{\sigma}$ under the CTOU model, with $\kappa = 5.827, \bar{\kappa} = 0.300, \bar{\theta} = 3.019, \sigma = 1.037, \zeta = -0.010$ and $\bar{\zeta} = 2.242$ as displayed in Table 3.1, at $T_1 = 30/365$ and $T_2 = 60/365$.

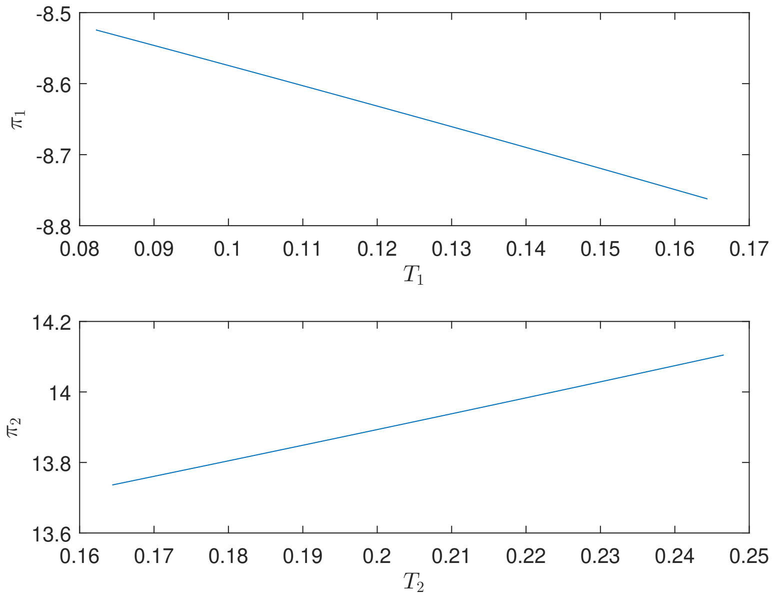


Figure 3.2: Optimal controls π_1^* and π_2^* as a function of T_1 and T_2 respectively, under the CTOU model, with parameters as displayed in Table 3.1, and T_1 ranges from [30/365,60/365], and T_2 ranges from [60/365,90/365].

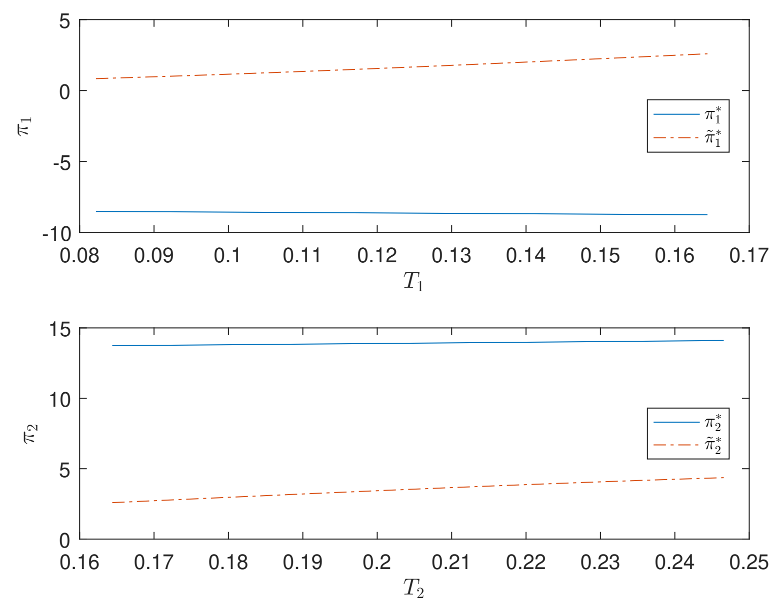


Figure 3.3: Optimal controls π_i^* with 2 contracts and $\tilde{\pi}_i^*$ with 1 contract, using parameters as displayed in Table 3.1, and T_1 ranges from [30/365,60/365], and T_2 ranges from [60/365,90/365].

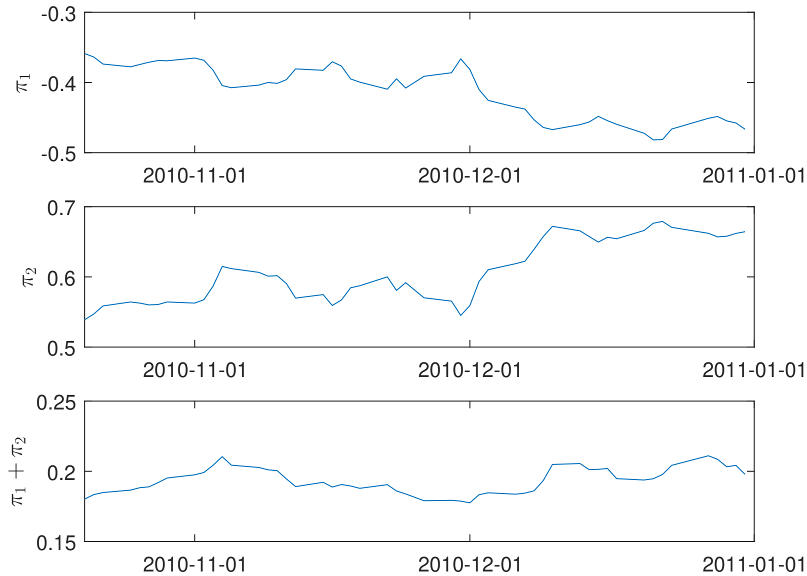


Figure 3.4: Optimal controls π_1^*, π_2^* and $\pi_1^*+\pi_2^*$ over the period Oct 2010-Dec 2010 using parameters as displayed in Table 3.1.

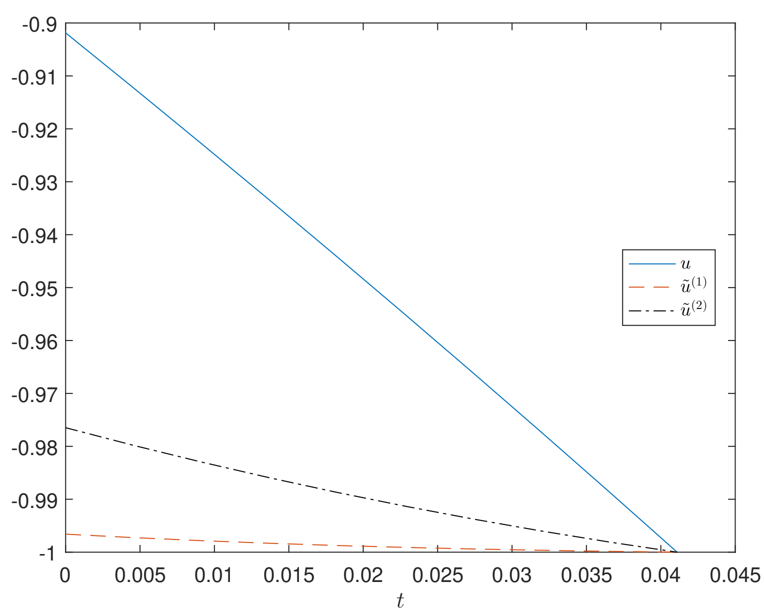


Figure 3.5: The value functions u, $\tilde{u}^{(1)}$ and $\tilde{u}^{(2)}$ at $w = 0$, with optimization horizon $T = 15/365$, maturity of F_1 is $T_1 = 30/365$, and maturity of $T_2 = 60/365$.

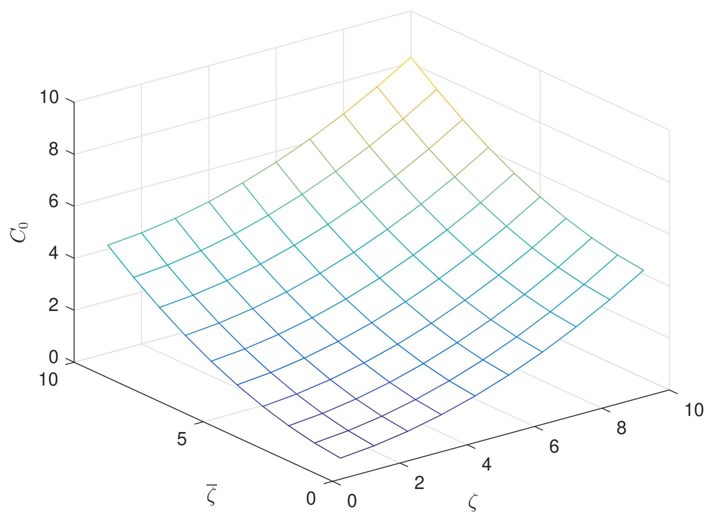


Figure 3.6: Certainty equivalent C_0 as a function of the market prices of risk ζ and $\bar{\zeta}$ under the CTOU model, with parameters as displayed in Table 3.1.

3.4 Summary

We have analyzed the problem of dynamically trading two futures contracts with the same underlying. Under a two-factor mean-reverting model for the spot price, we derive the futures price dynamics, solve the portfolio optimization problem in closed form, and give explicit optimal trading strategies. By studying the associated HJB equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. In addition to the analytic properties of our solutions, we also apply our results to VIX futures trading and present numerical examples to illustrate the optimal holdings.

In summary, the optimal controls are functions of time and the futures prices, but not functions of the spot price and its equilibrium level. They are inversely proportional to γ, as expected. Furthermore they depend only on the corresponding futures price F_i, but not the other futures price F_j for $j \neq i$. All of these features greatly simplify implementation.

We have focused only on volatility index futures to trade volatility in this chapter. Volatility trading, as described in [26], or in the book [112], in the early years mostly involved trading stock options directly, delta-hedged with the underlying stocks. These securities are publicly accessible to retail investors; on the other hand variance swaps have long been traded over-the-counter. More recently, products tied to realized variance began to appear in retail option exchanges. Furthermore, volatility indexes on a varieity of underlyings such as interest rates, single stocks, or even on VIX itself (VIX of VIX), have been created, and Exchange-Traded-Funds (ETFs) or Notes (ETNs) are the vehicles to obtain exposures to various aspects of volatility trading. The open question is whether the CTOU model, considered in this chapter, is still applicable for these indices; on the hand, since the underlying securities are no longer futures, the modeling of which will clearly require further refinement.

Since the VIX index is derived from S&P options, various models that incorporate dynamics of both S&P and VIX in the equations have been developed, see for example [27]. It will be fruitful to apply the stochastic approach for trading VIX futures under a more comprehensive joint model. Moreover, in addition to mean-reversion, since the VIX time series historically exhibits long periods of low activities, punctuated by sudden spikes, jump diffusion is obviously appropriate [108]. Another possibility is to incorporate self-excited jumps, followed by exponential decay, such as the dynamic contagion [39] model which generalizes the Hawkes model. They are frequently studied in the high frequency literature [11], but are potentially suitable for VIX modeling, since they capture the features of VIX well. The design of optimal trading strategies for securities other than VIX futures, and for models other than CTOU, is clearly a challenging and practical question.

3.5 Appendix

3.5.1 Drift of $df_t^{(i)}$ under CTOU

By Ito's Lemma, the drift of $df_t^{(i)}$, denoted by $m_i(t)$, is given by

$$m_i(t) = \frac{df^{(i)}}{dt} + \kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}}{dv} + \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}}{d\theta}$$

We have the following derivatives

$$\begin{aligned}\frac{df^{(i)}(t,v,\theta)}{dt} &= e^{-\kappa(T_i-t)}\kappa\,(v-\theta) - \frac{\left(e^{\kappa T_i + t\,\bar{\kappa}} - e^{\kappa t + T_i\bar{\kappa}}\right)^2\kappa^2\bar{\sigma}^2}{2e^{2T_i(\kappa+\bar{\kappa})}\left(\kappa-\bar{\kappa}\right)^2} \\ &\quad + \frac{\left(e^{-\kappa(T_i-t)} - e^{-\bar{\kappa}(T_i-t)}\right)\kappa\bar{\kappa}\left(\bar{\theta}-\theta\right)}{\kappa-\bar{\kappa}} - \frac{e^{-2\kappa(T_i-t)}\,\sigma^2}{2},\end{aligned}$$

$$\begin{aligned}&\kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}(t,v,\theta)}{dv} + \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}(t,v,\theta)}{d\theta} = \\ &\frac{e^{-(\kappa+\bar{\kappa})(T_i-t)}\left(e^{\kappa(T_i-t)}\kappa\left(\bar{\theta}\bar{\kappa} - \theta\bar{\kappa} + \bar{\sigma}\bar{\zeta}\right) + e^{\bar{\kappa}(T_i-t)}\left(\zeta\left(\kappa-\bar{\kappa}\right)\sigma - \left(\kappa\left(\kappa\left(v-\theta\right) + (\bar{\theta}-v)\bar{\kappa} + \bar{\sigma}\bar{\zeta}\right)\right)\right)\right)}{\kappa-\bar{\kappa}}.\end{aligned}$$

In turn, we obtain

$$m_i(t) = \frac{e^{-\bar{\kappa}(T_i-t)}}{\kappa-\bar{\kappa}}\kappa\bar{\sigma}\xi + \frac{e^{-\kappa(T_i-t)}}{(\kappa-\bar{\kappa})^2}(\kappa\bar{\sigma}\bar{\kappa}\xi - \kappa^2\bar{\sigma}\xi + \kappa^2\zeta\sigma - 2\kappa\zeta\bar{\kappa}\sigma + \zeta\bar{\kappa}^2\sigma)$$
$$+ \frac{\kappa^2\bar{\sigma}^2}{2(\kappa-\bar{\kappa})^2}(2e^{-(\kappa+\bar{\kappa})(T_i-t)} - e^{-2\bar{\kappa}(T_i-t)}) + \frac{e^{-2\kappa(T_i-t)}}{2(\kappa-\bar{\kappa})^2}(2\kappa\bar{\kappa}\sigma^2 - \kappa^2\bar{\sigma}^2 - \kappa^2\sigma^2 - \bar{\kappa}^2\sigma^2).$$

Interestingly, the drift is a deterministic function of time, and does not depend on v_t, θ_t, and $\bar{\theta}$. To see this, we collect v, θ and $\bar{\theta}$ in $m_i(t)$, and get

$$\frac{df^{(i)}(t,v,\theta)}{dt} = e^{-\kappa(T_i-t)}\kappa v - \frac{\kappa^2\bar{\sigma}^2\left(e^{\kappa T_i + t\bar{\kappa}} - e^{\kappa t + T_i\bar{\kappa}}\right)^2}{2e^{2T_i(\kappa+\bar{\kappa})}(\kappa-\bar{\kappa})^2} + \frac{\bar{\theta}\kappa\bar{\kappa}\left(e^{-\kappa(T_i-t)} - e^{-\bar{\kappa}(T_i-t)}\right)}{\kappa-\bar{\kappa}}$$
$$+ \frac{\kappa\theta\left(e^{-\bar{\kappa}(T_i-t)}\bar{\kappa} - e^{-\kappa(T_i-t)}\kappa\right)}{\kappa-\bar{\kappa}} - \frac{e^{-2\kappa(T_i-t)}\sigma^2}{2},$$

$$\kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}(t,v,\theta)}{dv} = e^{-\kappa(T_i-t)}\kappa\theta + e^{-\kappa(T_i-t)}(\zeta\sigma - \kappa v),$$

$$\bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}(t,v,\theta)}{d\theta} = \frac{\bar{\theta}\kappa\bar{\kappa}\left(e^{-\bar{\kappa}(T_i-t)} - e^{-\kappa(T_i-t)}\right)}{\kappa-\bar{\kappa}} + \frac{\theta\kappa\bar{\kappa}\left(e^{-\kappa(T_i-t)} - e^{-\bar{\kappa}(T_i-t)}\right)}{\kappa-\bar{\kappa}}$$
$$+ \frac{\kappa\bar{\sigma}\bar{\zeta}\left(e^{-\bar{\kappa}(T_i-t)} - e^{-\kappa(T_i-t)}\right)}{\kappa-\bar{\kappa}}.$$

When added together, the terms involving v, θ and $\bar{\theta}$ cancelled out, and we are left with (3.9).

3.5.2 Portfolio with a Single Futures Contract

We now discuss the case when the portfolio consists of only one futures contract. The system of SDEs for the wealth process and futures price is

$$\begin{bmatrix} dW_t \\ dF_t^{(1)} \end{bmatrix} = \begin{bmatrix} \tilde{\pi}_1\mu_1(t)F_t^{(1)} \\ \mu_1(t)F_t^{(1)} \end{bmatrix} dt + \begin{bmatrix} \tilde{\pi}_1\sigma_{v1}(t)F_t^{(1)} & \tilde{\pi}_1\sigma_{\theta 1}(t)F_t^{(1)} \\ \sigma_{v1}(t)F_t^{(1)} & \sigma_{\theta 1}(t)F_t^{(1)} \end{bmatrix} \begin{bmatrix} dZ_t^v \\ dZ_t^\theta \end{bmatrix}, \tag{3.39}$$

where we use the tilde notation to denote the single contract case. In order to avoid confusion when we later compare the optimal controls and the value function to the two contracts case, we keep the subscript i in $\tilde{\pi}_i \equiv \tilde{\pi}_i(t, F_i)$ to denote the optimal control in the single contract case when the contract has a maturity of $T_i, i = 1, 2$. The single contract case with the contract maturing in T_1, for example, can be interpreted as fixing π_2 identically to zero over the entire optimization horizon T.

We expect the value function $\tilde{u}(t, w, F_1)$ to solve the HJB equation

$$\tilde{u}_t + \sup_{\tilde{\pi}_1}[\,\tilde{\pi}_1\mu_1(t)F_1\tilde{u}_w + \tilde{\pi}_1\sigma_1(t)^2F_1^2\tilde{u}_{w1} + \frac{1}{2}\tilde{\pi}_1^2\sigma_1(t)^2F_1^2\tilde{u}_{ww}\,]$$
$$+ \frac{\sigma_1(t)^2}{2}F_1^2\tilde{u}_{11} + \mu_1(t)F_1\tilde{u}_1 = 0,$$

and the optimal control $\tilde{\pi}_1^*$ is given by

$$\tilde{\pi}_1^*(t, F_1) = \frac{\tilde{u}_w\mu_1(t) + F_1\tilde{u}_{w1}\sigma_1(t)^2}{F_1\tilde{u}_{ww}\sigma_1(t)^2}.$$

After substituting in $\tilde{\pi}_1^*$, we have the equation

$$\tilde{u}_t - \frac{\tilde{u}_w^2\mu_1(t)^2}{2\tilde{u}_{ww}\sigma_1(t)^2} - \frac{F_1\tilde{u}_w\tilde{u}_{w1}\mu_1(t)}{\tilde{u}_{ww}} + \frac{F_1(2\tilde{u}_1\tilde{u}_{ww}\mu_1(t) - F_1(\tilde{u}_{w1}^2 - \tilde{u}_{11}\tilde{u}_{ww})\sigma_1(t)^2)}{2\tilde{u}_{ww}} = 0.$$

Next, we apply the transformation

$$\tilde{u}(t, w, F_1) = -e^{-\gamma w} e^{\tilde{\Phi}(t, f_1)},$$

where $f_1 = \log F_1$, and the tilde on Φ to again denote the single contract case, to get the PDE for $\tilde{\Phi}$ as

$$-\tilde{\Phi}_t = \frac{\mu_1(t)^2}{2\sigma_1(t)^2} + \frac{\sigma_1(t)^2}{2}(\tilde{\Phi}_{11} - \tilde{\Phi}_1),$$

subject to $\tilde{\Phi}(T, f_1)$=0. We can see that $\tilde{\Phi}$ is a function of t only. Explicitly,

$$\frac{d\tilde{\Phi}}{dt} = -\frac{\mu_1(t)^2}{2\sigma_1(t)^2} = -\frac{\left(e^{-\bar{\kappa}\,(T_1-t)}\,\kappa\,\bar{\sigma}\,\xi - e^{-\kappa\,(T_1-t)}\,\left(\kappa\,\bar{\sigma}\,\xi - \kappa\,\zeta\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)\right)^2}{2\,(\kappa-\bar{\kappa})^2\,\left(\frac{\left(e^{-\kappa\,(T_1-t)} - e^{-\bar{\kappa}\,(T_1-t)}\right)^2\kappa^2\,\bar{\sigma}^2}{(\kappa-\bar{\kappa})^2} + e^{-2\,\kappa\,(T_1-t)}\,\sigma^2\right)}.$$

In turn, we numerically evaluate the integral

$$\tilde{\Phi}(t) = \int_t^T \frac{\mu_1(t')^2}{2\sigma_1(t')^2} dt'.$$

Now if we express the optimal control in terms of $\tilde{\Phi}$, we will have

$$\tilde{\pi}_1^*(t, F_1) = \frac{\mu_1(t) - \sigma_1(t)^2\tilde{\Phi}_1}{\gamma F_1 \sigma_1(t)^2} = \frac{\mu_1(t)}{\gamma F_1 \sigma_1(t)^2},$$

since $\tilde{\Phi}$ is a function of t only. We can see from (3.30) that $\tilde{\pi}_1^*$ is identical to that in the 2 contracts case, when $\rho_{12}(t)$ as defined in (3.16) equals zero. Explicitly, $\tilde{\pi}_1^*$ in the single contract case equals

$$\tilde{\pi}_1^*(t, F_1) = \frac{-\left(e^{-\bar{\kappa}(T_1-t)}\,\bar{\sigma}\,\kappa\,\xi\right) + e^{-\kappa\,(T_1-t)}\,\left(\bar{\sigma}\,\kappa\,\xi - \zeta\,\kappa\,\sigma + \zeta\,\bar{\kappa}\,\sigma\right)}{\gamma F_1\,\left(\bar{\kappa}-\kappa\right)\,\left(\frac{\left(e^{-\kappa\,(T_1-t)} - e^{-\bar{\kappa}\,(T_1-t)}\right)^2\bar{\sigma}^2\,\kappa^2}{(\kappa-\bar{\kappa})^2} + e^{-2\,\kappa\,(T_1-t)}\,\sigma^2\right)}. \tag{3.40}$$

3.5.3 CTOU with Correlation

We now generalize the CTOU to the case where Z_t^v and Z_t^θ are correlated with coefficient ρ. To this end, we solve the following linear parabolic PDE with constant coefficients for the futures contract $F(t, v, \theta)$ with maturity T_i:

$$\frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial v^2} + \frac{1}{2}\bar{\sigma}^2\frac{\partial F}{\partial \theta} + \kappa(\theta - v)\frac{\partial F}{\partial v} + \bar{\kappa}(\bar{\theta} - \theta)\frac{\partial F}{\partial \theta} + \rho\,\sigma\bar{\sigma}\frac{\partial^2 F}{\partial v \partial \theta} + \frac{\partial F}{\partial t} = 0,$$

coupled with the terminal condition $F(T_i, v, \theta) = e^v$. We assume $F(t, v, \theta)$ is of the form

$$F(t, v, \theta) = \exp(a(t)v + b(t)\theta + c(t)).$$

After substitution, we obtain the coupled set of linear ODE

$$a'(t) - \kappa a(t) = 0$$

$$b'(t) + \kappa a(t) - \bar{\kappa} b(t) = 0$$

$$c'(t) + \frac{1}{2}\sigma^2 a(t)^2 + \bar{\kappa}\bar{\theta} b(t) + \rho\,\bar{\sigma}\sigma a(t)b(t) + \frac{1}{2}\bar{\sigma}^2 b(t)^2 = 0$$

with the terminal conditions

$$a(T_i) = 1,\ b(T_i) = c(T_i) = 0.$$

We can solve for $a(t), b(t)$ and then $c(t)$ sequentially, and obtain

$$a(t) = \exp(-\kappa(T_i - t)),$$

$$b(t) = \frac{e^{\bar{\kappa}t - t(\bar{\kappa}-\kappa) - T_i(\bar{\kappa}-\kappa) - T_i\kappa}\left(e^{t(\bar{\kappa}-\kappa)} - e^{T_i(\bar{\kappa}-\kappa)}\right)\kappa}{\kappa - \bar{\kappa}},$$

$$\begin{aligned}
c(t) = &-\left(\frac{e^{-2\bar{\kappa}(t+3T_i)-(t+7T_i)\kappa+2(t+3T_i)(\bar{\kappa}+\kappa)}\bar{\kappa}\bar{\theta}}{\bar{\kappa}-\kappa}\right)\\
&-\frac{e^{4\bar{\kappa}(t+T_i)-2\bar{\kappa}(t+3T_i)}\bar{\sigma}^2\kappa^2}{4\bar{\kappa}(\bar{\kappa}-\kappa)^2}\\
&+\frac{e^{3\bar{\kappa}t+5\bar{\kappa}T_i-2\bar{\kappa}(t+3T_i)+t\kappa+7T_i\kappa-(t+7T_i)\kappa}\bar{\theta}\kappa\left(\bar{\kappa}^2-\kappa^2\right)}{(\bar{\kappa}-\kappa)^2\left(\bar{\kappa}+\kappa\right)}\\
&+\frac{e^{3\bar{\kappa}t+5\bar{\kappa}T_i-2\bar{\kappa}(t+3T_i)+2t\kappa+6T_i\kappa-(t+7T_i)\kappa}\bar{\sigma}\kappa\left(\bar{\sigma}\kappa+\left(\bar{\kappa}-\kappa\right)\rho\sigma\right)}{(\bar{\kappa}-\kappa)^2\left(\bar{\kappa}+\kappa\right)}\\
&-\frac{e^{2\bar{\kappa}t+6\bar{\kappa}T_i-2\bar{\kappa}(t+3T_i)+3t\kappa+5T_i\kappa-(t+7T_i)\kappa}\left(\bar{\sigma}^2\kappa^2+2\bar{\sigma}\left(\bar{\kappa}-\kappa\right)\kappa\rho\sigma+\left(\bar{\kappa}-\kappa\right)^2\sigma^2\right)}{4(\bar{\kappa}-\kappa)^2\kappa}\\
&+\frac{e^{2\bar{\kappa}t+6\bar{\kappa}T_i-2\bar{\kappa}(t+3T_i)+t\kappa+7T_i\kappa-(t+7T_i)\kappa}\left(\bar{\sigma}^2\kappa^2+\bar{\kappa}^2\left(4\bar{\theta}\kappa+\sigma^2\right)+\bar{\kappa}\kappa\left(4\bar{\theta}\kappa+\sigma\left(2\bar{\sigma}\rho+\sigma\right)\right)\right)}{4\bar{\kappa}\kappa\left(\bar{\kappa}+\kappa\right)}.
\end{aligned}$$

We can now apply Ito's lemma to obtain the futures SDE under $\mathbb{P}$, as

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)\,dt + \sigma_{vi}(t)\,dZ_t^v + \sigma_{\theta i}(t)\,dZ_t^\theta,$$

where

$$\sigma_{vi}(t) = \frac{-\left(e^{\bar{\kappa}t+T_i\kappa}\bar{\sigma}\kappa\rho\right) + e^{kT_i+t\kappa}\left(\bar{\sigma}\kappa\rho+\left(\bar{\kappa}-\kappa\right)\sigma\right)}{e^{T_i(\bar{\kappa}+\kappa)}\left(\bar{\kappa}-\kappa\right)},$$

$$\sigma_{\theta i}(t) = \frac{\left(e^{\bar{\kappa}T_i+t\kappa} - e^{\bar{\kappa}t+T_i\kappa}\right)\bar{\sigma}\kappa\sqrt{1-\rho^2}}{e^{T_i(\bar{\kappa}+\kappa)}\left(\bar{\kappa}-\kappa\right)}.$$

Again we define

$$\sigma_i(t)^2 \equiv \sigma_{vi}(t)^2 + \sigma_{\theta i}(t)^2.$$

The drift of $df_t^{(i)}$, denoted by $m_i(t)$, is

$$m_i(t) = \frac{df^{(i)}}{dt} + \kappa\left(\theta + \frac{\sigma\zeta}{\kappa} - v\right)\frac{df^{(i)}}{dv} + \bar{\kappa}\left(\bar{\theta} + \frac{\bar{\sigma}\bar{\zeta}}{\bar{\kappa}} - \theta\right)\frac{df^{(i)}}{d\theta}.$$

The drift of $dF_t^{(i)}$, denoted by $\mu_i(t)$, is

$$\begin{aligned}
\mu_i(t) =& m_i(t) + \frac{\sigma_i(t)^2}{2}\\
=& \frac{e^{-\bar{\kappa}(T_i-t)}\kappa\bar{\sigma}\bar{\zeta} - e^{-\kappa(T_i-t)}\left(\kappa\bar{\sigma}\bar{\zeta} - \kappa\zeta\sigma + \zeta\bar{\kappa}\sigma\right)}{\kappa-\bar{\kappa}},
\end{aligned}$$

which is identical to that in the case of no correlation.

We use the following matrix notation

$$\boldsymbol{\mu}(t) = \begin{bmatrix}\mu_1(t)\\ \mu_2(t)\end{bmatrix}, \boldsymbol{v}(t) = \begin{bmatrix}\sigma_{v1}(t) & \sigma_{\theta 1}(t)\\ \sigma_{v2}(t) & \sigma_{\theta 2}(t)\end{bmatrix}.$$

From Chapter 5, we know immediately that the value function is given by

$$\begin{aligned} u(t,w) &= -\exp\left(-\gamma w - \frac{1}{2}\boldsymbol{\mu}(t)'(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t)(T-t)\right) \\ &= -\exp\left(-\gamma w - \phi(t)\right), \end{aligned}$$

where

$$\phi(t) = \frac{1}{2}\frac{\zeta^2 + \bar{\zeta}^2 - 2\zeta\bar{\zeta}\rho}{1-\rho^2}.$$

The optimal control is

$$\boldsymbol{\pi}^*(t, F_1, F_2) = \frac{1}{\gamma}(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t) = \begin{bmatrix} \pi_1^* \\ \pi_2^* \end{bmatrix},$$

where

$$\begin{aligned} \pi_1^* &= \frac{e^{(-t+T_1)(\bar{\kappa}+\kappa)}\left(e^{\bar{\kappa}t+T_2\kappa}\bar{\sigma}\kappa\left(\zeta-\bar{\zeta}\rho\right)+e^{\bar{\kappa}T_2+t\kappa}\left(\bar{\sigma}\left(-\zeta\kappa+\kappa\bar{\zeta}\rho\right)+\left(\bar{\kappa}-\kappa\right)\left(\bar{\zeta}-\zeta\rho\right)\sigma\right)\right)}{\gamma F_1\left(e^{\bar{\kappa}T_2+T_1\kappa}-e^{\bar{\kappa}T_1+T_2\kappa}\right)\bar{\sigma}\kappa\left(\rho^2-1\right)\sigma} \\ \pi_2^* &= \frac{e^{(-t+T_2)(\bar{\kappa}+\kappa)}\left(e^{\bar{\kappa}t+T_1\kappa}\bar{\sigma}\kappa\left(\zeta-\bar{\zeta}\rho\right)+e^{\bar{\kappa}T_1+t\kappa}\left(\bar{\sigma}\left(-\zeta\kappa+\kappa\bar{\zeta}\rho\right)+\left(\bar{\kappa}-\kappa\right)\left(\bar{\zeta}-\zeta\rho\right)\sigma\right)\right)}{\gamma F_2\left(-e^{\bar{\kappa}T_2+T_1\kappa}+e^{\bar{\kappa}T_1+T_2\kappa}\right)\bar{\sigma}\kappa\left(\rho^2-1\right)\sigma}. \end{aligned}$$

Chapter 4

Commodity Futures Trading under the Schwartz Model

Missing in the literature are investigations of the trading of a portfolio of commodity futures contracts. Thus in this chapter, as published in [82], we shift the focus from futures on a mean-reverting spot index (such as VIX) to commodity futures, as motivated by the prevalence of managed futures funds. Managed futures funds constitute a significant segment in the universe of alternative assets. These funds are managed by professional investment individuals or management companies known as Commodity Trading Advisors (CTAs), and typically involve trading futures on commodities, currencies, interest rates, and other assets. Regulated and monitored by both government agencies such as the U.S. Commodity Futures Trading Commission and the National Futures Association, this class of assets has grown to over US$350 billion in 2017. One appeal of managed-futures strategies is their potential to produce uncorrelated and superior returns, as well as different risk-return profiles, as compared to the equity market [63, 49]. While the types of securities traded and strategies are conceivably diverse among managed futures funds, details of the employed strategies are often unknown. It was suggested in [68] that momentum-based strategies can help explain the returns of these funds.

In light of the tremendous growth, we aim to shed light on the opacity of this industry by proposing a feasible strategy to trade commodity futures, based on the well-established machinery of stochastic control theory. In this chapter, we analyze a stochastic dynamic control approach for portfolio optimization in which the commodity price dynamics and investor's risk preference are incorporated. We apply a no-arbitrage approach to construct futures prices from a stochastic spot model. Specifically, we adopt the well-known Schwartz two-factor model as described in [107], which takes into account the stochastic convenience yield in commodity prices. We determine the optimal futures trading strategies by solving the associated HJB equations in closed form. We use the Schwartz model due to its analytic tractability, which enables us to obtain closed-form expressions for the optimal holdings, despite the various shortcomings as pointed out in [25]. Parameter estimation for the Schwartz model is also well-studied, and is implemented using Kalman filtering, as reported in the R package *schwartz97*, and independently in [54].

The explicit formulae of our strategies allow for straight-forward financial interpretations and instant implementation. Moreover, our optimal strategies are explicit functions of the prices of the futures included in the portfolio, but do not require the continuous monitoring of the spot price or stochastic convenience yield. Related to the strategies, we also discuss the corresponding wealth process and certainty equivalent from futures trading. We provide some numerical examples and illustrate the optimal trading strategies using WTI crude oil futures data.

With the closed-form formulae, we found that the optimal number of contracts held, on both the long and short sides, are decreasing as the volatility of the stochastic convenience yield process increases. Moreover, the position sizes are still inversely proportional to risk aversion coefficient, similar to previous chapters. Based on realistic parameters calibrated from historical data, we found that positions change little with respect to maturities; in other words they are insensitive to time to maturity, ceteris paribus. We also found that the certainty equivalent for trading two contracts simultaneously is significantly greater than that derived from trading only a single contract, regardless of the choice of maturity. It will be shown later that this also holds when we divide initial wealth into two halves and trade two separate single maturity contracts simultaneously.

There is a host of research on the pricing of futures, but relatively few studies apply dynamic stochastic control methods to optimize futures portfolios. Among them, a study that examines how investors risk behavior affects the portfolio allocation to commodity futures is [6]. [19] consider trading a pair of futures but use the arithmetic Brownian motion. In a recent study, [5] study the problem of dynamically trading the price spread between a futures contract and its spot asset under a stochastic basis model. They model the basis process by a scaled Brownian bridge, and solve a utility maximization problem to derive the optimal trading strategies. These two related studies do not account for the well-observed no-arbitrage price relationships and term-structure in the futures market. Our work fills in the void by using a tractable spot model that can generate no-arbitrage futures prices and effectively capture their joint price evolutions. On the other hand, in contrast to a stochastic control approach to futures trading considered here, [75] introduced an optimal stopping approach to determine the optimal timing to open or close a futures position under three single-factor mean-reverting spot models. Futures portfolios are also often used to track the spot price movements, and we refer to [78, 79] for examples using gold and VIX futures. In this chapter we worked with the relatively simple two-factor Schwartz model for commodity;

we will leave to future research to investigate the application of our method to more complicated models such as [43].

4.1 Futures Price Dynamics

Let us denote the commodity spot price process by $(S_t)_{t\geq 0}$. Under the [107] model, the spot price is driven by a stochastic instantaneous convenience yield, denoted by $(\delta_t)_{t\geq 0}$ here. This convenience yield, which was originally used in the context of commodity futures, reflects the value of direct access minus the cost of carry and can be interpreted as the"dividend yield" for holding the physical asset. It is the "flow of services accruing to the holder of the spot commodity but not to the owner of the futures contract" as explained in [107].

For the spot asset, we consider its log price, denoted by X_t. Under the [107] model, it satisfies the system of stochastic differential equations (SDEs) under the physical probability measure $\mathbb{P}$:

$$\begin{aligned} X_t &= \log(S_t), && (4.1)\\ dX_t &= \left(\mu - \frac{\eta^2}{2} - \delta_t\right) dt + \eta dZ_t^s, && (4.2)\\ d\delta_t &= \kappa\left(\alpha - \delta_t\right) dt + \bar{\eta} dZ_t^\delta. && (4.3) \end{aligned}$$

Here, Z_t^s and Z_t^δ are two standard Brownian motions under $\mathbb{P}$ with instantaneous correlation $\rho \in (-1, 1)$. The stochastic convenience yield follows the Ornstein-Uhlenbeck model, which is mean-reverting with a constant equilibrium level α, volatility $\bar{\eta}$, and speed of mean-reversion equal to κ. We require that $\kappa, \bar{\eta}, \eta > 0$ and $\mu, \alpha \in \mathbb{R}$.

The investor's portfolio optimization problem will be formulated under the physical measure $\mathbb{P}$, but in order to price the commodity futures we need to work with the risk-neutral pricing measure $\mathbb{Q}$. To this end, we assume a constant interest rate $r \geq 0$, and apply a change of measure from $\mathbb{P}$ to $\mathbb{Q}$. The $\mathbb{Q}$-dynamics of the correlated Brownian motions (Z_t^s, Z_t^δ) are given by

$$\begin{aligned} d\tilde{Z}_t^s &= \frac{\mu - r}{\eta} dt + dZ_t^s, && (4.4)\\ d\tilde{Z}_t^\delta &= \frac{\lambda}{\bar{\eta}} dt + dZ_t^\delta. && (4.5) \end{aligned}$$

Consequently, the risk-neutral log spot price evolves according to

$$\begin{aligned} dX_t &= \left(r - \delta_t - \frac{\eta^2}{2}\right) dt + \eta d\tilde{Z}_t^s,\\ d\delta_t &= \kappa(\tilde{\alpha} - \delta_t) dt + \bar{\eta} d\tilde{Z}_t^\delta, \end{aligned}$$

where we have defined the risk-neutral equilibrium level for the convenience yield by

$$\tilde{\alpha} \equiv \alpha - \frac{\lambda}{\kappa}.$$

It is adjusted by the ratio of the market price of risk λ associated with Z_t^δ and the speed of mean reversion κ. With a constant λ, the convenience yield again follows the Ornstein-Uhlenbeck model under measure $\mathbb{Q}$ but with a different equilibrium level compared to that under measure $\mathbb{P}$.

We consider a commodity market that consists of n traded futures contracts with maturities $T_i, i = 1, \dots, n$. Let

$$F_t^{(i)} \equiv F^{(i)}(t, X_t, \delta_t) = \mathbb{E}[\, e^{X_T} \mid X_t, \delta_t\,]$$

be the price of the T_i-futures at time t, which is a function of time t, current log spot price X_t, and convenience yield δ_t. For any $i = 1, \dots, n$, the price function $F^{(i)}(t, X, \delta)$ satisfies the PDE

$$\frac{\eta^2}{2}\frac{\partial^2 F^{(i)}}{\partial X^2} + \rho\eta\bar{\eta}\frac{\partial^2 F^{(i)}}{\partial X \partial \delta} + \frac{\bar{\eta}^2}{2}\frac{\partial^2 F^{(i)}}{\partial \delta^2} + \left(r - \delta - \frac{\eta^2}{2}\right)\frac{\partial F^{(i)}}{\partial X} + \kappa(\tilde{\alpha} - \delta)\frac{\partial F^{(i)}}{\partial \delta} = -\frac{\partial F^{(i)}}{\partial t}, \qquad (4.6)$$

for $(t, x, \delta) \in [0, T_i) \times (-\infty, \infty) \times (-\infty, \infty)$, where we have compressed the dependence of $F^{(i)}$ on (t, X, δ). The terminal condition is $F^{(i)}(T_i, X, \delta) = \exp(X)$ for $x \in \mathbb{R}$. As is well-known (see [107, 37]), the futures price admits the exponential affine form:

$$F_t^{(i)} = \exp\left(X_t + A_i(t) + B_i(t)\delta_t\right) \tag{4.7}$$

for some functions $A_i(t)$ and $B_i(t)$ that depend only on time t and not the state variables. The functions $A_i(t)$ and $B_i(t)$ are found from the ODEs

$$r + \frac{\bar{\eta}}{2} B_i(t)^2 + B_i(t)(\alpha\kappa + \rho\eta\bar{\eta}) + A_i'(t) = 0, \tag{4.8}$$

$$B_i'(t) - \kappa B_i(t) - 1 = 0, \tag{4.9}$$

for $t \in [0, T_i)$, with terminal conditions $A_i(T_i) = 0$ and $B_i(T_i) = 0$. The ODEs (4.8) and (4.9) admit the following explicit solutions:

$$\begin{aligned} A_i(t) &= \left(r - \tilde{\alpha} + \frac{\bar{\eta}^2}{2\kappa^2} - \frac{\eta\bar{\eta}\rho}{\kappa}\right)(T_i - t) \\ &\quad + \frac{\bar{\eta}^2}{4}\frac{1 - e^{-2\kappa(T_i - t)}}{\kappa^3} + \left(\tilde{\alpha}\kappa + \eta\bar{\eta}\rho - \frac{\bar{\eta}^2}{\kappa}\right)\frac{1 - e^{-\kappa(T_i - t)}}{\kappa^2}, && (4.10) \\ B_i(t) &= -\frac{1 - e^{-\kappa(T_i - t)}}{\kappa}. && (4.11) \end{aligned}$$

Applying Ito's formula to (4.7), the T_i-futures price evolves according to the SDE

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)dt + \eta dZ_t^s + \bar{\eta} B_i(t) dZ_t^\delta, \tag{4.12}$$

under the physical measure $\mathbb{P}$, where the drift is given by

$$\begin{aligned} \mu_i(t) &= (\lambda + \tilde{\alpha}\kappa + \rho\bar{\eta}\eta)B_i(t) + \frac{\bar{\eta}^2}{2}B_i(t)^2 + \mu + A_i'(t) + \delta(B_i'(t) - \kappa B_i(t) - 1) && (4.13) \\ &= \mu - r - \frac{\lambda(1 - e^{-\kappa(T_i - t)})}{\kappa}. && (4.14) \end{aligned}$$

The last equality follows from (4.8) and (4.11). As a consequence, the drift of $F_t^{(i)}$ is independent of X_t and δ_t, meaning that the investor's value function (see (4.22) or (4.32)) will also be independent of X_t and δ_t. This turns out to be a crucial feature that greatly simplifies the investor's portfolio optimization problem and ultimately leads to an explicit solution.

To facilitate presentation, let us rewrite the linear combination of dZ_t^s and dZ_t^δ in (4.12) as

$$\sigma_i(t) dZ_t^{(i)} \equiv \eta dZ_t^s + \bar{\eta} B_i(t) dZ_t^\delta,$$

where $Z_t^{(i)}$ is a standard Brownian motion and

$$\sigma_i(t)^2 = \eta^2 + 2\rho\bar{\eta}\eta B_i(t) + \bar{\eta}^2 B_i(t)^2 \tag{4.15}$$

is the instantaneous volatility coefficient.

Under this model, futures prices are not independent and admit a specific correlation structure. For example, consider the T_1 and T_2 contracts. The SDE for the respective futures price is

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)dt + \sigma_i(t)dZ_t^{(i)}, \quad i \in \{1, 2\}, \tag{4.16}$$

The two Brownian motions, $Z_t^{(1)}$ and $Z_t^{(2)}$, are correlated with

$$dZ_t^{(1)}\, dZ_t^{(2)} = \rho_{12}(t)\, dt.$$

where

$$\rho_{12}(t) = \frac{\bar{\eta}^2 B_1(t)B_2(t) + (B_1(t) + B_2(t))\rho\eta\bar{\eta} + \eta^2}{\sigma_1(t)\sigma_2(t)} \tag{4.17}$$

is the instantaneous correlation that depends not only on the spot model parameters $(\rho, \eta, \bar{\eta})$ but also the two futures price functions through $B_1(t)$ and $B_2(t)$.

4.2 Utility Maximization Problem

We now present the mathematical formulation for the futures portfolio optimization problem. To begin, we discuss the case where the investor trades only futures with the same maturity in Section 4.2.1. Then, we extend the analysis to optimize a portfolio with two different futures in Section 4.2.2. We will also investigate in Section 4.2.4 the value of trading using the notion of certainty equivalent.

4.2.1 Single-Maturity Futures Portfolio

Suppose that the investor trades only futures of a single maturity T_i for some chosen $i \in \{1, 2, \ldots, n\}$. The trading horizon, denoted by T, must be equal to or shorter than the chosen maturity T_i, so we require $T \leq T_i$.

We will let $\tilde{\pi}_i(t, F_i)$ denote the number of T_i-futures contracts held in the portfolio. The investor can choose the size of the position in the T_i-futures, and the position can be long or short at any time. For brevity, we may write $\tilde{\pi}_i \equiv \tilde{\pi}_i(t, F_i)$.

Without loss of generality, we arbitrarily set $i = 1$ in our presentation of the optimization problem and solution. The investor is assumed to trade only the futures contract and not other risky or risk-free assets. The dynamic portfolio consists of $\tilde{\pi}_1(t, F_1)$ units of T_1-futures at time t. The self-financing condition means that the wealth process satisfies

$$d\tilde{W}_t = \tilde{\pi}_1(t, F_t^{(1)})\, dF_t^{(1)}. \tag{4.18}$$

Applying the futures price equations (4.7) and (4.12), we can express the system of SDEs for the wealth process and futures price as

$$\begin{bmatrix} d\tilde{W}_t \\ dF_t^{(1)} \end{bmatrix} = \begin{bmatrix} \tilde{\pi}_1 \mu_1(t) F_t^{(1)} \\ \mu_1(t) F_t^{(1)} \end{bmatrix} dt + \begin{bmatrix} \tilde{\pi}_1 \eta F_t^{(1)} & \tilde{\pi}_1 \bar{\eta} B_1(t) F_t^{(1)} \\ \eta F_t^{(1)} & \bar{\eta} B_1(t) F_t^{(1)} \end{bmatrix} \begin{bmatrix} dZ_t^s \\ dZ_t^\delta \end{bmatrix}, \tag{4.19}$$

$$= \begin{bmatrix} \tilde{\pi}_1 \mu_1(t) F_t^{(1)} \\ \mu_1(t) F_t^{(1)} \end{bmatrix} dt + \begin{bmatrix} \tilde{\pi}_1 \sigma_1(t) F_t^{(1)} \\ \sigma_1(t) F_t^{(1)} \end{bmatrix} dZ_t^{(1)}. \tag{4.20}$$

A control $\tilde{\pi}_1$ is said to be admissible if $\tilde{\pi}_1$ is real-valued progressively measurable, and is such that the system of SDE (4.19) admits a unique solution $(\tilde{W}_t, F_t^{(1)})$ and the admissibility condition

$$\mathbb{E}\left(\int_t^T \tilde{\pi}_1(s, F_s^{(1)})^2\, (F_s^{(1)})^2 ds \right) < \infty$$

is satisfied. We denote by $\tilde{\mathcal{A}}_t$ the set of admissible strategies in this case given an initial investment time t.

The investor's risk preference is described by the exponential utility function

$$U(w) = -e^{-\gamma w}, \quad \text{for } w \in \mathbb{R}, \tag{4.21}$$

where $\gamma > 0$ is the constant risk aversion parameter. For a given trading horizon, $[0, T]$, the investor seeks an admissible strategy that maximizes the expected utility of terminal wealth at time T by solving the optimization problem

$$\tilde{V}(t, w, F_1) = \sup_{\tilde{\pi}_1 \in \tilde{\mathcal{A}}_t} \mathbb{E}\left(U(\tilde{W}_T) \,|\, \tilde{W}_t = w, F_t^{(1)} = F_1 \right). \tag{4.22}$$

We note that the value function is only a function of time t, current wealth w, and current futures price F_1, and does not depend on the current spot price or convenience yield. Following the standard verification approach to dynamic programming [58, 105, 96], we assume the existence of a sufficiently smooth candidate solution $\tilde{u}(t, w, F_1)$, which will later be shown to be equal to the value function $\tilde{V}$ in (4.22).

To facilitate presentation, we define the following partial derivatives

$$\tilde{u}_t = \frac{\partial \tilde{u}}{\partial t}, \quad \tilde{u}_w = \frac{\partial \tilde{u}}{\partial w}, \quad \tilde{u}_{ww} = \frac{\partial^2 \tilde{u}}{\partial w^2},$$
$$\tilde{u}_1 = \frac{\partial \tilde{u}}{\partial F_1}, \quad \tilde{u}_{11} = \frac{\partial^2 \tilde{u}}{\partial F_1^2}, \quad \tilde{u}_{w1} = \frac{\partial^2 \tilde{u}}{\partial w \partial F_1}.$$

We expect the candidate function $\tilde{u}(t, w, F_1)$ to solve the HJB equation

$$\begin{aligned} &\tilde{u}_t + \sup_{\tilde{\pi}_1} \{ \, \tilde{\pi}_1 \mu_1(t) F_1 \tilde{u}_w + \tilde{\pi}_1 \sigma_1(t)^2 F_1^2 \tilde{u}_{w1} + \frac{1}{2} \tilde{\pi}_1^2 \sigma_1(t)^2 F_1^2 \tilde{u}_{ww} \, \} \\ &\quad + \frac{\sigma_1(t)^2}{2} F_1^2 \tilde{u}_{11} + \mu_1(t) F_1 \tilde{u}_1 = 0, \end{aligned} \tag{4.23}$$

for $(t, w, F_1) \in [0, T) \times \mathbb{R} \times \mathbb{R}_+$, with terminal condition $\tilde{u}(T, w, F_1) = e^{-\gamma w}$ for $(w, F_1) \in \mathbb{R} \times \mathbb{R}_+$. Performing the optimization in (4.25), we can express the optimal control $\tilde{\pi}_1^*$ as

$$\tilde{\pi}_1^*(t, F_1) = \frac{\tilde{u}_w \mu_1(t) + F_1 \tilde{u}_{w1} \sigma_1(t)^2}{F_1 \tilde{u}_{ww} \sigma_1(t)^2}. \tag{4.24}$$

Substituting this into (4.25), we obtain the nonlinear PDE

$$\tilde{u}_t - \frac{\tilde{u}_w^2 \mu_1(t)^2}{2\tilde{u}_{ww} \sigma_1(t)^2} - \frac{F_1 \tilde{u}_w \tilde{u}_{w1} \mu_1(t)}{\tilde{u}_{ww}} + \frac{F_1(2\tilde{u}_1 \tilde{u}_{ww} \mu_1(t) - F_1(\tilde{u}_{w1}^2 - \tilde{u}_{11}\tilde{u}_{ww})\sigma_1(t)^2)}{2\tilde{u}_{ww}} = 0. \tag{4.25}$$

Next, we conjecture that $\tilde{u}$ depends on t and w only, and apply the transformation

$$\tilde{u}(t, w) = -e^{-\gamma w - \tilde{\Phi}(t)}, \tag{4.26}$$

for some function $\tilde{\Phi}(t)$ to be determined. By direct substitution and computation, we obtain the ODE

$$\frac{d\tilde{\Phi}}{dt} = -\frac{\mu_1(t)^2}{2\sigma_1(t)^2} = -\frac{1}{2} \frac{(\lambda(1 - e^{-\kappa(T_1 - t)}) - \kappa(\mu - r))^2}{(1 - e^{-\kappa(T_1 - t)})^2 \bar{\eta}^2 - 2(1 - e^{-\kappa(T_1 - t)})\kappa\rho\eta\bar{\eta} + \kappa^2\eta^2}, \tag{4.27}$$

subject to $\tilde{\Phi}(T)$=0. In turn, we obtain $\tilde{\Phi}(t)$ by integration

$$\tilde{\Phi}(t) = \int_t^T \frac{\mu_1(t')^2}{2\sigma_1(t')^2} dt', \qquad 0 \leq t \leq T.$$

Applying (4.26) to (4.24), we obtain the optimal strategy

$$\tilde{\pi}_1^*(t, F_1) = \frac{\mu_1(t) - \sigma_1(t)^2 \tilde{\Phi}_1}{\gamma F_1 \sigma_1(t)^2} = \frac{\mu_1(t)}{\gamma F_1 \sigma_1(t)^2}. \tag{4.28}$$

Using (4.11), (4.14), and (4.15), the optimal strategy $\tilde{\pi}_1^*$ in the single-contract case is explicitly given by

$$\tilde{\pi}_1^*(t, F_1) = \frac{1}{\gamma F_1} \frac{\kappa(\lambda(1 - e^{-\kappa(T_1 - t)}) - \kappa(\mu - r))}{(1 - e^{-\kappa(T_1 - t)})^2 \bar{\eta}^2 - 2(1 - e^{-\kappa(T_1 - t)})\kappa\rho\eta\bar{\eta} + \kappa^2\eta^2}. \tag{4.29}$$

We observe from (4.29) that $\tilde{\pi}_1^*$ is inversely proportional to γ and F_1. This means that a higher risk aversion will reduce the size of the investor's position. A higher futures price will also have the same effect. However, the total cash amount invested in the futures, i.e. $\tilde{\pi}_1^*(t, F_1)F_1$, does not vary with the futures price, and is in fact a deterministic function of time. Note that the investor's position is independent of the equilibrium level of the convenience yield α or $\tilde{\alpha}$, but it depends on the speed of mean reversion κ, volatility $\bar{\eta}$, and market price of risk λ of the convenience yield.

4.2.2 Trading Futures of Two Different Maturities

We now consider the utility maximization problem involving a pair of futures with different maturities. Without loss of generality, let T_1 and T_2 be the two maturities of the futures in the portfolio. The trading horizon T satisfies $T \le \min\{T_1, T_2\}$. The investor continuously trades only the two futures over time. The trading wealth satisfies the self-financing condition

$$dW_t = \pi_1(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(1)} + \pi_2(t, F_t^{(1)}, F_t^{(2)})\, dF_t^{(2)}, \tag{4.30}$$

where $\pi_i(t, F_t^{(1)}, F_t^{(2)})$, $i = 1, 2$, denote the number of T_i-futures held. If it is negative, the corresponding futures position is short. For notational simplicity, we may write $\pi_i \equiv \pi_i(t, F_t^{(1)}, F_t^{(2)})$. Writing the trading wealth and two futures prices together in terms of two fundamental sources of randomness $(Z_t^{(1)}, Z_t^{(2)})$, we get

$$\begin{bmatrix} dW_t \\ dF_t^{(1)} \\ dF_t^{(2)} \end{bmatrix} = \begin{bmatrix} \pi_1\mu_1(t)F_t^{(1)} + \pi_2\mu_2(t)F_t^{(2)} \\ \mu_1(t)F_t^{(1)} \\ \mu_2(t)F_t^{(2)} \end{bmatrix} dt + \begin{bmatrix} \pi_1\sigma_1(t)F_t^{(1)} & \pi_2\sigma_2(t)F_t^{(2)} \\ \sigma_1(t)F_t^{(1)} & 0 \\ 0 & \sigma_2(t)F_t^{(2)} \end{bmatrix} \begin{bmatrix} dZ_t^{(1)} \\ dZ_t^{(2)} \end{bmatrix}. \tag{4.31}$$

A pair of controls (π_1, π_2) is said to be admissible if it is real-valued progressively measurable, and such that the system of SDE (4.31) admits a unique solution $(W_t, F_t^{(1)}, F_t^{(2)})$ and the integrability condition $\mathbb{E}\big(\int_t^T [\pi_i(s, F_s^{(1)}, F_s^{(2)})F_s^{(1)}]^2 ds\big) < \infty$, for $i = 1, 2$, is satisfied. We denote by $\mathcal{A}_t$ the set of admissible controls with an initial time of investment t. Next, we define the value function $V(t, w, F_1, F_2)$ of the investor's portfolio optimization problem. The investor seeks an admissible strategy (π_1, π_2) that maximizes the expected utility from wealth at time T, that is,

$$V(t, w, F_1, F_2) = \sup_{(\pi_1,\pi_2)\in\mathcal{A}_t} \mathbb{E}\left(U(W_T) \,|\, W_t = w, F_t^{(1)} = F_1, F_t^{(2)} = F_2\right). \tag{4.32}$$

Following the standard verification approach to dynamic programming [58, 105, 96], we assume the existence of a sufficiently smooth candidate solution $u(t, w, F_1, F_2)$, which will later be shown to be equal to the value function V in (4.32).

HJB Equation and Closed-Form Solution

To facilitate presentation, we define the following partial derivatives

$$u_t = \frac{\partial u}{\partial t}, \quad u_w = \frac{\partial u}{\partial w}, \quad u_{ww} = \frac{\partial^2 u}{\partial w^2},$$
$$u_1 = \frac{\partial u}{\partial F_1}, \quad u_{11} = \frac{\partial^2 u}{\partial F_1^2}, \quad u_2 = \frac{\partial u}{\partial F_2}, \quad u_{22} = \frac{\partial^2 u}{\partial F_2^2},$$
$$u_{w1} = \frac{\partial^2 u}{\partial w \partial F_1}, \quad u_{w2} = \frac{\partial^2 u}{\partial w \partial F_2}, \quad u_{12} = \frac{\partial^2 u}{\partial F_1 \partial F_2}.$$

We determine the candidate solution $u(t, w, F_1, F_2)$ by solving the HJB equation

$$\begin{aligned} u_t + \sup_{\pi_1,\pi_2} \big[&(\pi_1\mu_1(t)F_1 + \pi_2\mu_2(t)F_2)u_w \\ &+ (\pi_1\sigma_1(t)^2F_1^2 + \pi_2\rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2)u_{w1} + (\pi_2\sigma_2(t)^2F_2^2 + \pi_1\rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2)u_{w2} \\ &+ \frac{1}{2}(\pi_1^2\sigma_1(t)^2F_1^2 + \pi_2^2\sigma_2(t)^2F_2^2 + \rho_{12}(t)\pi_1\pi_2\sigma_1(t)\sigma_2(t)F_1F_2)u_{ww}\big] + \mu_1(t)F_1u_1 + \mu_2(t)F_2u_2 \\ &+ \frac{\sigma_1(t)^2}{2}F_1^2u_{11} + \frac{\sigma_2(t)^2}{2}F_2^2u_{22} + \rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2u_{12} = 0, \end{aligned} \tag{4.33}$$

for $(t, w, F_1, F_2) \in [0, T) \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$, along with the terminal condition

$$u(T, w, F_1, F_2) = -e^{-\gamma w}, \qquad \text{for } (w, F_1, F_2) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+.$$

We introduce the linear operator $\mathcal{L}^{(\pi_1,\pi_2)}$ that will be needed in order to apply the verification theorem in Appendix A:

$$\begin{aligned}\mathcal{L}^{(\pi_1,\pi_2)}u &= (\pi_1\mu_1(t)F_1 + \pi_2\mu_2(t)F_2)u_w \\ &\quad + (\pi_1\sigma_1(t)^2F_1^2 + \pi_2\rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2)u_{w1} + (\pi_2\sigma_2(t)^2F_2^2 + \pi_1\rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2)u_{w2} \\ &\quad + \frac{1}{2}(\pi_1^2\sigma_1(t)^2F_1^2 + \pi_2^2\sigma_2(t)^2F_2^2 + \rho_{12}(t)\pi_1\pi_2\sigma_1(t)\sigma_2(t)F_1F_2)u_{ww} + \mu_1(t)F_1u_1 + \mu_2(t)F_2u_2 \\ &\quad + \frac{\sigma_1(t)^2}{2}F_1^2u_{11} + \frac{\sigma_2(t)^2}{2}F_2^2u_{22} + \rho_{12}(t)\sigma_1(t)\sigma_2(t)F_1F_2u_{12}\end{aligned}$$

The HJB equation can be written compactly as

$$u_t + \sup_{\pi_1,\pi_2}\mathcal{L}^{(\pi_1,\pi_2)}u = 0$$

Next, we apply the transformation

$$u(t,w,F_1,F_2) = -e^{-\gamma w - \Phi(t,f_1,f_2)}, \tag{4.34}$$

with $f_1 = \log F_1$ and $f_2 = \log F_2$. Substituting (4.34) into (4.33), we obtain the *linear* PDE for Φ:

$$\begin{aligned}0 = \Phi_t &+ \left(\frac{1}{2}\frac{\mu_1^2}{(1-\rho_{12}^2)\sigma_1^2} + \frac{1}{2}\frac{\mu_2^2}{(1-\rho_{12}^2)\sigma_2^2} - \frac{\rho_{12}\mu_1\mu_2}{(1-\rho_{12}^2)\sigma_1\sigma_2}\right) \\ &+ \frac{\sigma_1^2}{2}(\Phi_{11} - \Phi_1) + \frac{\sigma_2^2}{2}(\Phi_{22} - \Phi_2) + \rho_{12}\sigma_1\sigma_2\Phi_{12},\end{aligned} \tag{4.35}$$

with $\Phi(T, f_1, f_2)$=0. We have defined the partial derivatives

$$\Phi_t = \frac{\partial\Phi}{\partial t}, \quad \Phi_1 = \frac{\partial\Phi}{\partial f_1}, \quad \Phi_2 = \frac{\partial\Phi}{\partial f_2},$$

$$\Phi_{11} = \frac{\partial^2\Phi}{\partial f_1^2}, \quad \Phi_{22} = \frac{\partial^2\Phi}{\partial f_2^2}, \quad \Phi_{12} = \frac{\partial^2\Phi}{\partial f_1\partial f_2},$$

and suppressed the dependence on t, in μ_i, σ_i, and ρ_{12} to simplify the notation.

We can solve this linear PDE of Φ by using the ansatz

$$\Phi(t,f_1,f_2) = a_{11}(t)f_1^2 + a_1(t)f_1 + a_{22}(t)f_2^2 + a_2(t)f_2 + a_{12}(t)f_1f_2 + a(t)$$

to deduce that

$$a_{11}'(t) = a_{22}'(t) = a_{12}'(t) = 0,\ a_{11}(t) = a_{22}(t) = a_{12}(t) = 0,$$
$$a_1'(t) = a_2'(t) = 0,\ a_1(t) = a_2(t) = 0.$$

From this, we deduce that Φ is in fact a function of t only, independent of f_1 and f_2, and satisfies the first-order differential equation

$$\frac{d\Phi}{dt} = -\frac{\mu_1(t)^2\sigma_2(t)^2 + \mu_2(t)^2\sigma_1(t)^2 - 2\rho_{12}(t)\mu_1(t)\mu_2(t)\sigma_1(t)\sigma_2(t)}{2(1-\rho_{12}(t)^2)\sigma_1(t)^2\sigma_2(t)^2}.$$

Solving this and applying (4.14), (4.15), and (4.17), we obtain a closed-form expression for Φ. Precisely,

$$\Phi(t) = \frac{(T-t)\left((r-\mu)^2\bar{\eta}^2 + 2\lambda\,(r-\mu)\,\rho\bar{\eta}\eta + \lambda^2\,\eta^2\right)}{2\,(1-\rho^2)\,\bar{\eta}^2\eta^2}. \tag{4.36}$$

Applying (4.36) to (4.34), the candidate solution is given by

$$u(t,w) = -e^{-\gamma w - \Phi(t)}. \tag{4.37}$$

Interestingly, as in the single-futures case, the value function is independent of the speed of mean reversion κ and equilibrium level α of the convenience yield process. Intuitively, it suggests that the optimal strategy effectively removes the stochasticity of the convenience yield in the investor's maximum expected utility. This feature is evident again later in the characterization of the optimal wealth process. Moreover, the value function does not depend on the current futures prices (F_1, F_2). The simplicity of the value function is unexpected, especially since there are two stochastic factors and two futures in the trading problem. Nevertheless, it does not mean that the corresponding trading strategies are trivial. In fact, the strategies depend not only on other model parameters but also the futures prices, as we will discuss next.

By applying (4.34) and (4.36) to (4.33), we obtain the optimal trading strategies

$$\pi_1^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_1(t)F_1}\left(\frac{\mu_1(t)}{\sigma_1(t)} - \rho_{12}(t)\frac{\mu_2(t)}{\sigma_2(t)}\right), \tag{4.38}$$

$$\pi_2^*(t, F_1, F_2) = \frac{1}{\gamma(1-\rho_{12}(t)^2)\sigma_2(t)F_2}\left(\frac{\mu_2(t)}{\sigma_2(t)} - \rho_{12}(t)\frac{\mu_1(t)}{\sigma_1(t)}\right). \tag{4.39}$$

In this case with two futures, for either $i = 1, 2$, the corresponding optimal strategy π_i^* is a function of F_i, but does not depend on the price of the other futures F_j, for $i \neq j$.

We recall (4.14), (4.15), and (4.17), and express the optimal strategies explicitly in terms of model parameters. Precisely,

$$\pi_1^* = -\frac{e^{\kappa(T_1-t)}\left(\left(e^{t\kappa} - e^{\kappa T_2}\right)(r-\mu)\,\bar{\eta}^2 + \left(e^{t\kappa}\lambda + e^{\kappa T_2}\left(r\kappa - \lambda - \kappa\mu\right)\right)\rho\bar{\eta}\eta + e^{\kappa T_2}\kappa\lambda\eta^2\right)}{F_1\left(e^{\kappa T_1} - e^{\kappa T_2}\right)\gamma\left(1-\rho^2\right)\bar{\eta}^2\eta^2}, \tag{4.40}$$

$$\pi_2^* = \frac{e^{\kappa(T_2-t)}\left(\left(e^{t\kappa} - e^{\kappa T_1}\right)(r-\mu)\,\bar{\eta}^2 + \left(e^{t\kappa}\lambda + e^{\kappa T_1}\left(r\kappa - \lambda - \kappa\mu\right)\right)\rho\bar{\eta}\eta + e^{\kappa T_1}\kappa\lambda\eta^2\right)}{F_2\left(e^{\kappa T_1} - e^{\kappa T_2}\right)\gamma\left(1-\rho^2\right)\bar{\eta}^2\eta^2}. \tag{4.41}$$

Thus we see that the optimal controls π_1^* and π_2^* do not depend on the current spot price S_t or convenience yield δ_t, and is also independent on the equilibrium of the convenience yield α. For practical applications, this independence removes the burden to estimate or continuously monitor the spot price or convenience yield. Nevertheless, the optimal controls do depend on all the other parameters, namely $\mu, r, \kappa, \eta, \bar{\eta}, \rho,$ and λ.

Remark 2. *Naturally, one can consider trading futures with more than two maturities. However, in such case under the Schwartz two-factor model, there is an infinite number of solutions to the corresponding utility maximization problem and the additional futures are redundant, since we can replicate a third contract with two. To this end, remember that the T_i-futures price evolves according to the SDE*

$$\frac{dF_t^{(i)}}{F_t^{(i)}} = \mu_i(t)dt + \eta dZ_t^s + \bar{\eta}B_i(t)dZ_t^\delta,$$

under the physical measure $\mathbb{P}$, where

$$B_i(t) = -\frac{1-e^{-\kappa(T_i-t)}}{\kappa}$$

and

$$\mu_i(t) = \mu - r - \frac{\lambda(1-e^{-\kappa(T_i-t)})}{\kappa} = \mu - r + \lambda B_i(t).$$

We seek to find π_1 and π_2 such that

$$\begin{bmatrix} \mu_1 F_t^{(1)} & \mu_2 F_t^{(2)} \\ \eta F_t^{(1)} & \eta F_t^{(2)} \\ \bar{\eta}B_1F_t^{(1)} & \bar{\eta}B_2F_t^{(2)} \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} = \begin{bmatrix} \mu_3 F_t^{(3)} \\ \eta F_t^{(3)} \\ \bar{\eta}B_3F_t^{(3)} \end{bmatrix}. \tag{4.42}$$

Noting that $\mu_i(t)$ can be written as

$$\mu_i(t) = \frac{\mu - r}{\eta}\eta + \frac{\lambda}{\bar{\eta}}\bar{\eta}B_i,$$

we can see that the first equation in (4.42) is redundant and is a linear combination of the second and third, and therefore the system of equations in (4.42) has a unique solution for $\pi_i, i = 1, 2$.

Optimal Wealth Process

To derive the optimal wealth process, we substitute the optimal futures positions, π_1^* and π_2^*, into the wealth equation (4.30) and get

$$\begin{aligned} dW_t &= \pi_1^* dF_t^{(1)} + \pi_2^* dF_t^{(2)} \\ &= \mu_W dt + (\pi_1^* F_t^{(1)} + \pi_2^* F_t^{(2)})\eta dZ_t^s + (\pi_1^* F_t^{(1)} B_1(t) + \pi_2^* F_t^{(2)} B_2(t))\bar{\eta} dZ_t^\delta \\ &\equiv \mu_W dt + \sigma_W dZ_t^W, \end{aligned}$$

where we have defined

$$\begin{aligned} \mu_W &= \pi_1^* F_t^{(1)} \mu_1(t) + \pi_2^* F_t^{(2)} \mu_2(t) \\ &= \frac{(r-\mu)^2 \bar{\eta}^2 + 2\lambda (r-\mu) \rho \bar{\eta} \eta + \lambda^2 \eta^2}{\gamma (1-\rho^2) \bar{\eta}^2 \eta^2} \end{aligned} \tag{4.43}$$

and

$$\begin{aligned} \sigma_W^2 &= (\pi_1^* F_t^{(1)} + \pi_2^* F_t^{(2)})^2 \eta^2 + (\pi_1^* F_t^{(1)} B_1(t) + \pi_2^* F_t^{(2)} B_2(t))^2 \bar{\eta}^2 \\ &\quad + 2\rho\eta\bar{\eta}(\pi_1^* F_t^{(1)} + \pi_2^* F_t^{(2)})(\pi_1^* F_t^{(1)} B_1(t) + \pi_2^* F_t^{(2)} B_2(t)) \\ &= \frac{(r-\mu)^2 \bar{\eta}^2 + 2\lambda (r-\mu) \rho \bar{\eta} \eta + \lambda^2 \eta^2}{\gamma^2 (1-\rho^2) \bar{\eta}^2 \eta^2} \\ &= \frac{\mu_W}{\gamma}. \end{aligned} \tag{4.44}$$

In (4.43) and (4.44) we have used (4.40) and (4.41).

Note that both μ_W and σ_W are constant. This implies that the wealth process, under the optimal trading strategy, is an arithmetic Brownian motion with constant drift and volatility. Moreover, these two constants do not depend on the speed of mean reversion κ and equilibrium level α of the convenience yield process. This is why the value function is also independent of these two parameters. The financial intuition is that the optimal strategy suggests trading in a way that removes the randomness stemmed from the convenience yield process. As a special case, when $\mu = r$ and $\lambda = 0$, the $\mathbb{P}$ measure is identical to $\mathbb{Q}$. This will lead to $\pi_i^* = 0, i = 1, 2$, and in turn a constant wealth, with $\mu_W = \sigma_W = 0$.

4.2.3 Verification Theorem

Based on the verification theorem in Appendix A, Theorem 7, in order to show that the candidate solution is the value function, it suffices to prove the uniform integrability of the family of random variables $\{u(\tau, W_\tau^*)\}_\tau$ for any stopping time $\tau \in [0, T]$. We will only present the 2 futures case since the 1 maturity case follows the same argument.

Theorem 4. *The candidate solution found in (4.37) is equal to the value function (4.32); namely,*

$$V(t, w) = u(t, w) = -\exp(-\gamma w - \Phi(t))$$

on $(t, w) \in [0, T] \times \mathbb{R}$, where $\Phi(t)$ is given by (4.36). Furthermore, the optimal control pair is given by (4.40) and (4.41).

Proof. Let $\epsilon > 0$ be an arbitrary positive number. From the verification theorem in Appendix A, it suffices to show that $\mathbb{E}(|u(\tau, W_\tau^*)|^{1+\epsilon}) < \infty$, uniformly with respect to any stopping time τ with $0 \leq \tau \leq T$. First, after applying the Cauchy-Schwarz inequality, we find that

$$\mathbb{E}(|u(\tau, W_\tau^*)|^{1+\epsilon}) \leq \mathbb{E}(\exp(-2(1+\epsilon)\gamma W_\tau^*))^{1/2} \times \tag{4.45}$$

$$\mathbb{E}(\exp(-2(1+\epsilon)\Phi(\tau)))^{1/2}. \tag{4.46}$$

For the first expectation (4.45), clearly

$$\begin{aligned}
&\mathbb{E}[\exp(-2(1+\epsilon)\gamma W_\tau^*)] \\
&\quad = \mathbb{E}\left[\exp\left(-2(1+\epsilon)\left(\int_0^\tau \mu_W ds + \int_0^\tau \sigma_W dZ_s^W\right)\right)\right] \\
&\quad = \mathbb{E}\left[\exp\left(-2(1+\epsilon)\left(\mu_W \tau + \sigma_W Z_\tau^W\right)\right)\right] \\
&\quad \leq c\mathbb{E}\left[\exp\left(-2(1+\epsilon)\sigma_W Z_\tau^W\right)\right]
\end{aligned} \tag{4.47}$$

for some constant c, since μ_W is a constant and $\tau \leq T$.

Now since Z_t^W is a martingale, we have that for any constant k, $\exp(kZ_t^W)$ is a submartingale by Jensen's inequality, since

$$\mathbb{E}_0(\exp(kZ_t^W)) \geq \exp(k\mathbb{E}_0(Z_t^W)) = \exp(k\mathbb{E}_0(Z_0^W)) = 1.$$

Moreover, $\exp(kZ_t^W)$ is positive. Therefore we can use Doob's martingale inequality [103]: for $\xi > 0$,

$$\begin{aligned}
\mathbb{E}[\exp(kZ_\tau^W)] &= \left\| \exp\left(\frac{k}{1+\xi} Z_\tau^W\right) \right\|_{1+\xi}^{1+\xi} \\
&\leq \left\| \sup_{0\leq t\leq T} \exp\left(\frac{k}{1+\xi} Z_t^W\right) \right\|_{1+\xi}^{1+\xi} \\
&\leq \left(1 + \frac{1}{\xi}\right)^{1+\xi} \sup_{0\leq t\leq T} \mathbb{E}[\exp(kZ_t^W)] \\
&\leq c
\end{aligned}$$

where c is another positive constant, independent of the stopping time τ.

For the second expectation (4.46), recalling that

$$\Phi(t) = \frac{(T-t)\left((r-\mu)^2 \bar{\eta}^2 + 2\lambda(r-\mu)\rho\bar{\eta}\eta + \lambda^2\eta^2\right)}{2(1-\rho^2)\bar{\eta}^2\eta^2},$$

we can clearly see that $\mathbb{E}(\exp(-2(1+\epsilon)\Phi(\tau)))$ is bounded as well. Hence we proved that $\{u(\tau, W_\tau^*)\}_\tau$ is uniformly integrable for any stopping time $\tau \in [0, T]$. $\square$

Since the uniform integrability for $\{\tilde{u}(\tau, W_\tau^*, F_\tau^{(1)})\}_\tau$ in the 1 futures case can be proved using the exact same argument, we can conclude as well that the candidate solution in the 1 futures case (4.26) is the value function (4.22).

4.2.4 Certainty Equivalent

Next, we consider the *certainty equivalent* associated with the trading opportunity in the futures. The certainty equivalent is the cash amount that derives the same utility as the value function. First, we consider the single-futures case. Recall from (4.21) and (4.26) that the investor's utility and value functions are both of exponential form. Therefore, the certainty equivalent is given by

$$\tilde{C}^{(i)}(t,w) \equiv U^{-1}(\tilde{u}(t,w)) = w + \frac{\tilde{\Phi}^{(i)}(t)}{\gamma}. \tag{4.48}$$

Here, the superscript (i) refers to the futures with maturity T_i in the portfolio. From (4.48), we observe that the certainty equivalent is the sum of the investor's wealth w and the time-deterministic component $\tilde{\Phi}^{(i)}(t)/\gamma$, which is positive and inversely proportional to the risk aversion parameter γ. All else being equal, a more risk averse investor has a lower certainty equivalent, valuing the futures trading opportunity less. Interestingly, the certainty equivalent does not depend on the current futures prices F_1 but it does depend on the model parameters that appear in the futures price dynamics.

Similarly, the certainty equivalent from dynamically trading two futures with different maturities is given by

$$C(t,w) \equiv U^{-1}(u(t,w)) = w + \frac{\Phi(t)}{\gamma}, \tag{4.49}$$

where $u(t,w)$ is the value function in (4.37) and Φ is given by (4.36).

Since the certainty equivalents in both the single-futures and two-futures cases have the same linear dependence on wealth w, we will for simplicity set $w = 0$ in our numerical examples to compare across these cases. To this end, we denote $\tilde{C}_0^{(i)}(t) \equiv \tilde{C}^{(i)}(t,0)$ and $C_0(t) \equiv C(t,0)$.

4.3 Numerical Implementation

We now examine our model through a number of numerical examples using simulated and empirical data. For our examples, we will use the estimated parameters values found in [54]. They are displayed here in Table 4.1. The drift parameter μ of the spot price was not given in [54], so we set $\mu = 1\%$ for our examples. We use federal funds rate as a proxy for the instantaneous interest rate r which, during the calibration period, hovered around 0.1%.[1] The default value for the risk aversion coefficient γ is 1% unless noted otherwise.

μ	κ	η	$\bar{\eta}$	ρ	λ	r
0.010	0.800	0.450	0.500	0.750	0.050	0.001

Table 4.1: The Schwartz model parameters estimated by Ewald, Zhang, and Zong.

In Figure 4.1, we show the dependence of the optimal positions, π_1^* and π_2^*, respectively in the T_1-futures and T_2-futures in the two-futures case on the volatility parameter $\bar{\eta}$ of the convenience yield process, for three different risk aversion levels. Observe that π_1^* at all three levels of γ is positive and decreasing in $\bar{\eta}$ while π_2^* is negative and increasing in $\bar{\eta}$. With the parameters given in Table 4.1, we are long the T_1-futures $F^{(1)}$ and short the T_2-futures $F^{(2)}$. When we rearrange the formulae (4.40) and (4.41) for π_1^* and π_2^*, respectively, and collect terms involving $\bar{\eta}$, we see that for both $i = 1, 2$, the optimal strategies are of the form $A_i + B_i/\bar{\eta} + C_i/\bar{\eta}^2$, which means that the absolute value of each strategy π_i^* decreases as $\bar{\eta}$ increases, with other variables held constant. The practical consequence is that the number of contracts held, on both the long and short sides, are decreasing as the volatility of the stochastic convenience yield process δ_t increases. This is in line with a risk-averse trader's intuition that less exposures on both legs of the traded pair should be preferred, if the volatility of the stochastic convenience yield is high. Furthermore, the positions increase in size (more positive for π_1^* and more negative for π_2^*) as risk aversion decreases. This is obvious given the inverse relationship between γ and π_i^* as seen in Eq (4.38) and (4.39).

[1] Data from www.macrotrends.net.

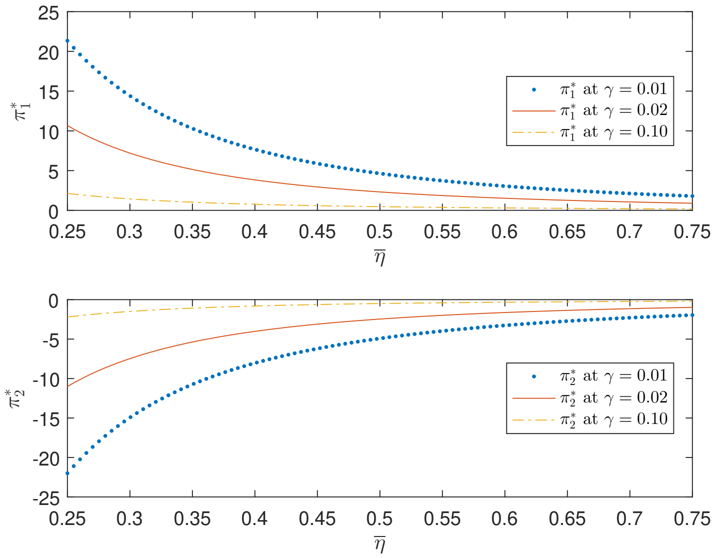


Figure 4.1: Optimal positions, π_1^* and π_2^*, respectively in the T_1-futures and T_2-futures in the two-futures case plotted for $\overline{\eta} \in [0.25, 0.75]$, at three levels of risk aversion γ. Common parameters are displayed in Table 4.1, with $F_1 = 100$ and $F_2 = 100$.

Figure 4.2 illustrates how the optimal futures positions, π_1^* and π_2^*, vary with respect to maturity. First of all, the two positions are of different signs and their sizes are very close. As maturity T_1 or T_2 lengthens, the size of the corresponding futures position increases, with π_1^* becoming more positive and π_2^* more negative. However, the change is very small as the scale on the y-axis shows, so one can interpret this as the positions are not very sensitive to the futures maturities.

In Figure 4.3 we compare the optimal trading strategies, π_1^* and π_2^* for two futures to the optimal strategy $\tilde{\pi}_i^*$ for trading a single futures. We plot the strategies as functions of η, the volatility of the spot price, using same set of parameters as in Table 4.1. When trading a single contract, the corresponding optimal strategy, $\tilde{\pi}_1^*$ and $\tilde{\pi}_2^*$, are both very small near zero. However, it can be seen that they do increase slightly in size when η becomes small, as volatility decreases.

This is in contrast to the two-contract case where the optimal strategies are π_1^* and π_2^*. Both increases, in opposite directions, as η increases. This shows that despite the increase in risk as η increases, paired positions in π_1^* and π_2^*, of opposite signs, will increase as volatility of the spot process increases.

It is also interesting to note the size of the positions in the single contract cases as compared to the pair-trading case. When we are constrained to trade only single contracts, that is when the admissible set is $\tilde{\mathcal{A}}_t$ as opposed to $\mathcal{A}_t$, the position is much smaller. Under the current model, the presence of multiple contracts of different maturities significantly increases trade volume and allows the trader to take much bigger hedged trades.

In Figure 4.4 we plot the optimal strategies as functions of γ, the risk aversion coefficient. Obviously, given the inverse relationship between γ and π_i^* as seen in Eq (4.38) and (4.39), as well as between γ and $\tilde{\pi}_i^*$ as seen in Eq (4.28), the optimal positions are expected to decrease in magnitude. What is interesting to note is the insensitivity of $\tilde{\pi}_i^*$ with respect to γ, in comparison to π_i^*. This means that in the single futures case, the position will be small regardless of the level of risk aversion.

Having analyzed the parameter dependence of the optimal strategies in details, now we turn to their path behavior based on historical data. We consider the June 2014 and July 2014 WTI crude oil futures. We show the empirical optimal positions over the period March 2014 to June 2014. This period is chosen to correspond to the post-calibration period of [54]. Applying our explicit formulae for the strategies, we compute π_1^*, π_2^*, and $\pi_1^* + \pi_2^*$ based on the daily settlement prices of these contracts as well as the parameters in Table 4.1. As shown in Figure 4.5, the optimal strategy π_1^* is positive throughout this

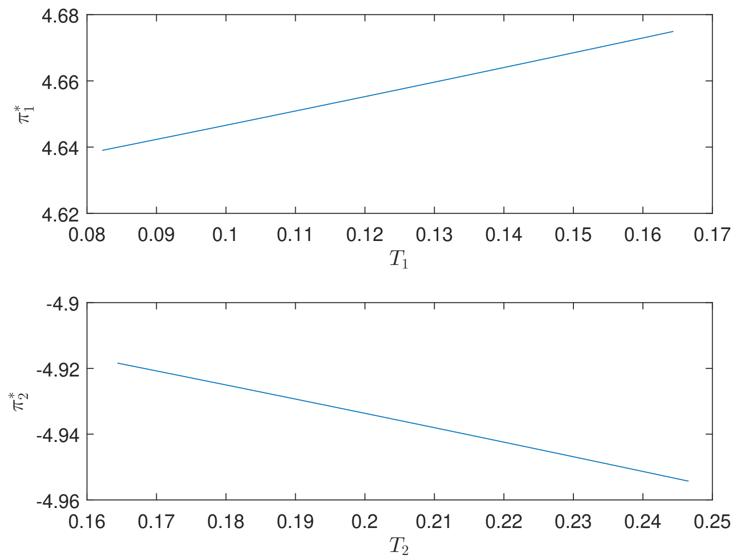


Figure 4.2: Optimal positions, π_1^* in the T_1-futures and π_2^* in the T_2-futures in the two-futures portfolio, plotted as a function of T_1 and T_2 respectively, with parameters as displayed in Table 4.1, and $F_1 = 100$ and $F_2 = 100$.

period, corresponding to a long position in the front-month contract, and the opposite holds for π_2^*. Taken together, the sum of both positions is negligibly small, corresponding to a net neutral position. Overall, the positions changed little when the parameters η and $\bar{\eta}$ are kept fixed. The only variables that change are F_i and $T_i - t$, of which we have already seen the relative insensitivity in Figure 4.2.

We now turn our attention to the certainty equivalents. With reference to Section 4.2.4, we plot in Figure 4.6 the following certainty equivalents: $\tilde{C}^{(1)}$ in the single-futures case with T_1-futures traded, $\tilde{C}^{(2)}$ in the single-futures case with T_2-futures traded, and C in the two-futures case with T_1-futures and T_2 futures traded. Their numerical values are given in Table 4.2.

$C_0(0)$	$\tilde{C}_0^{(1)}(0)$	$\tilde{C}_0^{(2)}(0)$
0.8962	0.1418	0.1782

Table 4.2: Values of certainty equivalent: $\tilde{C}^{(1)}$ in the single-futures case with T_1-futures traded, $\tilde{C}^{(2)}$ in the single-futures case with T_2-futures traded, and C in the two-futures case with T_1-futures and T_2 futures traded. The certainty equivalents are evaluated at $t = 0$ and $w = 0$.

We observe from Figure 4.6 that the certainty equivalent for trading two contracts simultaneously is significantly greater than that derived from trading only a single contract regardless of the choice of maturity. In fact, the certainty equivalent C is much larger than the sum of the two certainty equivalents $\tilde{C}^{(1)}$ and $\tilde{C}^{(2)}$, or $2\,\tilde{C}^{(i)}$ for both $i = 1, 2$, as seen in Table 4.2. This makes sense since the single-contract case can be viewed as two-contracts case but with one strategy constrained at zero. Effectively, the single-contract case is restricting the admissible set from $\mathcal{A}_t$ to $\tilde{\mathcal{A}}_t$, thus reducing the maximum expected utility as well as the certainty equivalent. Our result confirms the intuition that more choices of trading instruments are preferable to fewer.

Lastly, we examine the behavior of C at different risk aversion levels with focus on its sensitivity with respect to the market price of risk λ. In Figure 4.7, we see that the certainty equivalent at time 0, C_0, is increasing and quadratic in λ, and tends to infinity as λ increases. This holds for all three values of γ

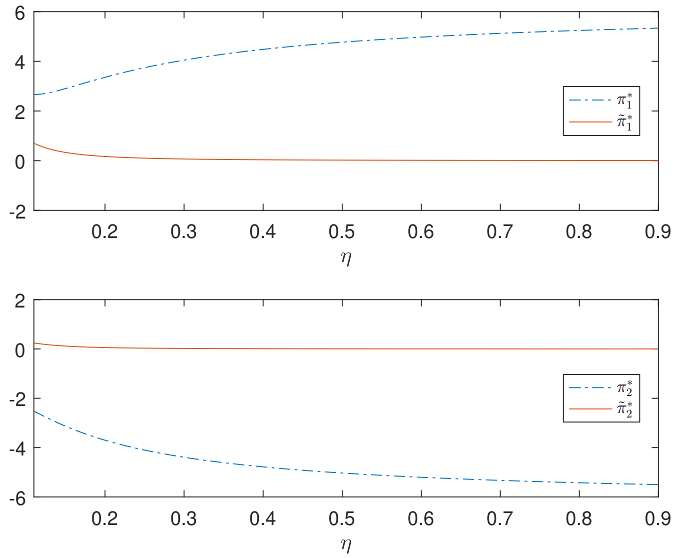


Figure 4.3: Optimal futures position π_i^* (dashed) in the 2-contract portfolio and $\tilde{\pi}_i^*$ (solid) in the single-contract portfolio (with the T_i-futures) plotted over $\eta \in [0.1, 0.9]$. Parameters are taken from Table 4.1, with $F_1 = 100$ and $F_2 = 100$.

shown, but a lower risk aversion suggests that the certainty equivalent is higher and faster growing in λ.

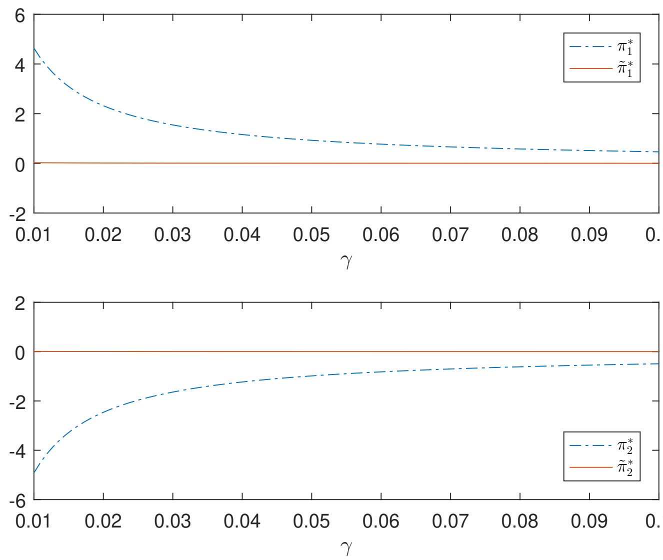


Figure 4.4: Optimal futures position π_i^* (dashed) in the 2-contract portfolio and $\tilde{\pi}_i^*$ (solid) in the single-contract portfolio (with the T_i-futures) plotted over $\gamma \in [0.01, 0.1]$. Parameters are taken from Table 4.1, with $F_1 = 100$ and $F_2 = 100$.

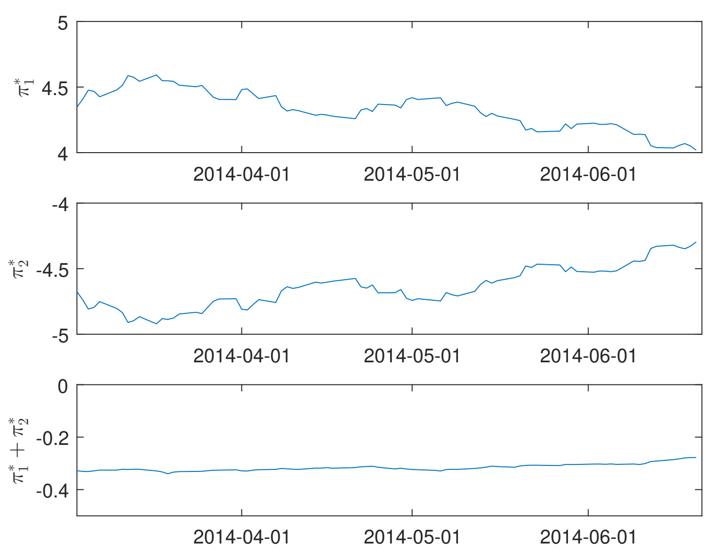


Figure 4.5: Optimal strategies π_1^*, π_2^* and $\pi_1^* + \pi_2^*$ based on historical WTI crude oil futures data over the period Mar 2014 - Jun 2014 using parameters as displayed in Table 4.1.

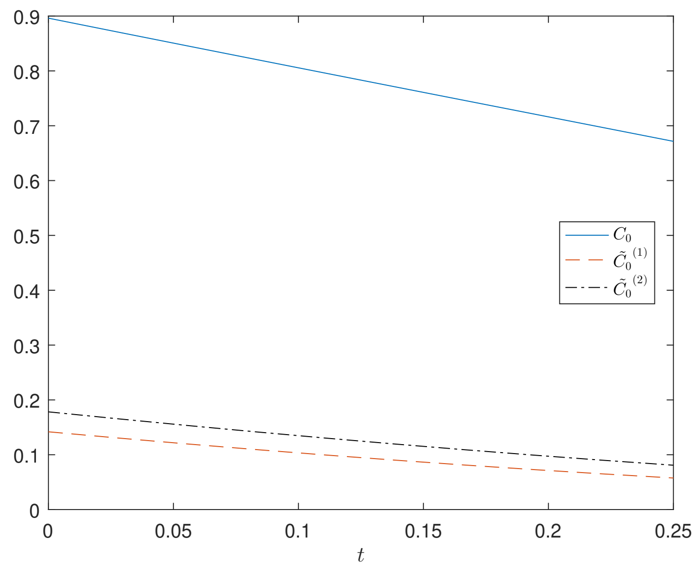


Figure 4.6: The certainty equivalents C_0 for the two-futures portfolio, as well as $\tilde{C}_0^{(1)}$ and $\tilde{C}_0^{(2)}$ for the single-futures portfolios, respectively with T_1-futures and T_2-futures (see (4.48)). The certainty equivalents are evaluated at time $t = 0$ with initial wealth $w = 0$. The trading horizon is $T = 1$, maturity of F_1 is $T_1 = 13/12$, and maturity of $T_2 = 14/12$. Other common parameters are from Table 4.1, along with $F_1 = 100$ and $F_2 = 100$.

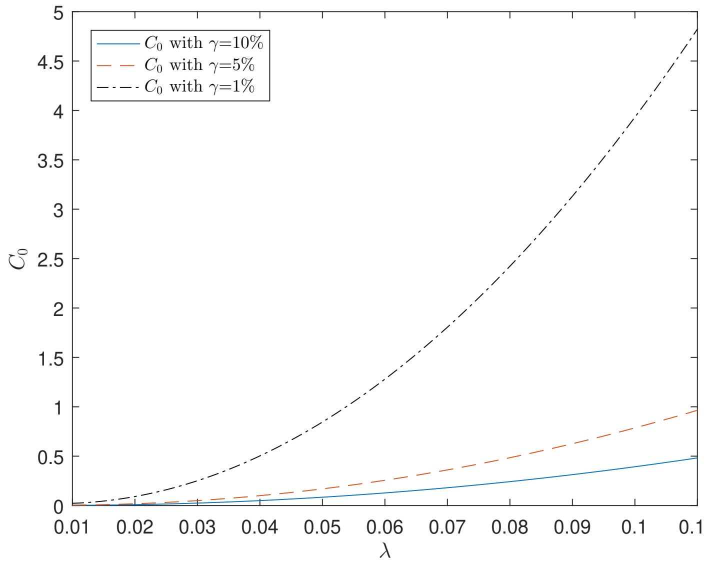


Figure 4.7: Certainty equivalent C_0, at time $t = 0$ with zero initial wealth $W_0 = 0$, as a function of the market price of risk λ, with parameters as displayed in Table 4.1.

4.4 Summary

We have analyzed the problem of dynamically trading two futures contracts with the same underlying. Under a two-factor mean-reverting model for the spot price, we derive the futures price dynamics and solve the portfolio optimization problem in closed form and give explicit optimal trading strategies. By studying the associated HJB equation, we solve the utility maximization explicitly and provide the optimal trading strategies in closed form. The optimized wealth process is again shown to possess a positive drift, suggesting it is profitable to follow the optimal strategies derived here. In addition to the analytic properties of our solutions, we also apply our results to commodity futures trading and present numerical examples to illustrate the optimal holdings.

This chapter provides a feasible strategy to quantitatively trade commodities futures, in which a convenience yield factor is present, in contrast to the model in the previous chapter. The formula is readily implementable, and the main reason for easy implementation is the independence of the optimal strategy on any hidden state variables. On the other hand, for models which lead to trading strategies that depend on hidden variables, it will be necessary to rely on the filtered state estimates that result from the Kalman filter methodology. Another possibility is to formulate the control problem as described in [17] and to solve stochastic partial differential equations (SPDE).

The simplicity of the trading strategy greatly facilitates actual deployment, for example by managed futures fund managers, which is the focus in this chapter here. Although this type of asset management strategies is opaque, it is safe to assume that they involve trading more than one commodity. There exists an extensive literature on the complicated relationships between closely-related commodities, and on their spread trading: for example the so-called *crush spread* which consists of soybean, soybean meal and soybean oil [110] in the agricultural space, or the so-called crack spread which consists of crude oil and related refinery products [55] in the energy space. Optimal trading in these related commodities are important topics for further research.

The main criticism of the well-known Schwartz model studied in this chapter is the poor fit of two-factor models to observed term structures. However, we will remedy the shortcomings of two-factor models when we consider the Cortazar and Naranjo's n-Factor Model [37] in the next chapter.

Chapter 5

Generalization to Cortazar and Naranjo's n-Factor Model

In the preceding three chapters we have been working exclusively with models with two factors. Naturally, the next step is to extend in a tractable way to an arbitrary n-factor model. In the two stocks pairs trading setting, as originated in [116], the generalization to a larger number of stocks was carried out in [87] and [29], and more recently in [65]. Thus in this chapter we solve for the optimal trading strategies in the commodity futures market under a multi-factor model, which was proposed by [37].

While the factors in this type of multi-factor models are not easily interpreted to have economic meaning, the flexibility of multi-factor models permits good fit to empirical term structure as displayed in the market. For example, in the Schwartz model the factors are the spot price process and the stochastic convenience yield process, which have clear economic meaning. However, the inadequate fit of two-factor models, as for example observed in [25], overwhelms the advantage of interpretability. Especially in deep and liquid futures markets such as crude oil or gold, with more than 10 contracts of various maturities actively traded at any given time, the lack of fit render other uses, such as derivative pricing, forecasting, or trading, which is the focus here, of such simple two-factor models impossible. Trading strategies are arguably less reliable if the model prices from which the strategies are derived cannot even match observed market prices, or if model implied volatility and correlation term structures cannot match that implied by historical market prices.

In addition to flexibility, another advantage of the Cortazar and Naranjo model is the ease of estimation, via Kalman filtering, as described in full details in [37]. As used in the Schwartz paper [107], Kalman filtering methodology can handle multi-factor models with hidden state variables and measurement errors. Inclusion of measurement errors is necessary since the number of available market prices is generally higher than the number of state variables that need to be estimated. In addition, Kalman filtering is capable of using a large price panel in the estimation process, avoiding the necessity of making an arbitrary selection of contracts to include in the estimation.

While many other approaches exist in the literature, we work with the Cortazar and Naranjo model, which leads to tractable formulae that will be given in later sections. Their sequel model [36] introduced stochastic volatility into the commodity price processes; while this feature increases the model complexity, it results in better fit to observed *option* prices. Other attempts to match market observables include [93], who allow various parameters (namely, volatility of spot price, speed of mean-reversion, the mean-reversion parameter, and the diffusion parameter of the spot convenience yield) to be time-varying deterministic functions. [73] assumes the commodity market operates in an economy with incomplete information, and derives closed-form solutions for forward and futures prices, as well as for European options on forward and futures contracts. In [3], a regime-switching model for crude oil and natural gas is proposed: one regime is found to have high convenience yield and high volatility, while the other is found to have low convenience yield and low volatility. For other nonlinear models of commodity prices which require more computationally intensive machine learning techniques to estimate, see for example [15] and references therein. Regarding the literature of optimal trading of commodity futures contracts, numerous references are given in [81] and [82].

We are primarily concerned with up to four factors in this chapter since, as shown in [37], four factors are sufficient to fit both price and volatility term structures. With only three factors, the model fits the term structure of futures prices, but four is required to fit the volatility term structure as well. We find that, as expected, the position sizes decrease as volatility increases. Moreover, the certainty equivalent increases as n increases from $n = 2$ to $n = 4$, meaning that the utility of trading, using the optimal strategies, is higher in markets with more maturities available. Finally, with $n = 4$ factors, we observe that trading any other combinations of futures maturities other than trading maximally all 4 available maturities results in lower certainty equivalents, which again means inclusion of more futures maturities in a portfolio is more economically beneficial.

5.1 Futures Price Dynamics

We now briefly review the model setup as described in [37]. Let $\mathbf{x}_t$ denote the n-dimensional vector of state variables $(x_t^{(1)}, \ldots, x_t^{(n)})$. The SDE for these n factors under $\mathbb{P}$ is

$$d\mathbf{x}_t = -\mathbf{K}\mathbf{x}_t dt + \Sigma d\mathbf{w}_t. \tag{5.1}$$

The $n \times 1$ vector $d\mathbf{w}_t$ is a vector of correlated Brownian motion increments

$$d\mathbf{w}_t = \left[dw_t^{(1)} \ \cdots \ dw_t^{(n)} \right]'$$

such that $d\mathbf{w}_t \cdot d\mathbf{w}_t' = \mathbf{\Omega} \, dt$, where the (i,j) element of the symmetric positive definite matrix $\mathbf{\Omega}$ is $\rho_{ij} \in [-1, 1]$, the instantaneous correlation, and the prime notation denotes the transpose of vectors and matrices.

The matrices $\mathbf{K}$ and Σ

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & k_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k_n \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}$$

are diagonal, and they consist of the constant speeds of mean-reversion, and constant volatility parameters, respectively. k_1 is exogenously set to zero so that the first state variable $x_t^{(1)}$ follows a random walk, which induces a unit root in the spot price process.

The $\mathbb{Q}$ dynamics, as in Eq. 3 in [37], is

$$d\mathbf{x}_t = -(\boldsymbol{\lambda} + \mathbf{K}\mathbf{x}_t)dt + \Sigma d\mathbf{w}_t^*, \tag{5.2}$$

where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)'$ is a $n \times 1$ vector of constant risk premiums, so that we can see the change of measure is effected through

$$\mathbf{w}_t^* = \mathbf{w}_t + \Sigma^{-1}\boldsymbol{\lambda}\, t. \tag{5.3}$$

The futures price, of maturity T, as a function of $\mathbf{x}$ and t, is given in [37] by

$$\begin{aligned} F(\mathbf{x}, t) = \exp\Bigg(& x_1 + \sum_{i=2}^{n} e^{-k_i(T-t)} x_i + \mu t + \left(\mu - \lambda_1 + \frac{\sigma_1^2}{2} \right)(T - t) \\ & - \sum_{i=2}^{n} \frac{1 - e^{-k_i(T-t)}}{k_i} \lambda_i + \frac{1}{2} \sum_{i*j \neq 1} \sigma_i \sigma_j \rho_{ij} \frac{1 - e^{-(k_i+k_j)(T-t)}}{k_i + k_j} \Bigg). \end{aligned} \tag{5.4}$$

The first derivatives are

$$\frac{\partial F}{\partial x_i} = F e^{-k_i(T-t)}, \tag{5.5}$$

for $i = 1, \dots, n$, but note that $k_1 = 0$. The second derivatives are

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = F e^{-(k_i+k_j)(T-t)}, \tag{5.6}$$

and the derivative with respect to time is

$$\begin{aligned} \frac{1}{F}\frac{\partial F}{\partial t} &= \sum_{i=2}^{n} e^{-k_i(T-t)} x_i k_i + \left(\lambda_1 - \frac{\sigma_1^2}{2} \right) \\ &\quad + \sum_{i=2}^{n} e^{-k_i(T-t)} \lambda_i - \frac{1}{2} \sum_{i*j \neq 1} \sigma_i \sigma_j \rho_{ij} e^{-(k_i+k_j)(T-t)} \\ &= \left(\lambda_1 - \frac{\sigma_1^2}{2} \right) + \sum_{i=2}^{n} e^{-k_i(T-t)} (x_i k_i + \lambda_i) - \frac{1}{2} \sum_{i*j \neq 1} \sigma_i \sigma_j \rho_{ij} e^{-(k_i+k_j)(T-t)}. \end{aligned} \tag{5.7}$$

We define the lower triangular matrix C with elements c_{ij} which result from Cholesky decomposition of the correlation matrix $\mathbf{\Omega} = CC'$; that is,

$$C = \begin{bmatrix} c_{11} & 0 & \dots & 0 \\ c_{21} & c_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}.$$

We also define $\tilde{\Sigma} \equiv \Sigma\, C$; hence, the SDE for $\mathbf{x}_t$ under $\mathbb{P}$ is

$$d\mathbf{x}_t = -\mathbf{K}\mathbf{x}_t dt + \tilde{\Sigma} d\mathbf{z}_t, \tag{5.8}$$

and under $\mathbb{Q}$ is

$$d\mathbf{x}_t = -(\boldsymbol{\lambda} + \mathbf{K}\mathbf{x}_t)dt + \tilde{\Sigma} d\mathbf{z}_t^*, \tag{5.9}$$

where the Brownian motion increment $\mathbf{z}_t$, with elements $z_t^{(i)}, i = 1, \ldots, n$, are now uncorrelated. Similarly, $\mathbf{z}_t^* = \mathbf{z}_t + \tilde{\Sigma}^{-1}\boldsymbol{\lambda}\, t$ is uncorrelated.

By Ito's Lemma, the SDE for the futures price process under $\mathbb{P}$, which is now denoted using subscript t as F_t as opposed to the function $F(\cdot, t)$, is

$$dF_t = \left(\frac{\partial F}{\partial t} - \nabla_{\mathbf{x}} F' \mathbf{K}\mathbf{x}_t + \frac{1}{2}\operatorname{Tr}(\tilde{\Sigma}' \nabla_{\mathbf{xx}} F \tilde{\Sigma})\right) dt + \nabla_{\mathbf{x}} F' \tilde{\Sigma} d\mathbf{z}_t, \tag{5.10}$$

where

$$\nabla_{\mathbf{x}} F = \begin{bmatrix} \frac{\partial F}{\partial x_1} & \cdots & \frac{\partial F}{\partial x_n} \end{bmatrix}', \qquad \nabla_{\mathbf{xx}} F = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 x_2} & \cdots & \frac{\partial^2 F}{\partial x_1 x_n} \\ \frac{\partial^2 F}{\partial x_1 x_2} & \frac{\partial^2 F}{\partial x_2^2} & \cdots & \frac{\partial^2 F}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_1 x_n} & \frac{\partial^2 F}{\partial x_2 x_n} & \cdots & \frac{\partial^2 F}{\partial x_n^2} \end{bmatrix}$$

are the gradient vector and the Hessian matrix respectively.

In particular, for the case where $n = 1$, we have

$$dF_t = \lambda_1 F_t dt + \sigma_1 F_t dz_t^{(1)}.$$

For $n = 2$, we have

$$\frac{dF_t}{F_t} = (\lambda_1 + \lambda_2 e^{-k_2(T-t)})dt + (\sigma_1 + e^{-k_2(T-t)}\rho_{12}\sigma_2)dz_t^{(1)} + e^{-k_2(T-t)}\sigma_2\sqrt{1-\rho_{12}^2}dz_t^{(2)}.$$

In general, for an arbitrary n, the SDE of the futures price process under $\mathbb{P}$ is

$$\frac{dF_t}{F_t} = \left(\sum_{i=1}^n \lambda_i e^{-k_i(T-t)}\right) dt + \sum_{i=1}^n \left(\sum_{j=i}^n e^{-k_j(T-t)} c_{ji}\sigma_j\right) dz_t^{(i)}. \tag{5.11}$$

We can readily see that the futures price SDE is driftless under $\mathbb{Q}$.

Now consider a collection of n contracts of different maturities available to trade, with n coinciding with the number of factors. We will denote by $F^{(k)}(\mathbf{x}, t)$ as the futures price function of the contract with maturity T_k, with $T_1 < \ldots < T_n$, and by $F_t^{(k)}$ as the stochastic process for this contract, for $k = 1, \ldots, n$. We have

$$\begin{aligned} \frac{dF_t^{(k)}}{F_t^{(k)}} &= \left(\sum_{i=1}^n \lambda_i e^{-k_i(T_k-t)}\right) dt + \sum_{i=1}^n \left(\sum_{j=i}^n e^{-k_j(T_k-t)} c_{ji}\sigma_j\right) dz_t^{(i)} \\ &\equiv \mu_k(t)dt + \sum_{i=1}^n v_{ki}(t) dz_t^{(i)}, \end{aligned}$$

where we define

$$\mu_k(t) \equiv \sum_{i=1}^n \lambda_i e^{-k_i(T_k-t)}, \tag{5.12}$$

and

$$v_{ki}(t) \equiv \sum_{j=i}^n e^{-k_j(T_k-t)} c_{ji}\sigma_j. \tag{5.13}$$

As usual, we define $f_t^{(k)} = \log F_t^{(k)}$ to be the log futures price process, so that

$$df_t^{(k)} = \left(\mu_k(t) - \frac{1}{2}\sum_{i=1}^{n} v_{ki}(t)^2\right)dt + \sum_{i=1}^{n} v_{ki}(t)dz_t^{(i)}.$$

In matrix notation,

$$d\boldsymbol{f}_t = \left(\boldsymbol{\mu}(t) - \frac{1}{2}D(\boldsymbol{v}(t)\boldsymbol{v}(t)')\right)dt + \boldsymbol{v}(t)d\boldsymbol{z}_t,$$

where $\boldsymbol{\mu}(t)$ is a $n \times 1$ vectors with elements $\mu_k(t)$, and $\boldsymbol{v}(t)$ is an $n \times n$ matrix with elements $v_{ki}(t)$, and $D(\boldsymbol{v}(t)\boldsymbol{v}(t)')$ is a $n \times 1$ vector whose elements are the diagonal of $\boldsymbol{v}(t)\boldsymbol{v}(t)'$.

We now let $\boldsymbol{\pi}(t,\boldsymbol{f}) = (\pi_1(t,\boldsymbol{f}),\ldots,\pi_n(t,\boldsymbol{f}))'$, where the elements $\pi_k(t,\boldsymbol{f})$ denote the amount of money invested in $F_t^{(k)}$. The wealth process is

$$dW_t = \sum_{k=1}^{n} \pi_k(t,\boldsymbol{f})\frac{dF_t^{(k)}}{F_t^{(k)}}, \tag{5.14}$$

and in matrix form, the system of variables is given by the set of SDE

$$\begin{bmatrix} dW_t \\ d\boldsymbol{f}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\pi}'\boldsymbol{\mu}(t) \\ \boldsymbol{\mu}(t) - \frac{1}{2}D(\boldsymbol{v}(t)\boldsymbol{v}(t)') \end{bmatrix} dt + \begin{bmatrix} \boldsymbol{\pi}'\boldsymbol{v}(t) \\ \boldsymbol{v}(t) \end{bmatrix} d\boldsymbol{z}_t. \tag{5.15}$$

From (5.15), it can be seen that the control problem under the n-factor model include the 2-factor models considered in previous chapters as special cases, due to the drift and volatility functions' sole dependence on time, and not on any of the state variables. This is in contrast to the situation in Chapter 2, where the drift in the wealth process depends on $(S_t^{(1)}, S_t^{(2)})$ via the cointegrating vector.

5.2 Utility Maximization Problem

We will work with the exponential utility function $U(w) \equiv -\exp(-\gamma w)$, with $\gamma > 0$ as the coefficient of risk aversion, as before. The trader fixes a finite optimization horizon $0 < T \leq T_1$ (which means that T has to be less than the maturity of the earliest expiring contract), and seeks to maximize the expected utility of wealth at T. We now define a vector $\boldsymbol{\pi}(t,\boldsymbol{f})$ of controls to be admissible if the elements of $\boldsymbol{\pi}(t,\boldsymbol{f})$ are real-valued, progressively measurable, and satisfies the integrability condition

$$\mathbb{E}\int_t^T \sum_{i=1}^{n} \pi_i(s,\boldsymbol{f})^2 ds < \infty.$$

The set of admissible controls at the initial time t is denoted by $\mathcal{A}_t$. Next, we define the value function $V(t,w,\boldsymbol{f})$ of the following stochastic control problem: the trader seeks an admissible strategy $\boldsymbol{\pi}(t,\boldsymbol{f}) \in \mathcal{A}_t$ that maximizes the utility of terminal wealth at time T; that is,

$$V(t,w,\boldsymbol{f}) = \sup_{\boldsymbol{\pi}\in\mathcal{A}_t} \mathbb{E}[U(W_T)]. \tag{5.16}$$

Following the standard verification approach to dynamic programming [58, 105, 96], we assume the existence of a sufficiently smooth candidate solution $u(t,w,\boldsymbol{f})$, which will later be shown to be equal to the value function V in (5.16).

To facilitate presentation, we suppress the dependence on t in $\boldsymbol{\mu}(t)$ and $\boldsymbol{v}(t)$, and the dependence on t and $\boldsymbol{f}$ in $\boldsymbol{\pi}(t,\boldsymbol{f})$, and define

$$u_t = \frac{\partial u}{\partial t}, \quad u_w = \frac{\partial u}{\partial w}, \quad u_{ww} = \frac{\partial^2 u}{\partial w^2},$$

$$\nabla_{\mathbf{f}} u = \begin{bmatrix} \frac{\partial u}{\partial f_1} & \cdots & \frac{\partial u}{\partial f_n} \end{bmatrix}', \qquad \nabla_{\mathbf{f}} u_w = \begin{bmatrix} \frac{\partial u_w}{\partial f_1} & \cdots & \frac{\partial u_w}{\partial f_n} \end{bmatrix}',$$

and

$$\nabla_{\mathbf{ff}} u = \begin{bmatrix} \frac{\partial u}{\partial f_1^2} & \frac{\partial u}{\partial f_1 f_2} & \cdots & \frac{\partial u}{\partial f_1 f_n} \\ \frac{\partial u}{\partial f_1 f_2} & \frac{\partial u}{\partial f_2^2} & \cdots & \frac{\partial u}{\partial f_2 f_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u}{\partial f_1 f_n} & \frac{\partial u}{\partial f_2 f_n} & \cdots & \frac{\partial u}{\partial f_n^2} \end{bmatrix}.$$

The value function $u(t, w, \boldsymbol{f})$ satisfies the HJB equation

$$\begin{aligned} u_t + \sup_{\boldsymbol{\pi}} \Bigg[& \boldsymbol{\pi'\mu} u_w + \left(\boldsymbol{\mu} - \frac{1}{2} D(\boldsymbol{vv'})\right)' \nabla_{\mathbf{f}} u + \boldsymbol{\pi' vv'} \nabla_{\mathbf{f}} u_w + \frac{1}{2} \boldsymbol{\pi' vv' \pi} u_{ww} \\ & + \frac{1}{2} \operatorname{Tr}\left(\boldsymbol{vv'} \nabla_{\mathbf{ff}} u\right) \Bigg] = 0, \end{aligned} \tag{5.17}$$

for $(t, w, f_1, \ldots, f_n) \in [0, T) \times \mathbb{R} \times \mathbb{R}^n$, along with the terminal condition

$$u(T, w, f_1, \ldots, f_n) = -e^{-\gamma w}, \quad \text{ for } (w, f_1, \ldots, f_n) \in \mathbb{R} \times \mathbb{R}^n.$$

The HJB equation can be written compactly as

$$u_t + \sup_{\pi} \mathcal{L}^{\pi} u = 0,$$

if we define the operator

$$\mathcal{L}^{\pi} u \equiv \boldsymbol{\pi'\mu} u_w + \left(\boldsymbol{\mu} - \frac{1}{2} D(\boldsymbol{vv'})\right)' \nabla_{\mathbf{f}} u + \boldsymbol{\pi' vv'} \nabla_{\mathbf{f}} u_w + \frac{1}{2} \boldsymbol{\pi' vv' \pi} u_{ww} + \frac{1}{2} \operatorname{Tr}\left(\boldsymbol{vv'} \nabla_{\mathbf{ff}} u\right).$$

We will first use the ansatz

$$u(t, w, \boldsymbol{f}) = -e^{-\gamma w} h(t, \boldsymbol{f})$$

to factor out w. Using the relations

$$u_t = -e^{-\gamma w} h_t, \quad u_w = \gamma e^{-\gamma w} h, \quad u_{ww} = -\gamma^2 e^{-\gamma w} h,$$

$$\nabla_{\mathbf{f}} u = -e^{-\gamma w} \nabla_{\mathbf{f}} h, \quad \nabla_{\mathbf{f}} u_w = \gamma e^{-\gamma w} \nabla_{\mathbf{f}} h, \quad \nabla_{\mathbf{ff}} u = -e^{-\gamma w} \nabla_{\mathbf{ff}} h,$$

we see that after substitution, the PDE (5.17) becomes

$$\begin{aligned} -h_t + \sup_{\boldsymbol{\pi}} \Bigg[& \boldsymbol{\pi'\mu} \gamma h - \left(\boldsymbol{\mu} - \frac{1}{2} D(\boldsymbol{vv'})\right)' \nabla_{\mathbf{f}} h + \gamma \boldsymbol{\pi' vv'} \nabla_{\mathbf{f}} h - \frac{1}{2} \gamma^2 \boldsymbol{\pi' vv' \pi} h \\ & - \frac{1}{2} \operatorname{Tr}\left(\boldsymbol{vv'} \nabla_{\mathbf{ff}} h\right) \Bigg] = 0, \end{aligned} \tag{5.18}$$

with terminal condition

$$h(T, \boldsymbol{f}) = 1.$$

From the first order condition, which is derived from differentiating the terms inside the supremum with respect to $\boldsymbol{\pi}$ and setting the equation to zero, we have

$$\gamma h \boldsymbol{\mu} + \gamma \boldsymbol{vv'} \nabla_{\mathbf{f}} h - \gamma^2 h \boldsymbol{vv' \pi} = 0,$$

so we can see that the optimal control can be expressed as

$$\boldsymbol{\pi}^*(t, \boldsymbol{f}) = \frac{1}{\gamma} \left((\boldsymbol{vv'})^{-1} \boldsymbol{\mu} + \frac{\nabla_{\mathbf{f}} h}{h} \right).$$

Substituting $\boldsymbol{\pi}^*$ back, the equation (5.18) becomes

$$-h_t + \frac{1}{2} \boldsymbol{\mu}' (\boldsymbol{vv'})^{-1} \boldsymbol{\mu} h + \frac{1}{2} D(\boldsymbol{vv'}) \nabla_{\mathbf{f}} h + \frac{\nabla_{\mathbf{f}} h' \boldsymbol{vv'} \nabla_{\mathbf{f}} h}{2h} - \frac{1}{2} \operatorname{Tr}\left(\boldsymbol{vv'} \nabla_{\mathbf{ff}} h\right) = 0. \tag{5.19}$$

Now if we let

$$h(t, \boldsymbol{f}) = \exp(-\Phi(t, \boldsymbol{f})),$$

we have these relations

$$h_t = -h\Phi_t, \quad \nabla_{\mathbf{f}} h = -h\nabla_{\mathbf{f}}\Phi, \quad \nabla_{\mathbf{ff}} h = h(\nabla_{\mathbf{f}}\Phi\nabla_{\mathbf{f}}\Phi' - \nabla_{\mathbf{ff}}\Phi).$$

After substitutions, we arrive at the linear, parabolic, PDE with only time-dependent coefficients

$$-\Phi_t + \frac{1}{2}\boldsymbol{\mu}'(\boldsymbol{v}\boldsymbol{v}')^{-1}\boldsymbol{\mu} + \frac{1}{2}D(\boldsymbol{v}\boldsymbol{v}')\nabla_{\mathbf{f}}\Phi - \frac{1}{2}\operatorname{Tr}(\boldsymbol{v}\boldsymbol{v}'\nabla_{\mathbf{ff}}\Phi) = 0, \tag{5.20}$$

with the terminal condition

$$\Phi(T, \boldsymbol{f}) = 0.$$

If we now use the ansatz

$$\Phi(t, \boldsymbol{f}) = \boldsymbol{f}'A(t)\boldsymbol{f} + \boldsymbol{f}'B(t) + C(t),$$

for a $n \times n$ matrix $A(t)$, a $n \times 1$ vector $B(t)$, and a scalar $C(t)$, we can see the derivatives are

$$\begin{aligned}
\frac{\partial\Phi}{\partial t} &= \boldsymbol{f}'\dot{A}(t)\boldsymbol{f} + \boldsymbol{f}'\dot{B}(t) + \dot{C}(t),\\
\nabla_{\mathbf{f}}\Phi &= \boldsymbol{f}'(A(t) + A(t)') + B(t),\\
\nabla_{\mathbf{ff}}\Phi &= A(t) + A(t)'.
\end{aligned}$$

After substituting the derivatives back, collecting terms of f_i^2 and f_i, and imposing the terminal condition $\Phi(T, \boldsymbol{f}) = 0$, we can deduce that $A(t) = B(t) = 0$, and

$$C(t) = \frac{T-t}{2}\boldsymbol{\mu}'(\boldsymbol{v}\boldsymbol{v}')^{-1}\boldsymbol{\mu}.$$

Therefore, the candidate solution is

$$u(t, w, \boldsymbol{f}) = -\exp\left(-\gamma w - \frac{1}{2}\boldsymbol{\mu}(t)'(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t)(T-t)\right) \tag{5.21}$$

which is independent of the log futures prices $\boldsymbol{f}$, and finally, the optimal controls are

$$\boldsymbol{\pi}^*(t, \boldsymbol{f}) = \frac{1}{\gamma}(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t) \tag{5.22}$$

which is also independent of $\boldsymbol{f}$.

The optimal wealth process is given by

$$\begin{aligned}
W_t^* &= \int_0^t \boldsymbol{\pi}^*(s)'\boldsymbol{\mu}(s)ds + \int_0^t \boldsymbol{\pi}^*(s)'\boldsymbol{v}(s)d\boldsymbol{z}_s\\
&= \frac{1}{\gamma}\int_0^t \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}\boldsymbol{\mu}(s)ds + \frac{1}{\gamma}\int_0^t \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}\boldsymbol{v}(s)d\boldsymbol{z}_s\\
&= \frac{1}{\gamma}\int_0^t \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}\boldsymbol{\mu}(s)ds + \frac{1}{\gamma}\int_0^t \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)')^{-1}d\boldsymbol{z}_s
\end{aligned}$$

We will show that the drift term is positive. To this end, define, for convenience, the square matrix $A \equiv \boldsymbol{v}(s)^{-1}$. Since

$$(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1} = (\boldsymbol{v}(s)')^{-1}\boldsymbol{v}(s)^{-1} = (\boldsymbol{v}(s)^{-1})'\boldsymbol{v}(s)^{-1} \equiv A'A,$$

we can see that $(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}$ is a positive definite matrix, from the observation that $\boldsymbol{\mu}'A'A\boldsymbol{\mu} = \|A\boldsymbol{\mu}\|^2 > 0$ for any non-zero vector $\boldsymbol{\mu}$. Therefore, we conclude that W_t^* is a submartingale; that is, the optimal wealth process, controlled by the optimal strategy $\boldsymbol{\pi}^*$, is expected to drift upward.

5.3 Verification Theorem

Based on the verification theorem in Appendix A, Theorem 7, in order to show that the candidate solution is the value function, it suffices to prove the uniform integrability of the family of random variables $\{u(\tau, W^*_\tau)\}$ for any stopping time $\tau \in [0,T]$.

Theorem 5. *The candidate solution found in (5.21) is equal to the value function (5.16); namely,*

$$V(t,w) = u(t,w) = -\exp\left(-\gamma w - \frac{1}{2}\boldsymbol{\mu}(t)'(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t)(T-t)\right)$$

on $(t,w) \in [0,T] \times \mathbb{R}$. *Furthermore, the optimal control is given by (5.22).*

Proof. Let $\epsilon > 0$ be an arbitrary positive number. From the verification theorem in Appendix A, it suffices to show that $\mathbb{E}(|u(\tau, W^*_\tau)|^{1+\epsilon}) < \infty$, uniformly with respect to any stopping time τ with $0 \le \tau \le T$. First, after applying the Cauchy-Schwarz inequality, we find that

$$\mathbb{E}(|u(\tau, W^*_\tau)|^{1+\epsilon}) \le \mathbb{E}(\exp(-2(1+\epsilon)\gamma W^*_\tau))^{1/2} \times \tag{5.23}$$
$$\mathbb{E}(\exp(-(1+\epsilon)\boldsymbol{\mu}(\tau)'(\boldsymbol{v}(\tau)\boldsymbol{v}(\tau)')^{-1}\boldsymbol{\mu}(\tau)(T-\tau)))^{1/2}. \tag{5.24}$$

For the first expectation (5.23), we apply Cauchy-Schwarz again to arrive at the inequality

$$\mathbb{E}[\exp(-2(1+\epsilon)\gamma W^*_\tau)]$$
$$= \mathbb{E}\left[\exp\left(-2(1+\epsilon)\left(\int_0^\tau \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}\boldsymbol{\mu}(s)ds + \int_0^\tau \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)')^{-1}d\boldsymbol{z}_s\right)\right)\right]$$
$$\le \mathbb{E}\left[\exp\left(-4(1+\epsilon)\int_0^\tau \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}\boldsymbol{\mu}(s)ds\right)\right]^{1/2} \times \tag{5.25}$$
$$\mathbb{E}\left[\exp\left(-4(1+\epsilon)\int_0^\tau \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)')^{-1}d\boldsymbol{z}_s\right)\right]^{1/2}. \tag{5.26}$$

The expectation in (5.25) is bounded by 1 for any stopping time τ, since $(\boldsymbol{v}(s)\boldsymbol{v}(s)')^{-1}$ is positive definite.

Now we look at (5.26). Define for notational convenience the martingale

$$M_t \equiv \int_0^t \boldsymbol{\mu}(s)'(\boldsymbol{v}(s)')^{-1}d\boldsymbol{z}_s.$$

Thus we have that for any constant k, $\exp(kM_t)$ is a submartingale by Jensen's inequality, since

$$\mathbb{E}_0(\exp(kM_t)) \ge \exp(k\mathbb{E}_0(M_t)) = \exp(k\mathbb{E}_0(M_0)) = 1.$$

Moreover, $\exp(kM_t)$ is positive. Therefore we can use Doob's martingale inequality [103]: for $\xi > 0$,

$$\begin{aligned}
\mathbb{E}[\exp(kM_\tau)] &= \left\| \exp\left(\frac{k}{1+\xi}M_\tau\right)\right\|_{1+\xi}^{1+\xi} \\
&\le \left\| \sup_{0\le t\le T} \exp\left(\frac{k}{1+\xi}M_t\right)\right\|_{1+\xi}^{1+\xi} \\
&\le \left(1+\frac{1}{\xi}\right)^{1+\xi} \sup_{0\le t\le T} \mathbb{E}[\exp(kM_t)] \\
&\le c
\end{aligned}$$

where c is a positive constant, independent of the stopping time τ.

For the second expectation (5.24), since we showed that $(\boldsymbol{v}(\tau)\boldsymbol{v}(\tau)')^{-1}$ is positive definite, we use this observation again to deduce that the second expectation is again bounded by 1. ∎

5.4 Numerical Implementation

With the closed-form expressions derived, we now use the parameters calibrated in Panel A of Table 1 in [37] to illustrate the optimal trading strategies for WTI crude oil futures contracts traded on NYMEX. We fix $\gamma = 0.01$ and consider n up to $n = 4$ since, as discussed, with four factors the model captures the volatility term structure accurately.

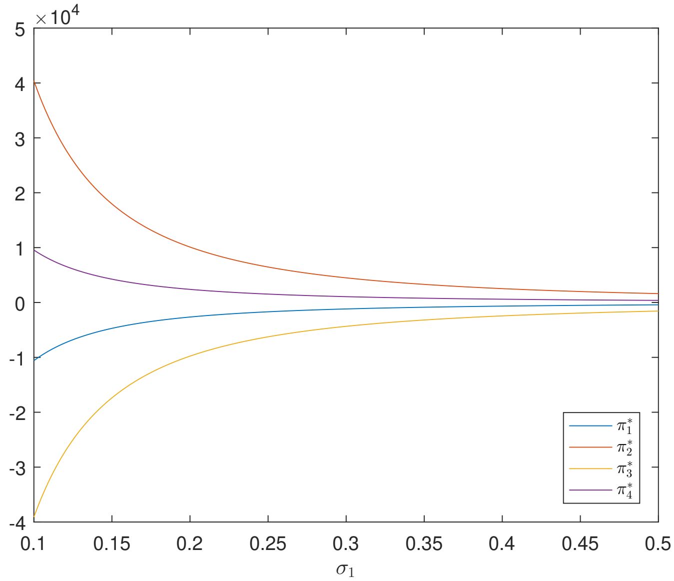


Figure 5.1: The optimal holdings π_i^* for $i = 1, \ldots, 4$ as functions of $\sigma_1 = \ldots = \sigma_4$.

In Fig 5.1 we plot $\boldsymbol{\pi}^*$ as function of the volatility parameter σ_1. We let the values of all 4 volatility parameters to be the same, $\sigma_1 = \ldots = \sigma_4$, and let all parameters range from $(0.15, 0.50)$, and plot $\pi_1^*, \ldots, \pi_4^*$. We set the maturities to be 3 months apart, starting with $T_1 = 3/12$, and so $T_2 = 6/12, T_3 = 9/12$, and $T_4 = 1$. As expected, the position sizes decrease as volatility increases, which agrees with the intuition of a risk-averse investor. Moreover, the decrease is exponentially fast, which suggests volatility will decrease position sizes faster than an increase in the coefficient of risk aversion.

Certainty equivalent is the inverse of the exponential utility function. Precisely, from (5.16) we have

$$\begin{aligned}\mathcal{C}(t, w) &= w + \frac{\Phi(t)}{\gamma} \\ &= w + \frac{1}{2\gamma}\boldsymbol{\mu}(t)'(\boldsymbol{v}(t)\boldsymbol{v}(t)')^{-1}\boldsymbol{\mu}(t)(T - t). \end{aligned} \tag{5.27}$$

Without loss of generality, we set $w = 0$. We plot $\mathcal{C}^{(n)}(0, 0)$ on Fig 5.2 for $n = 2, 3, 4$, based on different sets of parameters from [37] under different n, as a function of T at $t = 0$. Here, for any particular n, it means there are n factors as well as contracts of n different maturities available to trade. First of all, the certainty equivalent increases as a function of T, which means that the more time the trader has, the more profits can be made. Secondly, $\mathcal{C}^{(n)}$ increases from $n = 2$ to $n = 4$. The interpretation is that the utility derived from trading is higher in markets with more factors, and hence, more non-redundant contracts of different maturities (which cannot be replicated by other contracts).

Now it would be interesting to look at the certainty equivalent for trading $m < n$ contracts under a n factor model. Individually, we can immediately see that the drifts in (5.12), and volatilities in (5.13), remain identically the same, for the contract of maturity T_k. We will use the tilde notation $\tilde{\boldsymbol{\mu}}(t)$ and $\tilde{\boldsymbol{v}}(t)$ to denote the *smaller* versions when there are m contracts with n factors where, of course, the vector $\tilde{\boldsymbol{\mu}}(t)$ is now $m \times 1$, and the matrix $\tilde{\boldsymbol{v}}(t)$ is now a non-square $m \times n$ matrix. The vector of diagonal elements

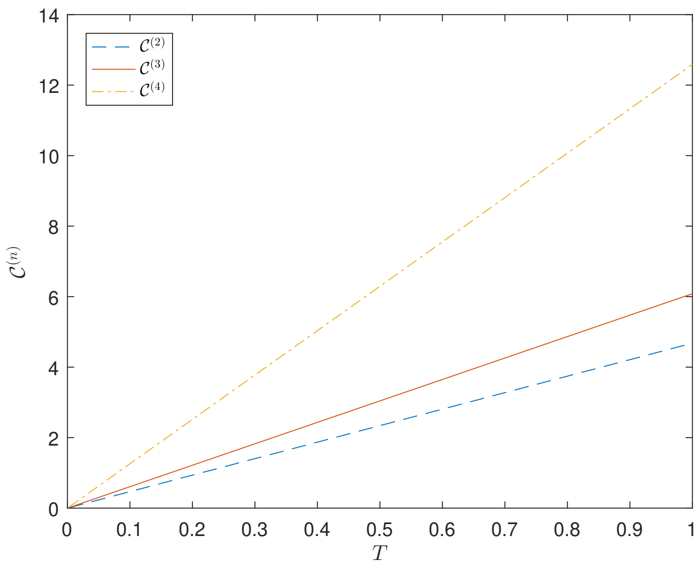


Figure 5.2: Certainty equivalent $\mathcal{C}^n$ for $n = 2, 3, 4$ as functions of T at $t = 0$.

$D(\tilde{\boldsymbol{v}}(t)\tilde{\boldsymbol{v}}(t)')$ is $m \times 1$. As an example, for a $n = 4$ factors market where only the first T_1 and third T_3 maturities are available, $m = 2$ in this case, and $\tilde{\boldsymbol{\mu}}(t)$ contains the first and third element of $\boldsymbol{\mu}(t)$, and $\tilde{\boldsymbol{v}}(t)$ contains the first and third row of $\boldsymbol{v}(t)$.

The vector of the log futures prices $\tilde{\boldsymbol{f}}$ and the optimal controls $\tilde{\boldsymbol{\pi}}(t, \tilde{\boldsymbol{f}})$ both have dimensions $m \times 1$. We will use $\tilde{W}_t$ to denote the wealth process, and the system of equations in (5.15) now becomes

$$\begin{bmatrix} d\tilde{W}_t \\ d\tilde{\boldsymbol{f}}_t \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{\pi}}'\tilde{\boldsymbol{\mu}}(t) \\ \tilde{\boldsymbol{\mu}}(t) - \frac{1}{2}D(\tilde{\boldsymbol{v}}(t)\tilde{\boldsymbol{v}}(t)') \end{bmatrix} dt + \begin{bmatrix} \tilde{\boldsymbol{\pi}}'\tilde{\boldsymbol{v}}(t) \\ \tilde{\boldsymbol{v}}(t) \end{bmatrix} d\boldsymbol{z}_t. \tag{5.28}$$

The steps to derive the solution and optimal control remain the same; in particular, the certainty equivalent (5.27) is now

$$\tilde{\mathcal{C}}(t, w) = w + \frac{1}{2\gamma}\tilde{\boldsymbol{\mu}}(t)'(\tilde{\boldsymbol{v}}(t)\tilde{\boldsymbol{v}}(t)')^{-1}\tilde{\boldsymbol{\mu}}(t)(T - t).$$

To illustrate, we will set again $n = 4$, and consider all possible combinations of subsets of $n = 4$ contracts. Denote by $\mathcal{C}^{(4)}_{1234}$ the certainty equivalent when there are all 4 maturities $T_k, k = 1, 2, 3, 4$ available, by $\mathcal{C}^{(4)}_{123}$ the certainty equivalent when there are maturities T_1, T_2 and T_3 available, by $\mathcal{C}^{(4)}_{12}$ the certainty equivalent when there are maturities T_1 and T_2 available, and so on.

We have 15 possible combinations, and we plot the certainty equivalents for all of these 15 possibilities in Fig 5.3 at $T = 0.6$ and at time $T = 0.8$. We can see, as expected, given the number of factors n is the same, the more contracts available, the higher the certainty equivalent is. In other words, under a n factor model, since the market with only m maturities, where $m < n$, available to trade is a more restricted universe than that with n maturities, the expected utility is higher in a less restricted environment. This implies contracts of different maturities are *not* redundant, and it is more beneficial, that is, more utility can be derived, to trade in a market with more maturities available. Moreover, comparing the top and bottom panels of Fig 5.3, we can see that the differences between certainty equivalents of different number of available contracts become relatively narrower as T decreases.

Finally, historical daily futures prices were obtained from Bloomberg from 1992 to 2001, coinciding with the period over which the parameters were calibrated. In contrast to previous graphs, we plot in Fig 5.4 the *number* of contracts, that is, $\pi_i^*(t)/F_i(t)$. In general, throughout the period, π_1^* and π_3^* are positive, and the opposite is true for π_2^* and π_4^*. Cyclicality can be observed as the contracts reach maturity. A peak in terms of position sizes around the beginning of 1999 can also be seen, coinciding with the front month futures contract trading at the low end around \$10.

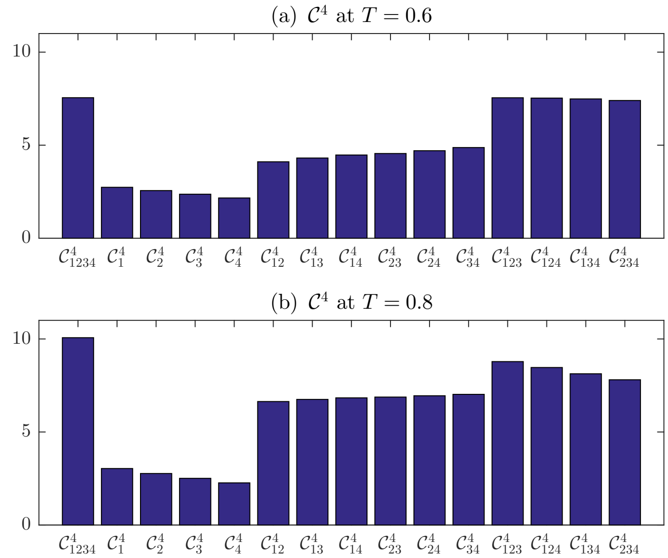


Figure 5.3: Certainty equivalent $\mathcal{C}^4$ for all 15 possible combinations, (a) at $T = 0.6$, and (b) at $T = 0.8$.

5.5 Summary

The stochastic approach to trading cointegrated securities in a more general multi-factor setting have been presented in [87, 29, 65]. Our model in this chapter is simpler due to the absence of the cointegrating vector, which was present in Chapter 2. Here, we have extended the investigation of optimal trading in commodity futures market under two-factor models to a multi-factor model. Closed-form expressions for the optimal controls and for the value function are obtained.

Using these formulae, we illustrated the optimal strategies using parameters calibrated from historical price data, and conclude that position sizes decrease as volatilities increase, and as risk aversion increases. Higher utility, as measured by certainty equivalent, can be derived from trading in multiple contracts, of different maturities, beyond pairs trading with only two, which is possible only in multi-factor models. Moreover, by fixing a particular n, which we illustrated using $n = 4$, we examined the effect of varying the number of different maturities to trade with. Again, intuitively, it should be more beneficial to be able to access a larger set of securities, and this intuition is confirmed quantitatively.

Given our observation that higher certainty equivalent can be achieved from trading in markets with more factors, the interesting question is to identify these markets. This has been studied extensively in empirical finance, in particular in the interest rate market; see for example the well-known three factor models in [88] or [42]. Techniques in multivariate statistics such as principal component analysis or factor analysis [104, 114] have been developed to quantify the number of factors. On the other hand, after correctly identifying the number of factors, the next issue is still the correct characterization of the factors' dynamics. For example, in their sequel paper [36], the authors improve their model by incorporating stochastic volatility.

After locating the market with a high number of factors, we take the next step in our quest for profit to consider all available contracts of different maturities in that market. From our numerical example, the highest certainty equivalent is achieved from trading every contract that is available. In reality, the critical concern is transaction costs – trading more contracts might not necessarily be more profitable in the presence of market frictions. It is equally important to identify the breakeven point in the number of contracts beyond which profitability might be eroded by transaction costs, and the determination of which requires incorporation of transaction costs into the model. It is an important consideration left for future work.

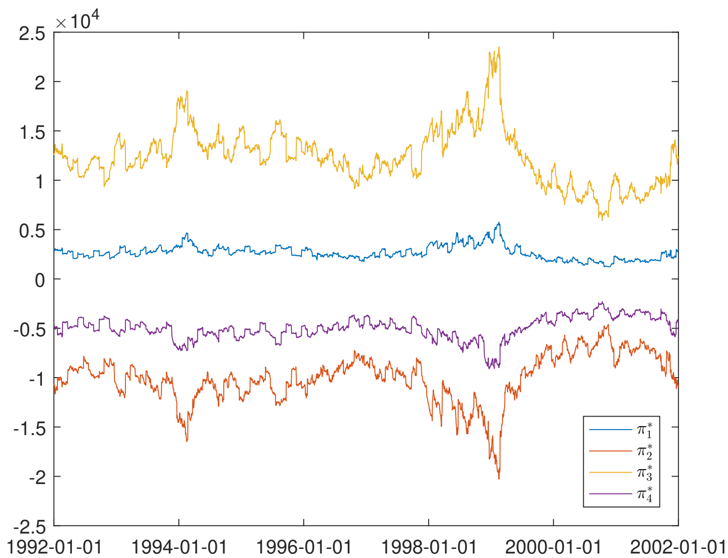


Figure 5.4: The optimal holdings π_i^* for $i = 1, \ldots, 4$ over the period 1992-2001.

Chapter 6

Conclusion and Future Work

In this thesis, we applied stochastic control theory to arrive at dynamic trading strategies for stocks and futures. Control problems were formulated and solved, and the optimal controls and candidate solutions to the HJB equations were obtained in closed form. Based on a verification theorem, the candidate solutions were shown to be equal to the value functions. We further illustrated the controls' and the solutions' properties, using parameters calibrated by existing works in the literature, and we could see that they were in-line with intuition, and were feasible for practical implementation due to their simplicity.

In Chapter 2, we applied these steps to the Duan-Pliska [44] model where the stocks are cointegrated. We considered a pair of related equities which belong to the financial sector. Pairs trading in the stock market has been well-established approach to trading stocks, where the objective is to execute trades when a pair of related stocks diverge from their historical equilibrium level, and to profit from the reverting back of the pair to their equilibrium level. The chapter was published in [116], and we applied stochastic control theory in the context of pairs trading, which resulted in a closed-form formula for the optimal strategy. We then furthermore added correlation into the stocks' dynamics, which complicated the formula but made the model more realistic.

Chapter 3 was published in [81], where we applied the same methodology to volatility index futures, when the spot index is modeled as a CTOU process [91]. In contrast to stocks, futures contracts on the same underlying, but of different maturities, typically exhibit a term structure. Furthermore, the VIX index exhibits mean-reversion. For trading strategies to be effective, they have to be based on models such as CTOU that captures these various features, and we presented the optimal strategy for trading volatility index futures in a closed-form formula. Based on parameters calibrated from VIX futures historical data, we found that traders should take bigger positions in the long end of the futures curve, in line with the monotonic decrease in the volatility term structure.

After studying futures on a volatility index, we proceeded to study commodity futures under the well-known two-factor Schwartz model [107] in Chapter 4, which was published in [82]. In contrast to models for volatility indices, models for the spot prices of commodities typically include convenience yield. which is absent, for example, in the VIX futures case. Furthermore, mean-reversion in the spot commodity price processes is not relevant under the Schwartz model. Hence we applied the same methodology to study a different market, and provided a closed-form formula for the optimal strategy under the Schwartz model. Based on parameters calibrated from WTI crude oil futures data, we saw that the optimal positions changed little with respect to maturities. Our results could shed light on how managed futures funds or commodity trading advisers operate their trading strategies.

All of the previous chapters are based on stochastic models driven by two factors. We then considered a more general n-factor model in Chapter 5, since models with more factors fit the term structures better. We therefore study optimal trading under the more general, yet tractable, n-factor Cortazar and Naranjo model [37]. We concluded in the chapter that both CTOU model and Schwartz model are nested under this general model when $n = 2$. We presented a tractable formula for the optimal trading strategies, which was illustrated with numerical examples based again on historical WTI crude oil futures prices.

In all chapters, we proved that the candidate solutions were equal to the value function, based on a standard verification theorem presented in Appendix A. Essentially, the theorem required a uniform integrability condition to hold, and the condition was proved individually in all of the preceding chapters. In the cointegrated and correlated stocks pairs trading case in Chapter 2, we derived the sufficient conditions, in the form of inequalities, for uniform integrability to hold. We saw that these inequalities are independent of the risk aversion parameter. On the other hand, the uniform integrability condition held unconditionally for all the futures trading cases.

In general, we found that the optimal positions decreased in magnitude as the volatilities in the underlying factors increased. We found that the magnitudes of the optimal positions were inversely proportional to the coefficients of risk aversion, similar to the solutions to the classical Merton portfolio optimization problem. Using the calibrated, hence realistic, model parameters, for all the 2-factor models, the positions are shown to have opposite signs, corresponding to one long and one short position, which is the typical setup in the context of pairs trading.

For the futures trading applications, we found that the certainty equivalent for trading two contracts simultaneously was significantly greater than that derived from trading only a single contract. Similar result extended to the n-factor model. Furthermore, it was shown that for the futures trading models, the optimized wealth processes controlled by the optimal strategies were submartingales with a positive drift. The optimal strategies were independent of the optimization horizon for all futures trading cases, but not for the cointegrated stock case. We also saw that the optimal holdings were independent of the hidden processes, thereby eliminating the need to estimate the state variables. The optimal holdings were seen to be independent of the optimization horizon, which is chosen arbitrarily. In the n-factor case, we saw that markets with a higher number of factors led to higher certainty equivalent. Moreover, given a particular number of factors, the certainty equivalent was higher when the trader could trade more contracts of different maturities; this confirmed the intuition that trading in a less restricted asset universe led to higher expected utility.

There are numerous directions for future research. Recall in this thesis, stocks in the equity market was studied in Chapter 2, VIX futures was studied in Chapter 3, commodity futures was studied in Chapter 4, which was further extended to a multi-factor model Chapter 5. The next step could be to continue the study, along this theme, to other instrument types and markets, such as interest rates products based on multifactor models (an example of which is [42]), or currencies in the foreign exchange market, which was explored in [119], or cryptocurrencies, which was explored in [87]. Related securities, other than cointegrated stocks from the same sector, are plentiful in the futures market; for example crush spreads [110] which consists of soybean, soybean meal, and soy oil, or crack spreads [55] which commonly refers to crude oil, gasoline and heating oil. Less researched financial instruments such as volatility indexes on commodities, single stocks, or interest rates, or volatility index on even VIX itself, are closely related to the securities in this thesis, and may provide opportunities for traders or researchers to venture into less crowded areas. The goal will be to exploit features unique to different instruments and markets, an example of which we have already seen in this thesis is the existence of term structures in futures market, which are absent in stocks.

Different asset classes requires different models. As already seen in Chapter 3, the CTOU model was used specifically to capture mean reversion in volatility indices. A subsequent study in the optimal trading in, as an example, the gold futures market, might require a three-factor model with the addition of stochastic interest rate to capture their complicated relationship; see for example [60] for a study of how gold price is affected by macro-economic variables. Another possibility is to study options on futures, which will require the incorporation of stochastic volatility in the spot dynamics, as formulated in [36]. Other studies include [86] which also focuses on stochastic volatility, or [91], which, in addition to stochastic volatility, also explores jumps diffusion. Fractional cointegration for commodity futures was explored in [43]. A regime-switching version of the Central Tendency Ornstein-Uhlenbeck for futures trading was presented in [84]. Regime-switching models designed for VIX options was presented in [61]. For VIX futures, where the VIX index is actually a function of the implied volatilities of S&P options, it would be interesting to investigate trading strategies under a joint model of VIX and the S&P index (see for example [33] or [27]), and to expand the asset universe to include the index, via futures or Exchange Traded Funds, and options, into the investor's portfolio.

We have already observed the independence of the strategy on the optimization horizon in futures trading; therefore extending the horizon to infinity is not relevant there. It would be interesting to consider problems of infinite horizon in stocks trading where the notion of contract maturity is not pertinent, but mathematically technical difficulties may arise in the proofs. Inclusion of consumption in the objective function, for example [127], is common from an micro-economic perspective.

Stochastic control theory encompasses many different formulations of the control problems. Other formulations, structurally different from those in this thesis, such as optimal stopping [80, 83] to determine when to exit trades, optimal switching [101] to determine which assets to trade, or optimal control of partially observable system to handle hidden variables [17], are promising approaches to apply the theory in the future to design trading strategies, for different markets under possibly more complicated models.

Besides the modeling aspect, in reality, a critical factor to profitability in trading in the real world is transaction costs. We have seen in Chapter 5 where trading more contracts of different maturities leads to higher certainty equivalent; however the presence of transaction costs or other market frictions will certainly impact profitability. [41],[40], or [117] are early attempts to incorporate this realistic aspect into portfolio optimization problems. A recent study that account for transaction costs is [76], which also investigates optimal strategies in the context of pairs trading similar to the setting in this thesis. These papers present the starting points to study financial control problems that incorporate transaction costs; but further investigation is required when the stochastic models become more complicated in different markets.

Other considerations from a finance perspective in trading include the effect of taxation. Moreover, futures are typically traded with leverage, and margin requirement is also a core issue. On the other hand, as opposed to speculators or investors, for traders acting on behalf of actual users of futures market, such as oil producers or airline companies, utility maximization is not limited to expected utility of wealth alone, and other aspects such as storage costs or the rate of consumption of the actual commodities are important business-related factors. Incorporating these practical aspects into futures portfolio optimization may not be straightforward, but will certainly have practical implications.

From the perspective of a market maker, stochastic control theory is also the natural methodology to design trading policies to maximize profits and to manage inventory. Market making, in particular in the high frequency domain, has received tremendous attention due to recent technological advances and availability of data. In this setting, the optimal controls are the bid and ask quotes to be posted at the dealer's discretion. Beginning with the seminal paper [9] who computed numerically the optimal bid/ask spreads, [64] solved the HJB equations in closed form. See [66] for a formulation of the control problem as a Hamilton-Jacobi-Bellman quasi-variational inequality. Elaborate studies in limit order books (LOB) are presented in [34] and the references therein, and Hawkes processes for modeling the arrivals and cancellations of limit orders are studied in [11]. Similar approaches that applied stochastic control theory to market making, but focused on the options market, were studied in [113] and [48]. The formulations and models in these papers are based on stochastic control theory similar to those considered in this thesis, and further investigations to incorporate realistic aspects to make markets in real world continue to attract attention from practitioners and researchers.

However, under more complicated models, whether from the choice of stochastic models, or from the consideration of other operational constraints and costs, the value functions and the optimal controls will likely require numerical approximations. Moreover, the solutions of the HJB equations may not exist in the classical sense, and hence the notion of *viscosity solutions* is needed to define them in a weaker sense. The theory of viscosity solutions, and the numerical methods for calculating them, has been extensively developed since the paper [14] was published, who provided very general conditions for a numerical scheme to converge; namely, the scheme should be monotone, stable, and consistent. For detailed overviews of the theory, we again refer to [58]. For a recent collection of control problems in finance and numerical methods, we refer to [10], which relies on the theory of viscosity solutions. It focuses primarily on the theory of impulse control, which is yet another formulation of control problems. Impulse control has also been shown to effectively address the incorporation of transaction costs in relatively simple models; see for example [97]. This technique should be a promising approach for more complicated models with more complicated constraints, and it is left to be explored in the future.

Another major complication for practical implementation of optimal trading strategies is the estimation of model parameters. Under the relatively simple models considered in thesis, the well-established Kalman filtering methodology, and the two-step Engle-Granger procedure with Dickey-Fuller test for cointegration, served the purpose. However the effectiveness of the strategies obtained from more complicated models, or from more complicated formulation of the control problems, will inevitably hinge on accurate estimates of parameters. From the theoretical side, more refined statistical methods, such as unscented Kalman filter [123] for estimating nonlinear dynamic systems, or the Johansen [70] framework of vector-autoregression to detect unit roots and cointegration in systems of time series, continue to be developed. Techniques in multivariate statistics such as principal component analysis or the more specialized independent component analysis [114] should prove to be useful for factor identification, which is relevant in Chapter 5 when choosing which markets to trade in. Empirically, studies such as [32] raised questions on pairs trading based on detection of cointegration, and prompt further research on the co-movement of stocks in aggregate. Moreover, in the presence of extensive and well-developed derivative markets, the relationship between historically estimated versus derivatives implied parameters warrants further exploration, since the use of one versus the other or some combination of both will have a profound impact on the optimal trading strategies.

Finally, arguably the most important aspect of any trading strategies is, of course, the actual profitability, and a priori assessment of which can only come from the backtesting on historical data. Furthermore, different metrics of optimality, besides the form of utility functions (exponential utility function in this thesis, power utility function in [18], or log utility function), such as drawdowns, expected shortfall, value-at-risk, mean-variance criteria, or other risk measures, will lead to interesting comparisons between the different ensuing *optimal* trading strategies, where optimality is defined differently under different metrics.

Appendix A

Verification Theorem

Here we present a verification theorem that was used to prove that the candidate solutions derived from the solution of the HJB equations were indeed the value functions of the optimal control problems. The main assumption is a uniform integrability condition, which was proven previously in various chapters. All results are standard; for more details we can refer to [58, 105, 96]. First, recall the definition of uniform integrability:

Definition 1 (Uniform Integrability)**.** *A family of random variables $\{X_\tau\}_\tau$ indexed by τ is said to be* uniformly integrable *if*

$$\lim_{K\to\infty}\left(\sup_\tau \mathbb{E}[|X_\tau|\mathbf{1}_{\{|X_\tau|\geq K\}}]\right) = 0.$$

It is a standard result that uniform integrability enables us to exchange limits of expectations similar to Lebesgue's dominated convergence theorem.

In the proofs in all previous chapters, we rely on this sufficient condition for uniform integrability:

Theorem 6. *A sufficient condition for a family of random variables $\{X_\tau\}_\tau$ to be uniformly integrable is if*

$$\sup_\tau \mathbb{E}[|X_\tau|^{1+\epsilon}] < \infty$$

for some $\epsilon > 0$.

Proof. Let $M < \infty$ be the supremum. For all $n > 0$, there exists $C_n \in \mathcal{R}$ such that $|x|^{1+\epsilon} \geq nM|x|$ for all $|x| \geq C_n$; namely, we can choose $C_n = (nM)^{1/\epsilon}$ for each n. Therefore,

$$M \geq \mathbb{E}[|X_\tau|^{1+\epsilon}] \geq \mathbb{E}[|X_\tau|^{1+\epsilon}\mathbf{1}_{\{|X_\tau|\geq C_n\}}] \geq nM\mathbb{E}[|X_\tau|\mathbf{1}_{\{|X_\tau|\geq C_n\}}]$$

Hence, $\sup_\tau \mathbb{E}[|X_\tau|\mathbf{1}_{\{|X_\tau|\geq C_n\}}] \leq \frac{1}{n}$, and the uniform integrability of $\{X_\tau\}_\tau$ follows. □

In what follows we will let $\boldsymbol{X}_t$ denote, in general, the security price processes in previous chapters: $\boldsymbol{X}_t = (S_t^{(1)}, S_t^{(2)})$ in Ch.2, $\boldsymbol{X}_t = (F_t^{(1)}, F_t^{(2)})$ in Ch.3 and 4, and $\boldsymbol{X}_t = \boldsymbol{f}_t$ in Ch.5.

Theorem 7. *Let $u(t, w, \boldsymbol{x})$ be a function in class $C^{1,2,2}([0,T) \times \mathbb{R} \times \mathbb{R}^n)$ such that $\{u(\tau, W_\tau^*, \boldsymbol{X}_\tau)\}_\tau$ is a uniformly integrable family of random variables, where $\boldsymbol{\pi}^*$ is an admissible control with the property that $u_t + \mathcal{L}^{\pi *}u = 0$, and $\tau \in [0,T]$ is a stopping time for W_τ^* starting at $W_t^* = w$. If furthermore $u(T, w, \boldsymbol{x}) = U(w)$ for every admissible control $\boldsymbol{\pi}$ and $u_t + \mathcal{L}^\pi u \leq 0$, then*

$$u(t, w, \boldsymbol{x}) = V(t, w, \boldsymbol{x}) \quad \forall (t, w, \boldsymbol{x}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^n.$$

Proof. We need to show that for any admissible $\boldsymbol{\pi}$ such that $\mathcal{L}^{\pi} u \leq 0$, the expected utility of terminal wealth will be less than or equal to what the value function would indicate, namely,

$$E_t[U(W_T^{\boldsymbol{\pi}})] \leq u(t, w, \boldsymbol{x}), \tag{A.1}$$

and that the equality holds when the wealth process is controlled optimally by $\boldsymbol{\pi}^*$; that is,

$$V(t, w, \boldsymbol{x}) \equiv \sup_{\boldsymbol{\pi} \in \mathcal{A}_t} \mathbb{E}_t[U(W_T^{\boldsymbol{\pi}})] = \mathbb{E}_t[U(W_T^*)] = u(t, w, \boldsymbol{x}). \tag{A.2}$$

Let $\{\tau_n\}_n$ be a localizing sequence of stopping times for W_t, which starts at w at time t. Introducing the stopping times $T_n = \min(\tau_n, T)$, we get from Ito's formula

$$\begin{aligned} u(T_n, W_{T_n}^{\pi}, \boldsymbol{X}_{T_n}) &= u(t, w, \boldsymbol{x}) + \int_t^{T_n} \frac{\partial u}{\partial s} + \mathcal{L}^{\pi} u(s, W_s^{\pi}, \boldsymbol{X}_s) ds + \int_t^{T_n} \boldsymbol{G}(s) d\boldsymbol{z}_s \\ &\leq u(t, w, \boldsymbol{x}) + \int_t^{T_n} \boldsymbol{G}(s) d\boldsymbol{z}_s, \end{aligned}$$

since $u_t + \mathcal{L}^{\pi} u \leq 0$, where

$$\boldsymbol{G}(s) = [u_w \quad u_1 \quad u_2] \cdot \begin{bmatrix} \pi_1 \sigma_1 S_s^{(1)} & \pi_2 \sigma_2 S_s^{(2)} \\ \sigma_1 S_s^{(1)} & 0 \\ 0 & \sigma_2 S_s^{(2)} \end{bmatrix} \qquad \text{in Chapter 2,}$$

$$\boldsymbol{G}(s) = [u_w \quad u_1 \quad u_2] \cdot \begin{bmatrix} \pi_1 \sigma_{v1}(s) F_s^{(1)} + \pi_2 \sigma_{v2}(s) F_s^{(2)} & \pi_1 \sigma_{\theta 1}(s) F_s^{(1)} + \pi_2 \sigma_{\theta 2}(s) F_s^{(2)} \\ \sigma_{v1}(s) F_s^{(1)} & \sigma_{\theta 1}(s) F_s^{(1)} \\ \sigma_{v2}(s) F_s^{(2)} & \sigma_{\theta 2}(s) F_s^{(2)} \end{bmatrix} \qquad \text{in Chapter 3}$$

$$\boldsymbol{G}(s) = [u_w \quad u_1 \quad u_2] \cdot \begin{bmatrix} \pi_1 \sigma_1(s) F_s^{(1)} & \pi_2 \sigma_2(s) F_s^{(2)} \\ \sigma_1(s) F_s^{(1)} & 0 \\ 0 & \sigma_2(s) F_s^{(2)} \end{bmatrix} \qquad \text{in Chapter 4, and}$$

$$\boldsymbol{G}(s) = \boldsymbol{\pi}' \boldsymbol{v}(s) u_w + \nabla_{\boldsymbol{f}} u' \boldsymbol{v}(s) \qquad \text{in Chapter 5.}$$

Since the Ito's integral is a martingale, after taking expectation, we find

$$u(t, w, \boldsymbol{x}) \geq \mathbb{E}_t[u(T_n, W_{T_n}^{\pi}, \boldsymbol{X}_{T_n})]. \tag{A.3}$$

Obviously, equality holds if $\mathcal{L}^{\pi *} u = 0$, for $\pi = \pi^*$. On the other hand, due to the continuity of u, we have

$$\lim_{n \to \infty} u(T_n, W_{T_n}^{\pi}, \boldsymbol{X}_{T_n}) = u(T, W_T^{\pi}, \boldsymbol{X}_T) = U(W_T^{\pi}).$$

Now

$$\begin{aligned} \mathbb{E}_t[U(W_T^{\pi})] &= \mathbb{E}_t\left[\lim_{n \to \infty} u(T_n, W_{T_n}^{\pi}, \boldsymbol{X}_{T_n})\right] \\ &= \lim_{n \to \infty} \mathbb{E}_t\left[u(T_n, W_{T_n}^{\pi}, \boldsymbol{X}_{T_n})\right] \qquad \text{since } \{u(\tau, W_\tau^{\pi}, \boldsymbol{X}_\tau)\}_\tau \text{ is U.I.} \\ &\leq u(t, w, \boldsymbol{x}) \qquad \text{from (A.3),} \end{aligned}$$

and equality holds if $\mathcal{L}^{\pi *} u = 0$. We have therefore proved (A.1) and (A.2). □

So we see in order to verify that the candidate solution is the value function, it suffices to prove uniform integrability of $\{u(\tau, W_\tau^{\pi}, \boldsymbol{X}_\tau)\}_\tau$, which was shown in the *Verification* sections in previous chapters.

Bibliography

[1] S. Ahuja, G. Papanicolaou, W. Ren, and T.-W. Yang. Limit order trading with a mean reverting reference price. *Risk and Decision Analysis*, 6(2):121–136, 2017.

[2] C. Alexander and D. Korovilas. Volatility exchange-traded notes: curse or cure. *Journal of Alternative Investments*, 16(2):52–70, 2013.

[3] A. Almansour. Convenience yield in commodity price modeling: a regime switching approach. *Energy Economics*, 53:238–247, 2016.

[4] B. Angoshtari. On the market-neutrality of optimal pairs-trading strategies. *ArXiv e-prints*, Aug. 2016.

[5] B. Angoshtari and T. Leung. Optimal dynamic basis trading. working paper, 2018.

[6] M. J. Anson. Maximizing utility with commodity futures diversification. *Journal of Portfolio Management*, 25(4):86–94, 1999.

[7] D. Aronson. *Evidence-Based Technical Analysis: Applying the Scientific Method and Statistical Inference to Trading Signals*. Wiley, 2006.

[8] A. G. Asuero, A. Sayago, and A. G. Gonzalez. The correlation coefficient: an overview. *Critical Reviews in Analytical Chemistry*, 2006.

[9] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.

[10] P. Azimzadeh. *Impulse Control in Finance: Numerical Methods and Viscosity Solutions*. PhD thesis, University of Waterloo, 2017.

[11] E. Bacry and J.-F. Muzy. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14(7):1147–1166, 2014.

[12] J. Baldeaux and A. Badran. Consistent modelling of VIX and equity derivatives using a 3/2 plus jumps model. *Applied Mathematical Finance*, 21(4):299–312, 2014.

[13] P. Bank and M. VoB. Linear quadratic stochastic control problems with stochastic terminal constraint. *SIAM Journal on Control and Optimization*, 56(2):672–699, 2018.

[14] G. Barles and P. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis*, 4:2347–2349, Jan 1991.

[15] J. Baruník and B. Malinská. Forecasting the term structure of crude oil futures prices with neural networks. *Applied Energy*, 164:366–379, 2016.

[16] D. S. Bates. U.S. stock market crash risk, 1926-2010. *Journal of Financial Economics*, 105:229–259, 2012.

[17] A. Bensoussan. *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1992.

[18] F. E. Benth and K. H. Karlsen. A note on Merton's portfolio selection problem for the Schwartz mean-reversion model. *Stochastic Analysis and Applications*, 23(4):687–704, 2005.

[19] M. Bichuch and S. Shreve. Utility maximization trading two futures with transaction costs. *SIAM Journal of Financial Mathematics*, pages 26–85, 2013.

[20] Z. Bodie, A. Kane, and A. J. Marcus. *Investments*. McGraw-Hill Education, 2017.

[21] M. Boguslavsky and E. Boguslavskaya. Arbitrage under power. *Risk*, pages 69–73, 2004.

[22] R. Brenner and K. Kroner. Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *Journal of Financial and Quantitative Analysis*, 30(1):23–42, 1995.

[23] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*. Springer Finance, 2 edition, 2006.

[24] J. F. Caldeira. Selection of a portfolio of pairs based on cointegration: a statistical arbitrage strategy. *Rev. Bras. Financas (Online), Rio de Janeiro*, 11(1):49–80, 2013.

[25] R. Carmona and M. Ludkovski. Spot convenience yield models for energy markets. In G. Yin and Y. Zhang, editors, *AMS Mathematics of Finance vol. 351 of Contemporary Mathematics*, pages 65–80. 2004.

[26] P. Carr and D. Madan. *Towards a theory of volatility trading*, page 458–476. Cambridge University Press, 2001.

[27] P. Carr and D. B. Madan. Joint modeling of VIX and SPX options at a single and common maturity with risk management applications. *IIE Transactions*, 46(11):1125–1131, 2014.

[28] A. Cartea, L. Gan, and S. Jaimungal. Trading co-integrated assets with price impact. *Mathematical Finance*, 29(2):542–567, 2019.

[29] A. Cartea and S. Jaimungal. Algorithmic trading of co-integrated assets. *International Journal of Theoretical and Applied Finance*, 19(06):1650038, 2016.

[30] K. Chen, M. C. Chiu, and H. Y. Wong. Time-consistent mean-variance pairs-trading under regime-switching cointegration. *Available at SSRN: https://ssrn.com/abstract=3250340*, 2018.

[31] Chicago Board Options Exchange. *VIX: CBOE volatility index white paper*, 2019. https://www.cboe.com/micro/vix/vixwhite.pdf.

[32] M. Clegg. On the persistence of cointegration in pairs trading. *Available at SSRN: https://ssrn.com/abstract=2491201*, 2014.

[33] R. Cont and T. Kokholm. A consistent pricing model for index options and volatility derivatives. *Mathematical Finance*, 23(2):248–274, 2013.

[34] R. Cont, S. Stoikov, and R. Talreja. A stochastic model for order book dynamics. *Operations Research*, 58(3):549–563, 2010.

[35] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics Series, 2003.

[36] G. Cortazar, M. Lopez, and L. Naranjo. A multifactor stochastic volatility model of commodity prices. *Energy Economics*, 67:182–201, 2017.

[37] G. Cortazar and L. Naranjo. An N-factor Gaussian model of oil futures prices. *Journal of Futures Markets*, 26(3):243–268, 2006.

[38] A. d'Aspremont. Identifying small mean-reverting portfolios. *Quantitative Finance*, 11(3):351–364, 2011.

[39] A. Dassios and H. Zhao. A dynamic contagion process. *Advances in Applied Probability*, 43(3):814–846, 09 2011.

[40] M. H. Davis, V. Panas, and T. Zariphopoulou. European option pricing with transaction costs. *SIAM Journal of Control and Optimization*, 31(2), March 1993.

[41] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713, 1990.

[42] F. X. Diebold and C. Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337 – 364, 2006.

[43] S. Dolatabadi, M. O. Nielsen, and K. Xu. A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. *Journal of Empirical Finance*, 38(B):623–639, 2016.

[44] J.-C. Duan and S. Pliska. Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, 28:727–754, 2004.

[45] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, 3 edition, 2001.

[46] C. L. Dunis, P. W. Middleton, A. Karathanasopolous, and K. Theofilatos, editors. *Artificial Intelligence in Financial Markets: Cutting Edge Applications for Risk Management, Portfolio Optimization and Economics*. Palgrave Macmillan, 2016.

[47] R. D. Edwards and J. Magee. *Technical Analysis of Stock Trends*. John Magee, Boston, Mass., 5 edition, 1966.

[48] S. El Aoud and F. Abergel. A stochastic control approach to option market making. *Market Microstructure and Liquidity*, 01(01):1550006, 2015.

[49] G. Elaut, P. Erdos's, and J. Sjodin. An analysis of the risk-return characteristics of serially correlated managed futures. *Journal of Futures Markets*, 36(10):992–1013, 2016.

[50] R. Elliot, J. Hoek, and W. Macolm. Pairs trading. *Quatitative Finance*, 5(3):271–276, 2005.

[51] W. Enders. *Applied Econometric Time Series*. Wiley, 3 edition, 2009.

[52] S. Endres and J. Stübinger. A flexible regime switching model with pairs trading application to the S&P 500 high-frequency stock returns. Technical report, 2018.

[53] R. Engle and C. Granger. Cointegration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276, 1987.

[54] C.-O. Ewald, A. Zhang, and Z. Zong. On the calibration of the Schwartz two-factor model to WTI crude oil options and the extended Kalman filter. *Annals of Operations Research*, Jan 2018.

[55] J. Fink and K. Fink. Do seasonal tropical storm forecasts affect crack spread prices? *Journal of Futures Markets*, 34(5):420–433, 2014.

[56] T. G. Fischer, C. Krauss, and A. Deinert. Statistical arbitrage in cryptocurrency markets. *Journal of Risk and Financial Management*, 12(1), 2019.

[57] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer-Verlag, 1975.

[58] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, 1993.

[59] E. Gatev, W. Goetzmann, and K. Rouwenhorst. Pairs trading: performance of a relative value arbitrage rule. *The Review of Financial Studies*, 19:797–827, 2006.

[60] C. G.Ntim, J. English, J. Nwachukwu, and Y. Wang. On the efficiency of the global gold markets. *International Review of Financial Analysis*, 41:218–236, 2015.

[61] S. Goutte, A. Ismail, and H. Pham. Regime-switching stochastic volatility model: estimation and calibration to VIX options. *Applied Mathematical Finance*, 24(1):38–75, 2017.

[62] W. H. Greene. *Econometric Analysis*. Prentice Hall, 5 edition, 2003.

[63] G. N. Gregoriou, G. Hubner, and M. Kooli. Performance and persistence of Commodity Trading Advisors: further evidence. *Journal of Futures Markets*, 30(8):725–752, 2010.

[64] O. Guéant, C.-A. Lehalle, and J. Fernandez-Tapia. Dealing with the inventory risk: a solution to the market making problem. *Mathematics and Financial Economics*, 7(4):477–507, September 2013.

[65] J. Guijarro-Ordonez. Stochastic control in high-dimensional statistical arbitrage under an Ornstein-Uhlenbeck process. Working paper, 01 2019.

[66] F. Guilbaud and H. Pham. Optimal high-frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94, 2013.

[67] J. Hull. *Options, Futures, and Other Derivatives*. Pearson, 10 edition, 2017.

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

[69] X. Jin and Y. Hou. Optimal investment with default risk. *FAME Research Paper*, (46), 2002.

[70] S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580, 1991.

[71] Y. Kitapbayev and T. Leung. Optimal mean-reverting spread trading: nonlinear integral equation approach. *Annals of Finance*, 13(2):181–203, 2017.

[72] C. Krauss. Statistical arbitrage pairs trading strategies: review and outlook. *IWQW Discussion Paper Series*, (9), 2015.

[73] A. N. Lai and C. Mellios. Valuation of commodity derivatives with an unobservable convenience yield. *Computers and Operations Research*, 66:402–414, 2016.

[74] H. Leland. Optimal portfolio implementation with transactions costs and capital gains taxes. *Working paper, University of California at Berkeley*, 2000.

[75] T. Leung, J. Li, X. Li, and Z. Wang. Speculative futures trading under mean reversion. *Asia-Pacific Financial Markets*, 23(4):281–304, 2016.

[76] T. Leung and X. Li. Optimal mean reversion trading with transaction costs and stop-loss exit. *International Journal of Theoretical & Applied Finance*, 18(3):15500, 2015.

[77] T. Leung and X. Li. *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications*. Modern Trends in Financial Engineering. World Scientific, Singapore, 2016.

[78] T. Leung and B. Ward. The golden target: analyzing the tracking performance of leveraged gold ETFs. *Studies in Economics and Finance*, 32(3), 2015.

[79] T. Leung and B. Ward. Dynamic index tracking and risk exposure control using derivatives. *Applied Mathematical Finance*, 25(2):180–212, 2018.

[80] T. Leung, K. Yamazaki, and H. Zhang. Optimal multiple stopping with negative discount rate and random refraction times under lévy models. *SIAM Journal on Control and Optimization*, 53(4):2373–2405, 2015.

[81] T. Leung and R. Yan. Optimal dynamic pairs trading of futures under a two-factor mean-reverting model. *International Journal of Financial Engineering*, 5(3):1850027, 2018.

[82] T. Leung and R. Yan. A stochastic control approach to managed futures portfolios. *International Journal of Financial Engineering*, 6(1):1950005, 2019.

[83] T. Leung and H. Zhang. Optimal trading with a trailing stop. *Applied Mathematics & Optimization*, pages 1–31, Feb 2019.

[84] T. Leung and Y. Zhou. Dynamic futures portfolio in a regime-switching two-factor framework. working paper, 2019.

[85] J. Li. Trading VIX futures under mean reversion with regime switching. *International Journal of Financial Engineering*, 3(3):1650021, 2016.

[86] T. N. Li and A. Tourin. Optimal pairs trading with time-varying volatility. *International Journal of Financial Engineering*, 03(03):1650023, 2016.

[87] P. Lintilhac and A. Tourin. Model-based pairs trading in the bitcoin markets. *Quantitative Finance*, 17(5):703–716, 2016.

[88] R. B. Litterman and J. Scheinkman. Common factors affecting bond returns. *The Journal of Fixed Income*, 1(1):54–61, 1991.

[89] A. Lo, H. Mamaysky, and J. Wang. Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, 55(4):1705–1765, 2000.

[90] X. Luo and J. E. Zhang. The term structure of VIX. *The Journal of Futures Markets*, 32:1092–1123, 2012.

[91] J. Mencia and E. Sentana. Valuation of VIX derivatives. *Journal of Financial Economics*, 108:367–391, 2013.

[92] R. Merton. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3(4):373–413, 1971.

[93] K. Miltersen. Commodity price modelling that matches current observables: a new approach. *Quantitative Finance*, 3:51–58, 02 2003.

[94] S. Mitra and A. Karathanasopoulos. Firm value and the impact of operational management. *Asia-Pacific Financial Markets*, 26(1):61–85, Mar 2019.

[95] S. Mudchanatongsuk, J. A. Primbs, and W. Wong. Optimal pairs trading: a stochastic control approach. In *Proceedings of the American Control Conference*, pages 1035–1039, 2008.

[96] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 6 edition, 2014.

[97] B. Oksendal and A. Sulem. Optimal consumption and portfolio with both fixed and proportional transaction costs. *SIAM Journal on Control and Optimization*, 40(6):1765–1790, 2002.

[98] B. Oksendal and A. Sulem. *Applied Stochastic Control of Jump Diffusions*. Springer, 2005.

[99] A. Papanicolaou and R. Sircar. A regime-switching heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance*, 14(10):1811–1827, 2014.

[100] H. Pham and Y. Jiao. Optimal investment with counterparty risk: a default-density modeling approach. *Finance and Stochastics*, 15(4):725–753, 2011.

[101] H. Pham and M.-M. Ngo. Optimal switching for the pairs trading rule: a viscosity solutions approach. *Journal of Mathematical Analysis and Applications*, 441(1):403–425, 2016.

[102] P. C. B. Phillips and S. Ouliaris. Asymptotic properties of residual based tests for cointegration. *Econometrica*, 58(1):165–193, 1990.

[103] P. Protter. *Stochastic Integration and Differential Equations.* Springer, 2 edition, 2003.

[104] A. Rencher. *Methods of Multivariate Analysis.* John Wiley & Sons, 2 edition, 2002.

[105] K. Ross. *Stochastic Control in Continuous Time.* Lecture Notes on Continuous Time Stochastic Control, 2008.

[106] W. Schoutens. *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley, 2003.

[107] E. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52(3):923–973, 7 1997.

[108] A. Sepp. VIX option pricing in a jump-diffusion model. *Risk Magazine*, pages 84–89, April 2008.

[109] D. Simon and J. Campasano. The VIX futures basis: evidence and trading strategies. *Journal of Derivatives*, 21(3):54–69, 2014.

[110] D. P. Simon. The soybean crush spread: empirical evidence and trading strategies. *Journal of Futures Markets*, 19(3):271–289, 1999.

[111] K. Simonsen. Mean reversion and first passage times in relative value trading: the Ornstein-Uhlenbeck process. *ABN AMRO Fixed Income Relative Value Research*, 2003.

[112] E. Sinclair. *Volatility Trading.* John Wiley & Sons, 2008.

[113] S. Stoikov and M. Sağlam. Option market making under inventory risk. *Review of Derivatives Research*, 12(1):55–79, Apr 2009.

[114] J. Stone. *Independent Componenet Analysis.* The MIT Press, 2004.

[115] C. Y. Tang and S. X. Chen. Parameter estimation and bias correction for diffusion processes. *Journal of Econometrics*, 149:65–81, 04 2009.

[116] A. Tourin and R. Yan. Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control*, 37:1947–2156, 2013.

[117] A. Tourin and T. Zariphopoulou. Viscosity solutions and numerical schemes for investment/consumption models with transaction costs. In L. C. G. Rogers and D. Talay, editors, *Numerical methods in finance*, pages 245–269. Cambridge University Press, 1997.

[118] A. Trolle and E. Schwartz. Unspanned stochastic volatility and the pricing of commodity derivatives. *The Review of Financial Studies*, 22(11):4423–4461, 2009.

[119] L. A. M. Veraart. Optimal investment in the foreign exchange market with proportional transaction costs. *Quantitative Finance*, 11(4):631–640, 2011.

[120] G. Vidyamurthy. *Pairs Trading: Quantitative Methods and Analysis.* Wiley Finance, 2004.

[121] J. Wachter. Portfolio and consumption decisions under mean-reverting returns: an exact solution for complete markets. *Journal of Financial and Quatitative Analysis*, 37(1):63–91, 2002.

[122] R. Walder. Dynamic allocation of treasury and corporate bond portfolios. *FAME Research Paper*, (64), 2002.

[123] E. A. Wan and R. van der Merwe. *The Unscented Kalman Filter*, chapter 7, pages 221–280. John Wiley & Sons, 2002.

[124] W. Xiong. Convergence trading with wealth effects: an amplification mechanism in financial markets. *Journal of Financial Economics*, 62:247–292, 2001.

[125] R. Yamamoto and N. Hibiki. Optimal multiple pairs trading strategy using derivative free optimization under actual investment management conditions. *Journal of the Operations Research Society of Japan*, 60:244–261, 2017.

[126] Z. X. Y. Yong, Jiongmin. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, 1999.

[127] T. Zariphopoulou. Optimal investment and consumption models with non-linear stock dynamics. *Mathematical Methods of Operations Research*, 50:271–296, 1999.

[128] J. Zhang, T. Leung, and A. Aravkin. Mean reverting portfolios via penalized maximum likelihood estimation and optimization. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, pages 5795–5800, 2018.