OPTIMAL PAIR-TRADING DECISION RULES FOR A CLASS OF NON-LINEAR BOUNDARY CROSSINGS BY ORNSTEIN-UHLENBECK PROCESSES

Emmanuel Edem Kwaku Tamakloe, BEd., MS

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2021

APPROVED:

Kai-Sheng Song, Major Professor
Joseph Iaia, Committee Member
Nam Trang, Committee Member
Ralf Schmidt, Chair of the Department of Mathematics
Pamela Padilla, Dean of the College of Science
Victor Prybutok, Dean of the Toulouse Graduate School

Tamakloe, Emmanuel Edem Kwaku. *Optimal Pair-Trading Decision Rules for a Class of Non-Linear Boundary Crossings by Ornstein-Uhlenbeck Processes*. Doctor of Philosophy (Mathematics), December 2021, 147 pp., 51 tables, 83 figures, 1 appendix, 40 numbered references.

The most useful feature used in finance of the Ornstein-Uhlenbeck (OU) stochastic process is its mean-reverting property: the OU process tends to drift towards its long-term mean (its equilibrium state) over time. This important feature makes the OU process arguably the most popular statistical model for developing best pair-trading strategies. However, optimal strategies depend crucially on the first passage time (FPT) of the OU process to a suitably chosen boundary and its probability density is not analytically available in general. Even for crossing a simple constant boundary, the FPT of the OU process would lead to crossing a square root boundary by a Brownian motion process whose FPT density involves the complicated parabolic cylinder function. To overcome the limitations of the existing methods, we propose a novel class of non-linear boundaries for obtaining optimal decision thresholds. We prove the existence and uniqueness of the maximizer of our decision rules. We also derive simple formulas for some FPT moments without analytical expressions of its density functions. We conduct some Monte Carlo simulations and analyze several pairs of stocks including Coca-Cola and Pepsi, Target and Walmart, Chevron and Exxon Mobil. The results demonstrate that our method outperforms the existing procedures.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my major professor, Dr. Kai-Sheng Song, for his patience and unrelenting support that has brought me this far. I appreciate the wealth of knowledge and experience with which he guided this research. He indeed went out of his way to make sure this project was successful and I cannot thank him enough.

I would also like to thank my committee members, Dr. Joseph Iaia and Dr. Nam Trang, for their various contributions to my learning process. In addition, I would like to again thank Dr. Nam Trang and the Department of Mathematics of the University of North Texas for providing me with a high-performing laptop when I needed one to perform my research.

My gratitude also goes out to my entire family back home in Ghana for their support and motivation in my academic pursuit.

I am thankful to my wife, Patience Tamakloe, for her support in this journey and I dedicate this work to our first son, Paul Elikplim Kojo Tamakloe.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

With the introduction of algorithmic trading and high frequency trading in the late 1980s, trading in the financial market has transformed over the years, and this is true for all forms of trading strategies, including arbitrage.

1.1. Pairs Trading

Various trading strategies are employed in financial markets, one of which is arbitrage. Billingsley defined arbitrage as "the process of buying assets in one market and selling them in another to profit from unjustifiable price differences" [6]. While this is true, the assets need not be in different markets. An arbitrage exists as long as there is a deviation in price, which is expected to close up over time. There are various forms of arbitrage, some of which have been identified in [6]. One form of arbitrage is the statistical arbitrage. Goncu and Akyildirim defined a statistical arbitrage as follow:
A statistical arbitrage is a zero initial cost, self-financing trading strategy $\{v(t) : t \geq 0\}$ with cumulative discounted value $v(t)$ such that

(1) $v(0) = 0$

(2) $lim_{t\to\infty} E[v(t)] > 0$,

(3) $lim_{t\to\infty} P(v(t) < 0) = 0$, and

(4) $lim_{t\to\infty} \frac{var(v(t))}{t} = 0$ if $P(v(t) < 0) > 0, \forall t < \infty$ [17]

"A statistical arbitrage refers to trading strategies that generate almost sure profits asymptotically via trading signals generated from quantitative models" [17]. These models are usually mean-reversion models. and the assets are usually short-term financial instruments. Pairs trading is a form of statistical arbitrage. It is widely assumed to be the "ancestor" of statistical arbitrage [4].

In its most common form, pairs trading involves forming a portfolio of two related stocks whose relative pricing departs from its "equilibrium" [20]. The idea of pairs trading is based on the assumption that if the prices of a pair of financial instruments, for instance

stocks, moved together in the past, then this behavior is likely to continue in the future [40]. The strategy employs a lot of quantitative and computational techniques to realize results.

1.2. Cointegration

It is well known that asset price time series are generally nonstationary. In conformity with the efficient market hypothesis, they exhibit unit-root property. Thus the current price is the best predictor of the next price [15]. However, there exists co-movements among prices of different assets. As a result, one may find a linear combination of the asset prices, which is stationary. This idea lends itself to the concept of cointegration, which was documented in [7]. The term cointegration was later coined by Granger in [18].

Let X_t be an asset price. It is said to be of integration order zero if it is stationary, and we denote that by $X_t \sim I(0)$. If X_t is nonstationary but $\nabla X_t := X_t - X_{t-1}$ is stationary, then X_t is said to be of integration order 1. In this case, we say that X_t has a unit root and is denoted by $X_t \sim I(1)$. Now, suppose we have two related stocks P and Q. Let P_t and Q_t be the time series of their prices. Suppose they both have integration order 1. If there exists a non-zero constant η such that $P_t - \eta Q_t \sim I(0)$, then P_t and Q_t are said to be cointegrated [15]. Thus there exist a linear combination of the two time series whose integration order is zero. In other words this linear combination is stationary.

We can model the cointegration of P_t and Q_t as:

$$ln(P_t) - ln(P_{t_0}) = \alpha(t - t_0) + \eta[ln(Q_t) - ln(Q_{t_0})] + \epsilon_t, \quad t \geq 0, \tag{1.1}$$

where α is the drift rate, η is some constant and ϵ_t is a stationary or mean-reverting process [4].

1.3. Mean Reversion Models

As stated in the previous section, the random variable ϵ_t in equation (1.1) is assumed to be stationary, or mean reverting. A mean-reversion process is a stochastic process that tends to revert toward its equilibrium position or long-term mean, or a long-term trend, whenever there is a deviation from this position or trend. Thus the long-term mean or trend

acts as an attractor and together with the random component makes the process oscillate around this mean or trend [26].

The long-term trend may be deterministic or stochastic. A one-dimensional mean reversion process with deterministic long-term trend and constant parameters can be modeled as:

$$dX_t = \lambda(\mu(t) - X_t)dt + \sigma X_t^{\gamma} dB_t; \quad t \in [0, T] \tag{1.2}$$

where $X_0 = x$ is the initial position, the constant $\lambda > 0$ is the reversion rate, the constant $\sigma > 0$ is the scale parameter for the volatility and the constant $\gamma \in [0, 3/2]$ is the sensitivity parameter of the variance of the process to the level of X_t. The deterministic function $\mu(t)$ is the long-term trend, and $\{B_t\}_{t\geq 0}$ is a one-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ [26].

Some examples of the above model with constant long-term mean include:

(1) Vasicek / Ornstein-Uhlenbeck Process (OU)

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t \tag{1.3}$$

(2) Cox-Ingersoll-Ross Square Root model (CIR SR)

$$dX_t = \lambda(\mu - X_t)dt + \sigma\sqrt{X_t} dB_t \tag{1.4}$$

(3) Brennan-Schwartz model

$$dX_t = \lambda(\mu - X_t)dt + \sigma X_t dB_t \tag{1.5}$$

Some examples with time-dependent long-term trend

(4) Hull-White model

$$dX_t = \lambda(\mu(t) - X_t)dt + \sigma dB_t \tag{1.6}$$

For (1), see [35]. (2) can be found in [12]. Michael Brennan and Eduardo Schwartz constructed (3) in [9]. Also, a summary of these can be found in [11] and [34].

1.4. Ornstein-Uhlenbeck Processes

Given two stocks P and Q and their price time series P_t and Q_t as defined in section 1.2, we model the spread of their log-returns by the cointegration model (1.1), where the random component ϵ_t is assumed to follow an Ornstein-Uhlenbeck process [4, 40, 17].

We may or may not ignore the drift rate α. Both [17] and [40] suggest that it is usually ignorable compared to the fluctuations of the residual ϵ_t, and indeed did not include it in their models. Following their approach, we define $X_t := \epsilon_t + ln(P_0) - \eta ln(Q_0)$.

Thus the cointegration model can therefore be represented as:

$$X_t = ln(P_t) - \eta ln(Q_t) \tag{1.7}$$

We note that since $ln(P_0) - \eta ln(Q_0)$ is constant, then X_t is also an Ornstein-Uhlenbeck(OU) process. Thus it satisfies equation (1.3), which is stated again here for completeness.

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t$$

The interpretation of equation (1.7) together with the above equation is that at some time $t_* > 0$ where $X_{t_*} \neq \mu$, the trader may initiate a trade by doing one of the following:

(I) If $X_{t_*} > 0$, then at time t_*, asset P is relatively overvalued in comparison with asset Q and hence its current price is relatively above the long-term equilibrium price. Thus the return on asset P as the values of both assets shift towards the long-term equilibrium is expected to decrease while the return on asset Q will increase since it is relatively undervalued. In this case, at time t_*, the trader will short 1 dollar of asset P and long η dollars of asset Q, and then clear position when X_t reaches the long-term mean μ.

(II) If $X_{t_*} < 0$, then at time t_*, asset P is relatively undervalued in comparison with asset Q and hence its current price is relatively below the long-term equilibrium price. Thus the return on asset P as the values of both assets shift towards the long-term equilibrium is expected to increase while the return on asset Q will decrease since it is relatively overvalued. For this case, at time t_*, the trader will long 1 dollar of asset P

and short η dollars of asset Q, and then clear position when X_t reaches the long-term mean μ.

There are various ways of deciding a value for η. Some common ones are as follows:

(1) As a linear regression coefficient of the logarithm of one asset price time series against the other [40].

(2) $\eta = log(P_{t_*}/Q_{t_*})$, also known as the no borrowing/lending case, since the amount for the long and short positions offset each other [17].

(3) $\eta = \beta_P/\beta_Q$, where the β's are obtained from the market at the beginning of each trade cycle [17].

1.5. The Generalized Ornstein-Uhlenbeck Process

Suppose we choose to maintain the drift rate α in the cointegration. Then we will define X_t by $X_t := \alpha(t - t_0) + \epsilon_t + ln(P_0) - \eta ln(Q_0)$, where the residual ϵ_t follows a mean reverting OU process as before. In this case, the long-term equilibrium behavior of X_t has a linear trend coming from the term $\alpha(t - t_0)$.

The cointegration model again becomes

$$X_t = ln(P_t) - \eta ln(Q_t),$$

$-\alpha t_0 + ln(P_0) - \eta ln(Q_0)$ is constant, but X_t follows a trending Ornstein-Uhlenbeck process (Also called the trend-stationary Ornstein-Uhlenbeck process). The stochastic differential equation for this model is represented as:

$$d(X_t - \mu t) = -\lambda(X_t - \mu t)dt + \sigma dB_t \tag{1.8}$$

or

$$dX_t = (\mu - \lambda(X_t - \mu t))dt + \sigma dB_t \quad [34] \tag{1.9}$$

Similar to the OU process case, at some time $t_* > 0$ where $X_{t_*} \neq \mu t_*$, the trader takes position as discussed in I and II above, and clears position when X_{t_*} crosses the long-term equilibrium trend.

1.6. The Problem Statement

At this point, the question that comes to mind is how does one determine the time t_* at which to take position? This is among the questions we seek to address in this dissertation. In general, the way this is done is to find an "appropriate" level of the spread X_t to start the trade cycle and the first time this level is crossed will be our choice for t_*. We first explore various literature on this problem in the next chapter.

CHAPTER 2

OVERVIEW AND LIMITATIONS OF EXISTING METHODS

As explained in the statement of the problem, pairs trading strategy is about choosing the appropriate time to enter a trade position and when to exit, according to the procedure explained in section 1.4. The question of choosing the appropriate time to enter trade position also corresponds to that of choosing an appropriate spread level at which to enter the trade position. Thus we initiate trade the first time the level is crossed. So the problem can be framed in the context of first passage time (FPT), which is discussed in the next section.

2.1. First Passage Time

The first passage time (FPT) of a stochastic process to a boundary is the amount of time it takes the process to reach the boundary for the first time, given its initial value. Formally defined as follows:

Given a stochastic process $\{X_t\}_{t\geq 0}$ with $X_0 = x_0$, and a real-valued function $b(t)$ defined on $t \geq 0$, such that $b(0) \geq x_0$, we define the first passage time (FPT) of X_t to $b(t)$ by,

$$\tau_b := \inf\{t > 0 : X_t \geq b(t) | X_0 = x_0\}$$

The probability distribution and density functions of the first passage time of various stochastic processes are very much studied because of their wide range of applications in various fields. A few of these applications besides pairs trading are listed below;

i In biology, first passage time densities are used in the study of stochastic birth-death model for cell populations [21].

ii In chemical physics, it is used for instance in studying models for the dissociation of diatomic molecules in which dissociation occurs when the molecules acquire a certain critical energy E_c through collisions [37].

iii The most prevalent use is found in quantitative finance, where first passage times are used in credit risk analysis (times of default) as well as in defining exotic contingent claims (so-called barrier options) [14].

Despite the numerous applications, expressions of first passage time densities are known only in very few specific cases. These include Brownian motions to constant boundaries, linear boundaries, square root boundaries [8, 29] and square boundaries [19], and Ornstein-Uhlenbeck (OU) processes to constant boundaries [3]. There are also results for certain transformations of these boundaries. See for instance [2]. However, with the exception of the constant and the linear Brownian motion boundaries, the rest are all in terms of infinite series of some advanced mathematical functions, such as the parabolic cylinder function, the Hermite polynomial and the Airy function [1], which makes them unwieldy.
In the case of Brownian motions, one common method by which results are obtained involves solving Kolmogorov's differential equation to obtain the transition density of the Brownian motion and using the result to find the first passage time density [13]. Another approach is through what is known as the method of images. It involves solving an implicit equation that involves the integral of some exponential function with respect to a positive σ-finite measure, and then the first passage time density is defined as some function of the solution [25, 23].

For Ornstein-Uhlenbeck processes, the case of the constant boundary corresponds to the square root boundary of the Brownian motion [3]. Alili et al showed how this is achieved via Doob's transform [36].

In the absence of general analytic expressions for the first passage time density or distribution, one has to resort to computational methods, which is a common practice. Some numeral methods used are based on integral equations, such as the Volterra integral equations [31, 10] and the Fredholm equations [23, 33].

There are also other numerical methods that rely on Monte Carlo simulations [22].

2.2. Pairs Trading Strategies

The random component of the cointegration model for pairs trading follows an Ornstein-Uhlenbeck process. Hence in deciding the optimal level for taking position, most of the strategies rely on the first passage time of the Ornstein-Uhlenbeck process to a one-sided or two-sided boundary, and/or the first passage time of the Ornstein-Unlenbeck process from a boundary to the long term mean. We look into some of these methods next. The methods differ in two ways: how the optimal level or boundary is chosen and how the trade is implemented.

2.2.1. Conventional Method

In the conventional method of pairs trading, the spread between the log-returns of the prices of the pair of assets is modeled with the cointegration relation. The random part is assumed to follow a mean-reverting stochastic process, such as the OU process. The optimal buy/sell thresholds are determined with the appropriate mathematical and statistical technique and a buy or sell position is taken when the set threshold is reached and the trader waits for the process to return to the long-term mean in order to clear position, and this cycle is repeated throughout the trading time horizon. The optimal levels are usually determined by the two standard deviation rule [5]. However, other levels of standard deviation may be used, see for instance [4].

2.2.2. Continuous Time Trading

> A continuous trading strategy comprises a sequence of individual trades performed on a continuous time stochastic process [5].

Bertram also added that "a continuous time trading strategy is defined by entering a trade when $X_t = a$, exiting the trade at $X_t = m$, and waiting until the process returns to $X_t = a$, to complete the trading cycle". The difference here is that the exit level is not necessarily the long-term mean. Bertram expressed the return as a function of the entry level a, exit level b and the transaction cost. The optimal thresholds are found by solving for the entry and exit levels that maximize the expected return as well as the Sharpe ratio.

2.2.3. Zeng and Lee's Strategy

Zeng and Lee presented an alternative method in which they sought to find optimal entry and exit thresholds that maximize the expected profit per unit time in the long run [40]. Similar to [5], the exit level is not necessarily the long term mean. Using the elementary renewal theorem, they were able to obtain an expression for the expected profit per unit time in terms of the entry level, the exit level, and the transaction cost. Then applying the relevant optimization techniques, they obtained implicit expressions for the thresholds. One of the cases they considered reduced to the conventional method, in that the exit level matched with the long term mean, while the other two cases they considered resulted in their "New Optimal Rule". They noted that when there is no transaction cost, the maximal return of the new rule is the same as the maximal return of the conventional method. But the new rule outperforms the conventional rule if the transaction cost is greater than zero.

2.2.4. Goncu and Akyildirim's Strategy

Another strategy, which is more like the conventional method in terms of exit level, is presented by [17]. The objective is to find the optimal threshold that maximizes the probability of successful trade within a given time horizon. Successful trade in this context is defined as successful mean reversion and closing of spread position with the given time. They used the first passage time density of the Ornstein-Uhlenbeck process from a position above the long term mean to the long term mean in deriving the optimal threshold for the strategy. Although transaction cost was considered in the paper, it does not affect the optimization technique use in this strategy.

2.3. Limitations of Existing Methods

We have seen some useful pairs trading strategies in the previous section. As useful as they are, they do have some limitations which we discuss in this section.

2.3.1. The Cointegration Drift Rate

As pointed out earlier on, a common assumption of the strategies discussed in the previous section is that the drift rate α in the cointegration model (1.1) is often negligible

and therefore ignored. While this is true, there are several empirical cases where we found out that maintaining the drift rate was necessary to ensure that the pair of asset price time series were cointegrated, thus (1.7) represents a stationary process. We present some examples below. As we see in figure 2.1, the residual plot for the cointegration between logarithmic returns of Pepsi stock price time series and Coca Cola stock price time series from September 16th, 2008 to July 7th, 2009 shows that a trend is present. This shows time dependency, hence nonstationarity in the spread, and therefore suggests that the logarithmic returns of the two stock prices over the given period may not be really cointegrated. In this example, η is taken to be the coefficient of linear regression. On the other hand, considering a stationarity test, such as the augmented Dickey-Fuller test, for the same data set, as presented in table 2.1, the p-values suggest stationarity in the residuals with or without trend. The "drift" used in the table refers to the vertical intercept of the line, while the "trend" refers to the rate of change of the line [30].

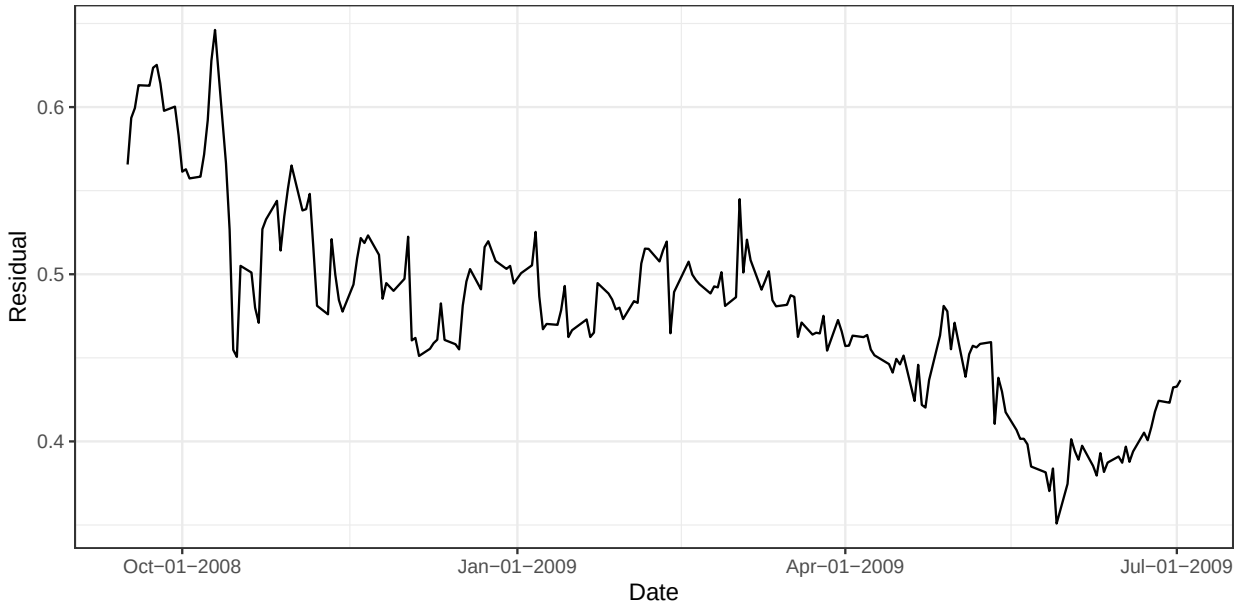

FIGURE 2.1. Cointegration between logarithmic returns of Pepsi and Coca Cola stock prices from September 16, 2008 to July 7, 2009

Linear model type	Lag	ADF	p-value
Type 1: no drift, no trend	0	-2.6	≤ 0.01
Type 2: with drift, no trend	0	-2.6	0.0977
Type 3: with drift and trend	0	-4.39	≤ 0.01

TABLE 2.1. Augmented Dickey-Fulller test for stationarity of the cointegration between Pepsi and Coca Cola log-returns.

Note 1: Alternate hypothesis: Stationary

Notwithstanding, if we consider the augmented Dickey-Fuller test statistic presented in the table, we notice that the model with drift and trend (Type 3) has significantly lower test statistic than the model with no drift nor trend. This may be worth considering in choosing whether or not to maintain the drift rate α in the cointegration model. We will explore this further in chapter 3.

Additionally, we also noticed that out of sample prediction was better in some cases when we involve the drift rate. It also provides more crossings and thus has the potential to generate more trade opportunities.

2.3.2. Constant Boundary

The strategies we have discussed so far all employ a constant (non-time-dependent) threshold. To the best of our knowledge, no literature on pairs trading strategies employs any form of time dependent threshold.

Since the residual in the cointegration model is assumed to follow a mean-reverting process, one would expect that over time there would be a reduction in the deviation from the long term mean or trend. This is particularly true for short-term trades as evidenced in the examples to follow. We pick four pairs of stocks that are considered to exhibit similar patterns and were used by [40]; [17] in their papers.

Thus with a constant threshold, there is a high chance that at the early part of the time horizon, the trader may enter trade position at a level too close to the long term mean and hence miss some profit and/or miss trade opportunities toward the end of the time

horizon due to reduction in deviation from the long term mean over time. It is worth noting that this phenomenon is very common, but we only picked a few instances for each of the four pairs to make our point as shown in figures 2.2, 2.3, 2.4 and 2.5.

2.3.3. Zeng and Leng's Strategy and Transaction Cost

Except for Goncu and Akyildirim's strategy, most known pairs trading strategies depend on transaction cost. In fact Zeng and Lee stated in their paper that at zero transaction cost their result yield the same maximal return as the conventional method [40]. On top of this we found out that a zero transaction cost would mean a threshold of zero for Zeng and Lee's method. This can be checked by solving equation 80. However, trades nowadays have either zero or close to zero transaction cost, which means Zeng and Lee's strategy would not be applicable in such circumstances.

2.3.4. A Discussion on the Method of Goncu and Akyildirim

As stated earlier Goncu and Akyildirim's method is based on the investment time horizon, T. The function they sought to maximize is:

$$P(\tau < T) = \int_0^T \sqrt{\frac{2}{\pi}} \frac{|c|e^{-t}}{(1-e^{-2t})^{3/2}} exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right) dt$$

We show a plot of this function for the case of $T = 0.5$ in figure 2.6. It is clear from this plot that the optimal value of the function is obtained when c is approximately equal to zero, which makes intuitive sense, in that the closer the threshold is to the long term mean, the higher the chance of reaching it within the time horizon T. But this would mean their threshold is essentially zero, hence provides no trade opportunity in the given time horizon. So, the strategy fails. Besides this, we also have concerns regarding the optimization technique employed by the authors in arriving at the optimal expression for level c. The approach does not guarantee in general that $P(\tau)$ is a probability function for the resulting value of c. We also note that for large time horizon T, say $T = 2$, the resulting thresholds for their strategy are too far from the spread to yield any trades at all. We show an example in figure 2.7 for 2-year time horizon for the German utility companies EOAN and RWE.

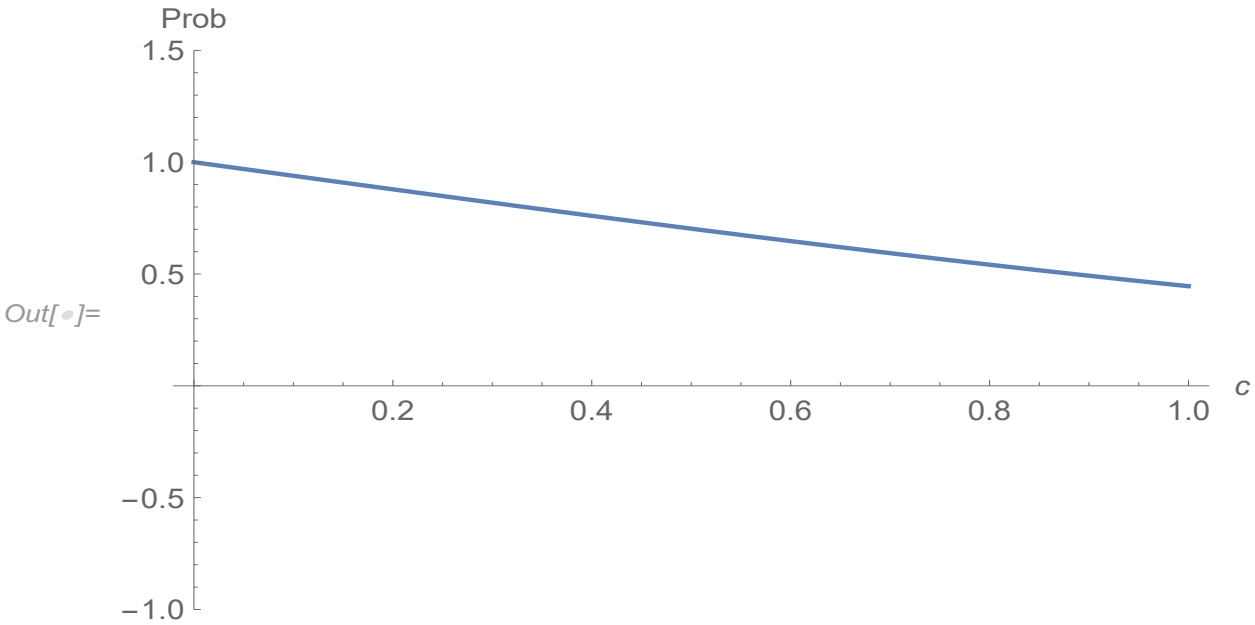

FIGURE 2.6. Objective function of Goncu and Akyildirim for time horizon $T = 0.5$

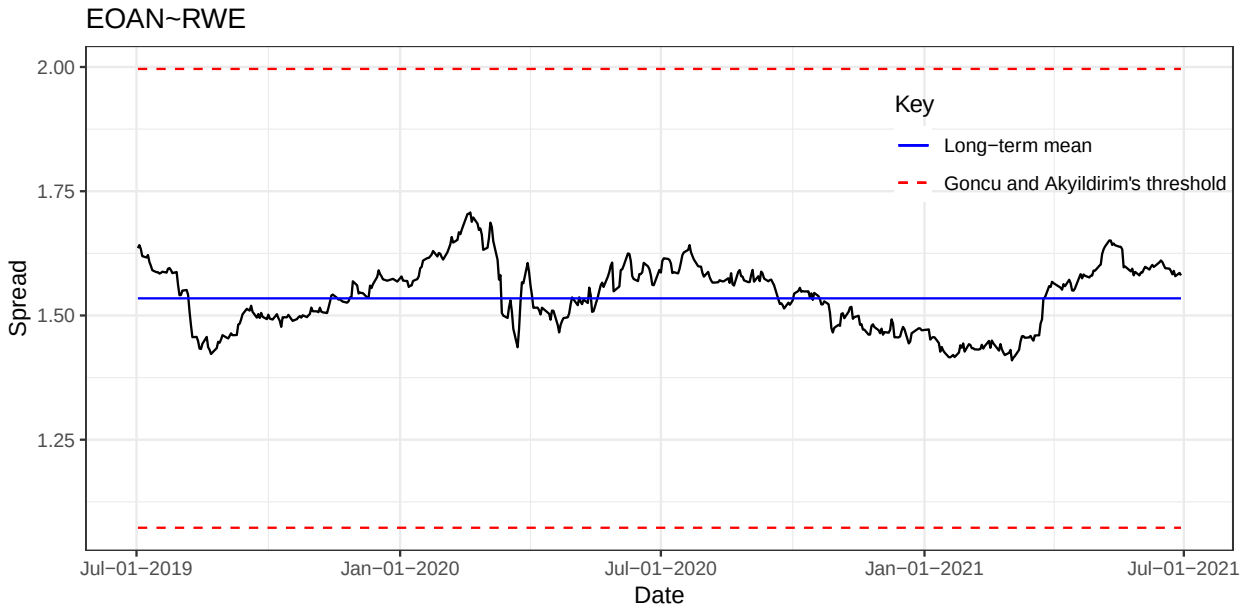

FIGURE 2.7. Goncu and Akyildirim's threshold for 2 year time horizon pairs trading of EOAN.DE and RWE.DE

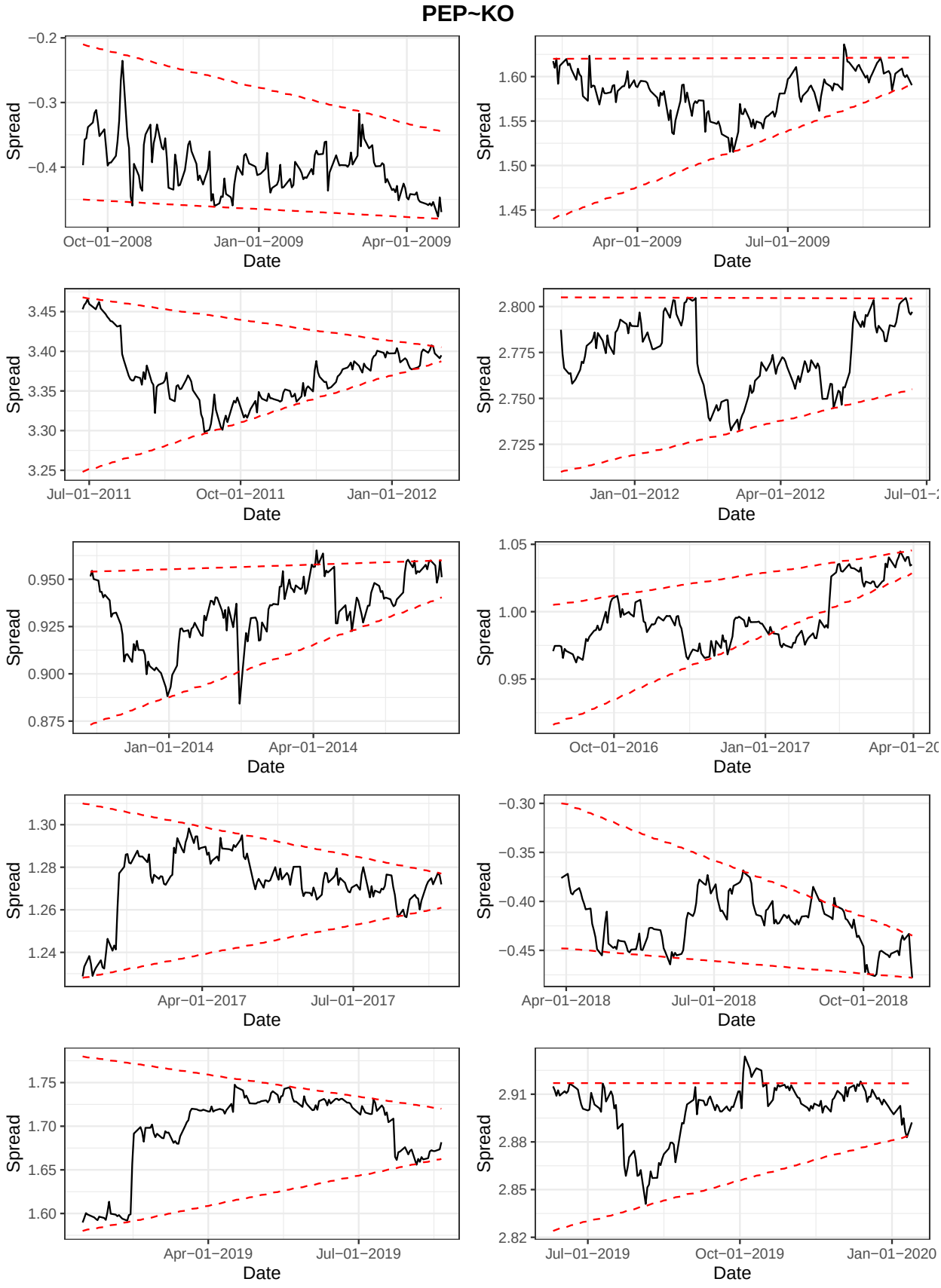

FIGURE 2.2. Spreads between logarithmic returns of PEP and KO stock prices for various half year periods

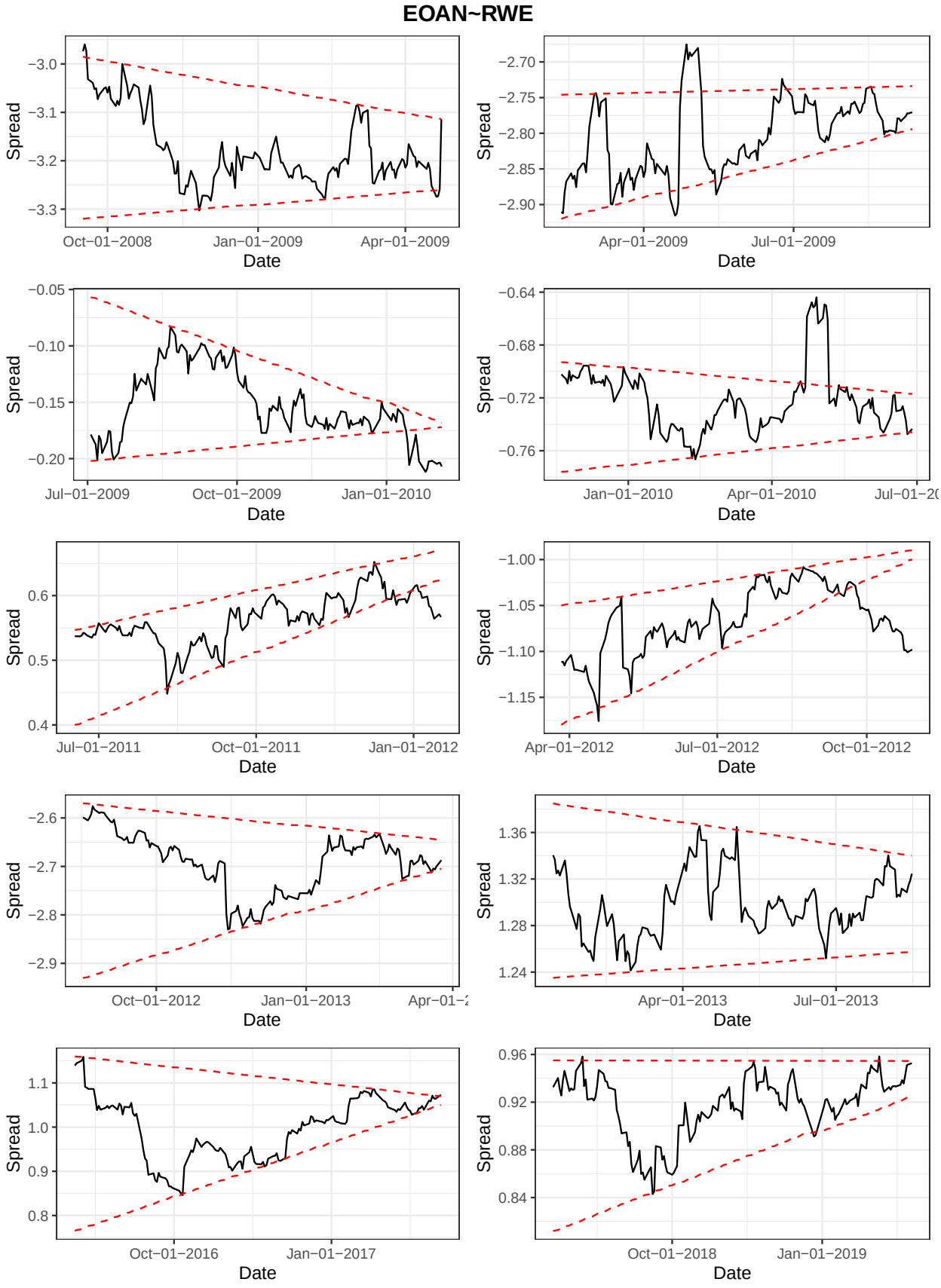

FIGURE 2.3. Spreads between logarithmic returns of EOAN and RWE stock prices for various half year periods

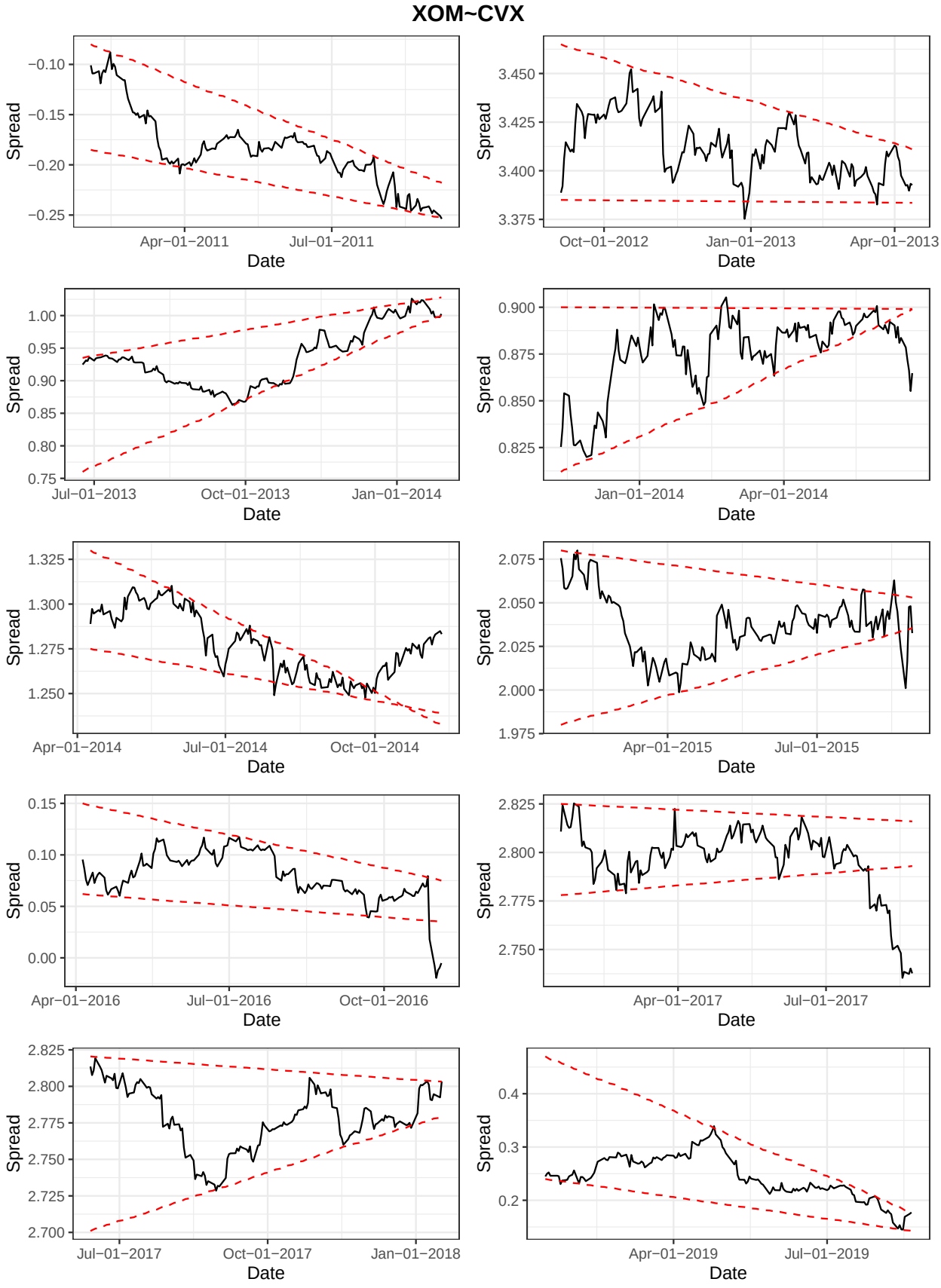

FIGURE 2.4. Spreads between logarithmic returns of XOM and CVX stock prices for various half year periods

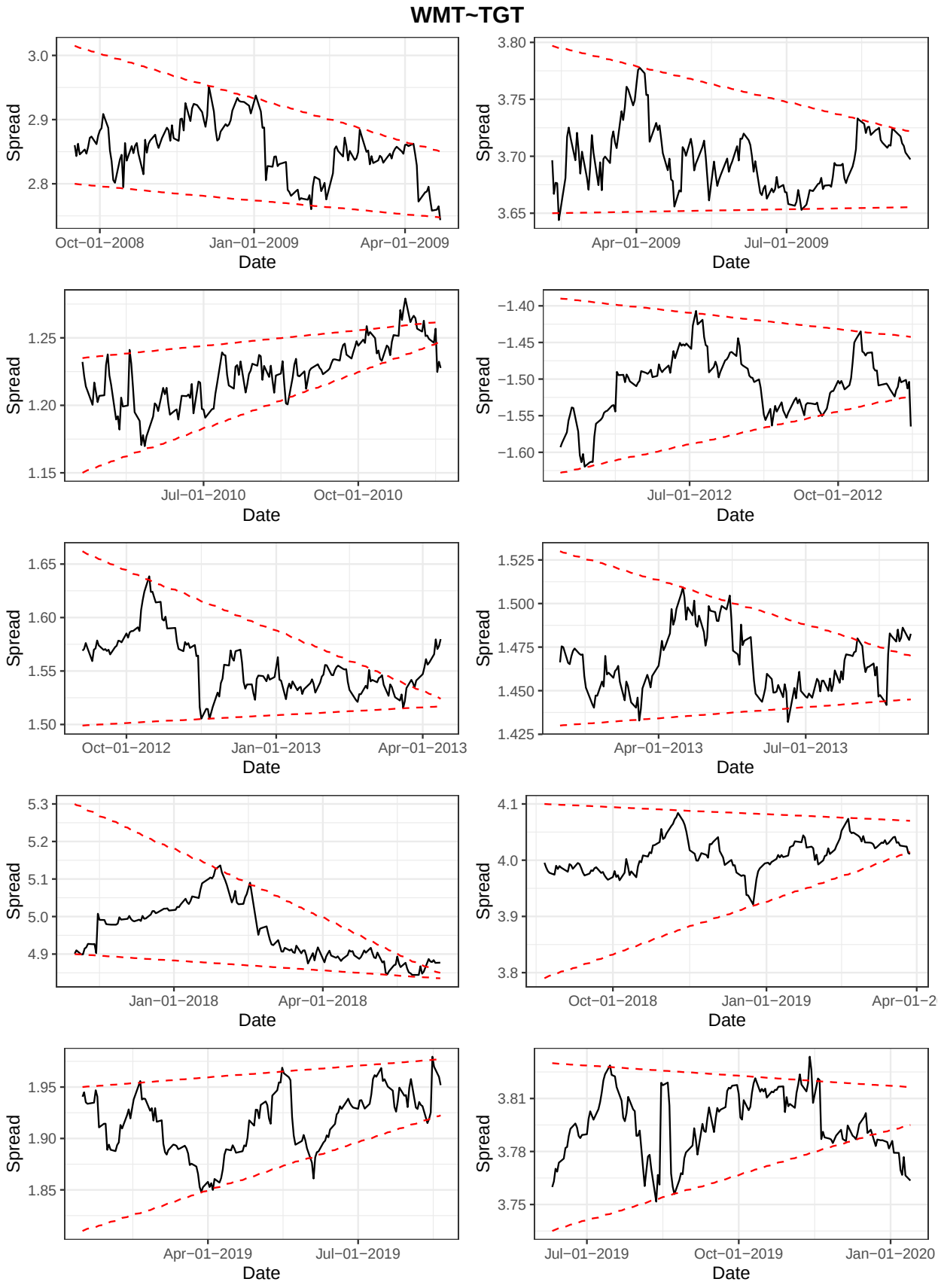

FIGURE 2.5. Spreads between logarithmic returns of WMT and TGT stock prices for various half year periods

CHAPTER 3

METHODOLOGY

From the literature reviewed in chapter 2, it is clear that the application of first passage time density in deriving optimal thresholds for pairs trading strategies is not completely new. However, existing strategies only consider constant thresholds and also overlook trend in the data by ignoring the drift rate in the cointegration model.

3.1. The Drift Rate Treatment

Our study will be in two folds; one in which we will consider the drift rate α and another in which we follow existing methods and ignore the drift rate. In the performance analyses, we will be comparing the returns in both cases against the some recent pairs trading strategies.

3.2. New Pair Trading Strategy

From figures 2.2, 2.3, 2.4 and 2.5, we see that constant boundaries are not always appropriate, if one wants to make maximum returns from pairs trading, especially for pairs for which the deviations reduce over the time horizon.

To address this issue, we present a new boundary that takes into account the mean-reversion rate of the Ornstein-Uhlenbeck process and as such drifts toward the long-term mean or trend over time.

We will study two cases of this threshold and come up with objective functions that we will use to obtain our optimal thresholds for the pairs trading strategy. We will show the existence and uniqueness of optimizers. This will be done by exploiting the association between OU processes and the standard Brownian motion, as well as their boundary crossing probabilities.

3.3. Performance

We will first test our strategy on an artificial pair of stocks, which we denote by P and Q, in chapter 5. We will then proceed in chapter 6 to recent real data collected for the

four pairs of stocks, Coca-Cola (KO)/Pepsi (PEP), (CVX)/(XOM), Target (TGT)/Walmart (WMT), and RWE AG (RWE.DE)/E.OnSe (EOAN.DE). We will consider both long-term and short-term performance of our optimal thresholds, both with and without trend and compare the returns against those obtained from two recent strategies from [40] and [17] discussed in Chapter 2.

3.4. Result on Certain Moments of First Passage Time

Another result that we present in this dissertation is about certain moments of the first passage time of the standard Brownian motion to some class of boundaries. First passage time probabilities and densities for Brownian motions are not known in most cases. But it is possible to obtain certain moments of the first passage time without knowing explicitly the corresponding probabilities or densities. We derive a simple formula that provides these moments for several cases of the standard Brownian motion boundaries. We will also show how this formula extends to the OU process.

CHAPTER 4

THEORETICAL RESULTS

4.1. Relationship between Boundary Crossing Probabilities of Ornstein-Uhlenbeck Processes and Brownian Motion

From previous chapters, it is apparent that Ornstein-Uhlenbeck processes, as well as other diffusion processes are related to Brownian motions.

We consider the Ornstein-Uhlenbeck process,

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t, \quad X_0 = x_0 \tag{1}$$

This process can be standardized as follows:

Let $\tilde{t} = \lambda t$ and $Z_{\tilde{t}} = (X_t - \mu)/\sqrt{\frac{\sigma^2}{2\lambda}}$

Thus,

$$X_t - \mu = \frac{\sigma}{\sqrt{2\lambda}} Z_{\tilde{t}}$$

$$dX_t = \frac{\sigma}{\sqrt{2\lambda}} dZ_{\tilde{t}}$$

and,

$$dt = \frac{1}{\lambda} d\tilde{t}$$

and by the scaling property of Brownian motion, $B_{\tilde{t}} = \sqrt{\lambda} B_t$.

Thus $dB_{\tilde{t}} = \sqrt{\lambda} dB_t$. Hence,

$$\frac{\sigma}{\sqrt{2\lambda}} dZ_{\tilde{t}} = -\lambda \frac{\sigma}{\sqrt{2\lambda}} Z_{\tilde{t}} \frac{1}{\lambda} d_{\tilde{t}} + \sigma \frac{1}{\sqrt{\lambda}} dB_{\tilde{t}}$$

$$dZ_{\tilde{t}} = -Z_{\tilde{t}} d_{\tilde{t}} + \sqrt{2} dB_{\tilde{t}}, \tag{2}$$

which is the standardized OU process, and is also called the dimensionless OU process, since $Z_{\tilde{t}}$ is not dependent on the parameters of the original OU process.

Similarly, for the generalized (trend-stationary) OU process with linear trend, which satisfies the SDE

$$d(X_t - (at+b)) = -\lambda(X_t - (at+b))dt + \sigma dB_t, \quad t \geq 0 \tag{3}$$

$$X_0 = x_0, \quad \lambda > 0, \quad \sigma > 0.$$

let $\tilde{t} = \lambda t$ and $Z_{\tilde{t}} = (X_t - (at+b))/\sqrt{\frac{\sigma^2}{2\lambda}}$

Thus,

$$dZ_{\tilde{t}} = \frac{\sqrt{2\lambda}}{\sigma} d(X_t - (at+b))$$

and

$$dt = \frac{1}{\lambda} d\tilde{t}$$

and by the scaling property of Brownian motion, $B_{\tilde{t}} = \sqrt{\lambda} B_t$.

Thus $dB_{\tilde{t}} = \sqrt{\lambda} dB_t$. Hence,

$$\begin{aligned} \frac{\sigma}{\sqrt{2\lambda}} dZ_{\tilde{t}} = & -\lambda \frac{\sigma}{\sqrt{2\lambda}} Z_{\tilde{t}} \frac{1}{\lambda} d_{\tilde{t}} + \sigma \frac{1}{\sqrt{\lambda}} dB_{\tilde{t}} \\ dZ_{\tilde{t}} = & -Z_{\tilde{t}} d_{\tilde{t}} + \sqrt{2} dB_{\tilde{t}}. \end{aligned}$$

4.2. Boundary Crossing Probabilities and First Passage Time Probabilities of the Standardized Ornstein-Uhlenbeck Process

Consider the standardized OU process

$$dZ_{\tilde{t}} = -Z_{\tilde{t}} d_{\tilde{t}} + \sqrt{2} dB_{\tilde{t}}, \quad Z_0 = z_0. \tag{4}$$

Let $Y_{\tilde{t}} = e^{\tilde{t}} Z_{\tilde{t}}$

Then by Ito's lemma, appendix A.3,

$$\begin{aligned} dY_{\tilde{t}} &= (e^{\tilde{t}} Z_{\tilde{t}} + (-Z_{\tilde{t}}) e^{\tilde{t}} + 0) d\tilde{t} + \sqrt{2} e^{\tilde{t}} dB_{\tilde{t}}, \\ &= \sqrt{2} e^{\tilde{t}} dB_{\tilde{t}} \end{aligned}$$

Thus $Y_{\tilde{t}}$ is Martingale.

$$\int_0^t dY_{\tilde{t}} = \sqrt{2}\int_0^t e^{\tilde{t}} dB_{\tilde{t}}$$

$$Y_t = Y_0 + \sqrt{2}\int_0^t e^{\tilde{t}} dB_{\tilde{t}} \tag{5}$$

It also follows that

$$E[Y_t] = Y_0 \tag{6}$$

$$\begin{aligned} Var[Y_t] &= 2E\left[\left(\int_0^t e^{\tilde{t}} dB_{\tilde{t}}\right)^2\right] \\ &= 2\int_0^t e^{2\tilde{t}} d\tilde{t}, \quad \text{by isometry of Ito integral} \\ &= \left[e^{2\tilde{t}}\right]_0^t \\ &= e^{2t} - 1 \end{aligned} \tag{7}$$

Thus Y_t is a continuous Gaussian process with $E[Y_t] = Y_0$ and $Var[Y_t] = e^{2t} - 1$. For $t \geq 0$, we define $s(t) := e^{2t} - 1$. $s(t)$ is a strictly increasing function of t, and hence admits an inverse $t(s) = \frac{1}{2}ln(1+s)$, $s \geq 0$.

Define $\tilde{W}_s := Y_{t(s)} - Y_0, \quad s \geq 0$.

Then

$$E[\tilde{W}_s] = 0, \quad \tilde{W}_0 = 0$$

and

$$\begin{aligned} Var[\tilde{W}_s] &= Var[Y_{t(s)}] \\ &= e^{2t(s)} - 1 \\ &= e^{2(\frac{1}{2}ln(1+s))} - 1 \\ &= s \end{aligned}$$

So $\{\tilde{W}_s, s \geq 0\}$ is a standard Brownian motion.

Now,

$$\begin{aligned} Z_t &= e^{-t}Y_t \\ &= e^{-t}(Y_0 + \tilde{W}_{s(t)}) \\ &= e^{-t}(Z_0 + \tilde{W}_{s(t)}) \end{aligned} \tag{8}$$

Let $b(t)$ be some continuous function of t such that $b(0) > Z_0$.

Then,

$$\begin{aligned} P(Z_t < b(t), \forall t \in [0,T]) &= P(e^{-t}(z_0 + \tilde{W}_{s(t)}) < b(t), \forall t \in [0,T]) \\ &= P(\tilde{W}_{s(t)} < -z_0 + e^t b(t), \forall t \in [0,T]) \\ &= P(\tilde{W}_s < -z_0 + e^{t(s)} b(t(s)), \forall s \in [0,S]) \end{aligned}$$

(where $S = s(T) = e^{2T} - 1$)

$$= P(\tilde{W}_s < -z_0 + (\sqrt{1+s}) b(t(s)), \forall s \in [0,S]) \tag{9}$$

Similarly, let $a(t)$ be some continuous function of t such that $a(0) < Z_0$.

Thus,

$$\begin{aligned} P(Z_t > a(t), \forall t \in [0,T]) &= P(e^{-t}(z_0 + \tilde{W}_{s(t)}) > a(t), \forall t \in [0,T]) \\ &= P(\tilde{W}_{s(t)} > -z_0 + e^t a(t), \forall t \in [0,T]) \\ &= P(\tilde{W}_s > -z_0 + e^{t(s)} a(t(s)), \forall s \in [0,S]) \end{aligned}$$

(where $S = s(T) = e^{2T} - 1$)

$$= P(\tilde{W}_s > -z_0 + (\sqrt{1+s}) a(t(s)), \forall s \in [0,S]) \tag{10}$$

Under some assumptions, Wang and Potzelberger proved a general form of this result for two-sided boundaries [36]. See appendix A.1.

The above results imply that one can obtain the first passage time probability of the standardized OU process Z_t to either the lower boundary $a(t)$ or the upper boundary $b(t)$

from the corresponding first passage time probability of the standard Brownian motion $\tilde{W}_s$ to the boundaries $-z_0 + (\sqrt{1+s})a(t(s))$ or $-z_0 + (\sqrt{1+s})b(t(s))$, respectively.

4.3. New Class of Non-Linear Boundaries and New Thresholds

In this section, we introduce a new class of boundaries for the OU process. Given the Ornstein-Uhlenbeck process X_t satisfying the SDE

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t, \quad X_0 = x_0,$$

we study the first passage time to the boundaries

$$g(t) = \mu \pm \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\rho\lambda t}), \tag{11}$$

where, $\lambda > 0$, $\sigma > 0$, $\beta > \gamma \geq 0$ and $\rho > -1$. We also assume $\mu + \frac{\sigma}{\sqrt{2\lambda}}(\beta - \gamma) > x_0$ and $\mu - \frac{\sigma}{\sqrt{2\lambda}}(\beta - \gamma) < x_0$. As we will soon see, these boundaries are almost straight lines for small intervals of t. Nonetheless, they are not linear boundaries.

The equivalent of this in the trend-stationary OU process is

$$g(t) = at^{\dagger} + b \pm \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\rho\lambda t}), \tag{12}$$

where $t^{\dagger}$ is an appropriate scaling of t, depending on the context of application (see subsections 5.2.6 and 6.1.4), with $a, b \in \mathbb{R}$ and other parameters satisfying same conditions as above, and here we assume $b + \frac{\sigma}{\sqrt{2\lambda}}(\beta - \gamma) > x_0$ and $b - \frac{\sigma}{\sqrt{2\lambda}}(\beta - \gamma) < x_0$.

Our interest is in the one sided boundary and as such it is sufficient to consider only the upper boundary;

$$g(t) = \mu + \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\rho\lambda t}) \tag{13}$$

or

$$g(t) = at^{\dagger} + b + \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\rho\lambda t}). \tag{14}$$

We define the first passage time of X_t to $g(t)$ by:

$$\tau_{g(t),x_o} := \inf\{t > 0 : X_t \geq g(t) | X_0 = x_0\} \tag{15}$$

We will consider two cases of this threshold (boundary), as listed in the subsequent subsections.

Let us first consider the standardized OU process $Z_{\tilde{t}}$ satisfying the SDE

$$dZ_{\tilde{t}} = -Z_{\tilde{t}}d_{\tilde{t}} + \sqrt{2}dW_{\tilde{t}}, \quad Z_0 = z_0.$$

The above boundary for the OU process is equivalent to the boundary

$$\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}} - \gamma e^{\rho\tilde{t}}, \tag{16}$$

of the standardized OU process.

THEOREM 4.1. *Let $\tau_{\tilde{g}(\tilde{t}),Z_o} := \inf\{\tilde{t} \geq 0 : Z_{\tilde{t}} \geq \tilde{g}(\tilde{t})|Z_0 = z_0\}$ be the first passage time of $Z_{\tilde{t}}$ to the boundary $\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}}$. Then the density of $\tau_{\tilde{g}(\tilde{t})}$ is given by*

$$f_{\tilde{g}(\tilde{t}),z_0}(\tilde{t}) = \frac{(-z_0+\beta)(2e^{2\tilde{t}})}{\sqrt{2\pi}(e^{2\tilde{t}}-1)^{3/2}} exp\left(\frac{-(-z_0+\beta)^2}{2(e^{2\tilde{t}}-1)}\right) \tag{17}$$

PROOF. By equation 9,

$$\begin{aligned}
P(Z_{\tilde{t}} < \tilde{g}(\tilde{t}), \forall t \in [0,T]) &= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})\tilde{g}(\tilde{t}(s)), \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})\beta e^{-t(\tilde{s})}, \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})\beta e^{-\frac{1}{2}ln(1+s)}, \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})\beta (1+s)^{-\frac{1}{2}}, \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + \beta, \forall s \in [0,S])
\end{aligned}$$

Thus the boundary crossing probability of the standardized OU process to the boundary $\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}}$ is equal to the boundary crossing probability of the standard Brownian motion $\tilde{W}_s$ to the constant boundary $g(s) = -z_0 + \beta$.

Let us define the first passage time of the standard Brownian motion to the boundary $g(s) = -z_0 + \beta$ by $\tau_{g(s),z_o} := \inf\{s \geq 0 : \tilde{W}_s \geq -z_0 + \beta|\tilde{W}_0 = 0\}$, and let $f_{g(s),0}$ denote its density function. It is well known that $f_{g(s),0}$ follows the Levy distribution and has density

$$f_{g(s),0}(s) = \frac{(-z_0+\beta)}{\sqrt{2\pi}s^{3/2}} exp\left(\frac{-(-z_0+\beta)^2}{2s}\right)$$

Hence,

$$f_{\tilde{g}(\tilde{t}),z_0}(\tilde{t}) = \frac{(-z_0+\beta)}{\sqrt{2\pi}(e^{2\tilde{t}}-1)^{3/2}} exp\left(\frac{-(-z_0+\beta)^2}{2(e^{2\tilde{t}}-1)}\right)(2e^{2\tilde{t}})$$

□

THEOREM 4.2. *Let* $\tau_{\tilde{g}(\tilde{t}),Z_o} := \inf\{\tilde{t} \geq 0 : Z_{\tilde{t}} \geq \tilde{g}(\tilde{t}) | Z_0 = z_0\}$ *be the first passage time of* $Z_{\tilde{t}}$ *to the boundary* $\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}} - \gamma e^{\tilde{t}}$. *Then the density of* $\tau_{\tilde{g}(\tilde{t})}$ *is given by*

$$f_{\tilde{g}(\tilde{t}),z_0}(\tilde{t}) = \frac{2|-z_0+\beta-\gamma|(e^{2\tilde{t}})}{\sqrt{2\pi(e^{2\tilde{t}}-1)^3}} exp\left(\frac{-(-z_0+\beta-\gamma e^{2\tilde{t}})^2}{2(e^{2\tilde{t}}-1)}\right) \tag{18}$$

PROOF. By equation 9,

$$\begin{aligned}
P(Z_{\tilde{t}} < \tilde{g}(\tilde{t}), \forall t \in [0,T]) &= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})\tilde{g}(\tilde{t}(s)), \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})(\beta e^{-t(\tilde{s})} - \gamma e^{t(\tilde{s})}), \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})(\beta e^{-\frac{1}{2}ln(1+s)} - \gamma e^{\frac{1}{2}ln(1+s)}), \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + (\sqrt{1+s})(\beta(1+s)^{-\frac{1}{2}} - \gamma(1+s)^{\frac{1}{2}}), \forall s \in [0,S]) \\
&= P(\tilde{W}_s < -z_0 + \beta - \gamma - \gamma s, \forall s \in [0,S])
\end{aligned}$$

Here the boundary crossing probability of the standardized OU process to the boundary $\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}} - \gamma e^{\tilde{t}}$ is equal to the boundary crossing probability of the standard Brownian motion $\tilde{W}_s$ to the linear boundary $g(s) = -z_0 + \beta - \gamma - \gamma s$.

Let us define the first passage time of the standard Brownian motion to the boundary $g(s) = -z_0 + \beta - \gamma - \gamma s$ by $\tau_{g(s),Z_o} := \inf\{s \geq 0 : \tilde{W}_s \geq -z_0 + \beta - \gamma - \gamma s | \tilde{W}_0 = 0\}$, and let $f_{g(s),0}$ denote its density function.

Following Karatzas and Shreve [24], we have that $f_{g(s),0}$ follows an inverse gaussian distribution and has density

$$f_{g(s),0}(s) = \frac{|(-z_0+\beta-\gamma)|}{\sqrt{2\pi s^3}} exp\left(\frac{-(-z_0+\beta-\gamma-\gamma s)^2}{2s}\right)$$

Hence,

$$f_{\tilde{g}(\tilde{t}),z_0}(\tilde{t}) = \frac{|(-z_0+\beta-\gamma)|}{\sqrt{2\pi(e^{2\tilde{t}}-1)^3}} exp\left(\frac{-(-z_0+\beta-\gamma-\gamma(e^{2\tilde{t}}-1))^2}{2(e^{2\tilde{t}}-1)}\right)(2e^{2\tilde{t}})$$

$$= \frac{2|(-z_0+\beta-\gamma)|(e^{2\tilde{t}})}{\sqrt{2\pi(e^{2\tilde{t}}-1)^3}} exp\left(\frac{-(-z_0+\beta-\gamma e^{2\tilde{t}})^2}{2(e^{2\tilde{t}}-1)}\right)$$

□

4.4. The Ornstein-Uhlenbeck Process First Passage Time Densities

4.4.1. Case 1: $\gamma = 0$

Let $\gamma = 0$, then the boundary 13 becomes

$$g(t) = \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t}) + \mu. \tag{19}$$

COROLLARY 4.3. *The density of the first passage time $\tau_{g(t),x_o}$ of the OU process X_t to the above boundary, $g(t)$, is given by*

$$f_{g(t),x_0} = \frac{(\frac{\sqrt{2\lambda}}{\sigma}(-x_0+\mu)+\beta)}{\sqrt{2\pi}(e^{2\lambda t}-1)^{3/2}} exp\left(\frac{-(\frac{\sqrt{2\lambda}}{\sigma}(-x_0+\mu)+\beta)^2}{2(e^{2\lambda t}-1)}\right)(2\lambda e^{2\lambda t}). \tag{20}$$

To see this, notice that in the transformation of the OU process X_t to the standard OU process $Z_{\tilde{t}}$, we set $\tilde{t} = \lambda t$ and $Z_{\tilde{t}} = (X_t - \mu)/\sqrt{\frac{\sigma^2}{2\lambda}}$.
Thus $z_0 = \frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu)$. So the result follows from theorem 4.1 after changing the time variable from $\tilde{t}$ to t, and replacing z_0 with $\frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu)$.

4.4.2. Case 2: $\gamma \geq 0$ and $\rho = 1$

Now, let $\gamma \geq 0$ and $\rho = 1$. Then the boundary 13 becomes

$$g(t) = \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\lambda t}) + \mu.$$

COROLLARY 4.4. *The density of the first passage time $\tau_{g(t),x_o}$ of the OU process X_t to the above boundary, $g(t)$, is given by*

$$f_{g(t),x_0} = \frac{|\frac{\sqrt{2\lambda}}{\sigma}(-x_0+\mu)+\beta-\gamma|}{\sqrt{2\pi(e^{2\lambda t}-1)^3}} exp\left(\frac{-(\frac{\sqrt{2\lambda}}{\sigma}(-x_0+\mu)+\beta-\gamma e^{2\lambda t})^2}{2(e^{2\lambda t}-1)}\right)(2\lambda e^{2\lambda t}). \tag{21}$$

This result also follows from theorem 4.2, by a similar argument as in case 1.

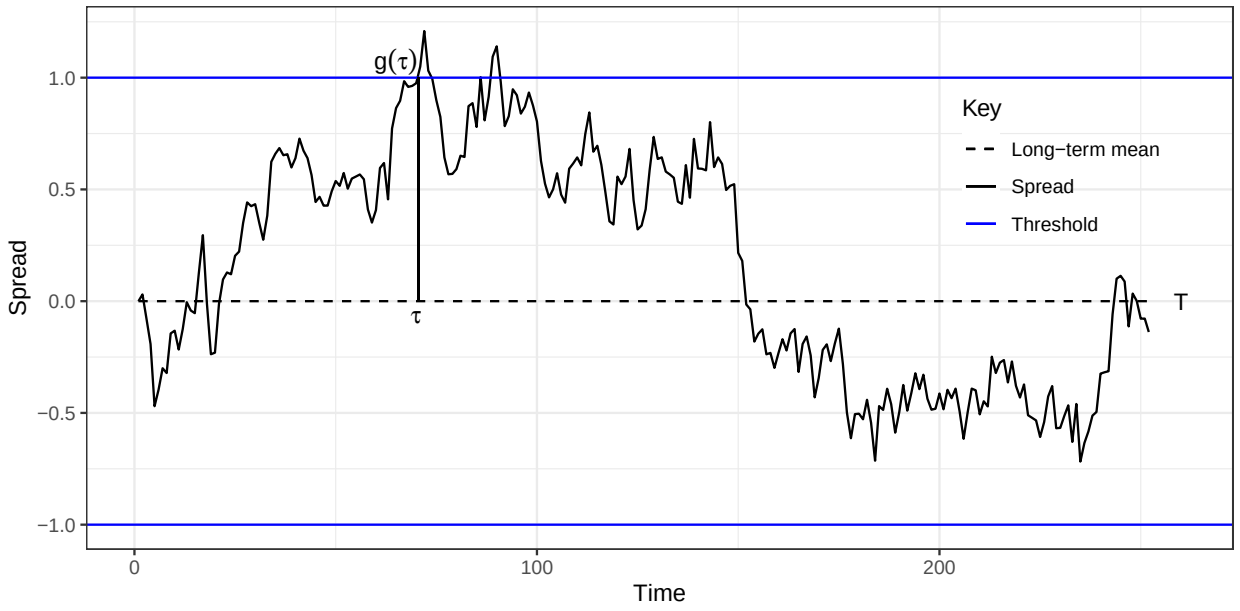

FIGURE 4.1. Spread and threshold.

Given the spread between the logarithmic returns of the pair of stocks, our objective is to find thresholds that maximize the expected return of a complete trade cycle. The trade cycle involves two stages. The first stage is the movement from the long-term mean or long-term trend to crossing either the upper or lower threshold, while the second stage involves the movement back to the long-term mean or long-term trend. Since the thresholds we are considering are symmetric 11, we will focus on a one-sided threshold for our optimization problem. The return on the trade cycle corresponds to the height of the spread at the time of crossing the threshold. We demonstrated this in figure 4.1 for the constant threshold for simplicity. Since we are considering a time-dependent threshold this height will vary over time.

In section 4.1, we showed that the OU process 1 and the trend-stationary OU process 3 can be converted to the standardized (dimensionless)OU process. Thus we will undertake the optimization in the dimensionless system, and then the optimal thresholds in the original process can be obtained by translation and scaling, as shown in section 4.3, equations 14, 15

and 16.

In our study, it came to light that although theoretically we can always find values for our parameters β and γ that maximize the expected return, practically, this can only be done for a reasonable time horizon, which usually should not be more than $1\frac{1}{4}$ years. As a result, our objective function will entail two pieces, the first piece focuses on time horizon $T \leq 1.25$ where we optimize the expected return, while the second piece focuses on $T > 1.25$ where we optimize a scaled form of the probability that the first passage time is greater than 1.25 but less than the time horizon. This is to ensure that β does not blow out, nor γ become too small such that we do not obtain any trades. In fact this second piece of the objective function is a nice behaved function that maximizes our return.

4.5.1. Case 1: $\gamma = 0$

The dimensionless system threshold for this case is

$$\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}}.$$

The objective function is

$$h(\beta) = \begin{cases} E\left[\beta e^{-\tau} I_{(0<\tau\leq T)}\right], & 0 < T \leq 1.25 \\ e^{-\beta} P(1.25 < \tau < T), & 1.25 < T < \infty \end{cases}$$

THEOREM 4.5. *For any $0 < T < \infty$, there exists some $\tilde{\beta} \in (0, \infty)$ which maximizes $h(\beta)$ and it is unique.*

LEMMA 4.6. *There exists some $\tilde{\beta} \in (0, \infty)$ which maximizes $E\left[\beta e^{-\tau} I_{(0<\tau\leq T)}\right]$ and it is unique.*

PROOF. Let $\Pi_T(\beta) = E\left[\beta e^{-\tau} I_{(0<\tau\leq T)}\right]$, where $\beta > 0$.

By theorem 4.1, the pdf of τ is given by

$$f_\tau(\tilde{t}) = \sqrt{\frac{2}{\pi}} \frac{\beta}{\sqrt{(e^{2\tilde{t}} - 1)^3}} exp\left(\frac{-\beta^2}{2(e^{2\tilde{t}} - 1)}\right), \quad \tilde{t} > 0$$

By change of variable, with $v = \beta\sqrt{\frac{e^{2\tilde{t}}}{e^{2\tilde{t}}-1}}$, we have,

$$\Pi_T(\beta) = 2\beta e^{\frac{\beta^2}{2}} \int_{C\beta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv, \quad \text{where} \quad C = \sqrt{\frac{e^{2T}}{e^{2T}-1}}$$

Thus

$$\Pi_T(\beta) = 2\beta e^{\frac{\beta^2}{2}} [1 - \Phi(C\beta)], \tag{22}$$

where Φ denotes the standard normal cumulative distribution function(CDF).

By the Mills ratio inequality, see appendix A.2, [32] and [39], it follows that

$$\frac{2}{\sqrt{(C\beta)^2+4}+C\beta}\phi(C\beta) \leq 1 - \Phi(C\beta) \leq \frac{4}{\sqrt{(C\beta)^2+8}+3C\beta}\phi(C\beta)$$

where ϕ is the standard normal density function.

Hence,

$$\frac{4\beta e^{\frac{\beta^2}{2}}\phi(C\beta)}{\sqrt{(C\beta)^2+4}+C\beta} \leq \Pi_T(\beta) \leq \frac{8\beta e^{\frac{\beta^2}{2}}\phi(C\beta)}{\sqrt{(C\beta)^2+8}+3C\beta}$$

The right hand side of the above inequality is equivalent to $\frac{8\beta}{\sqrt{2\pi}} e^{-\frac{(C^2-1)\beta^2}{2}} \frac{1}{\sqrt{(C\beta)^2+8}+3C\beta}$.

Since $C^2 - 1 = \frac{1}{e^{2T}-1} > 0$, it follows that,

$$\lim_{\beta\to\infty} \frac{8\beta}{\sqrt{2\pi}} e^{-\frac{(C^2-1)\beta^2}{2}} \frac{1}{\sqrt{(C\beta)^2+8}+3C\beta} = 0$$

Also,

$$\lim_{\beta\to 0} \frac{8\beta}{\sqrt{2\pi}} e^{-\frac{(C^2-1)\beta^2}{2}} \frac{1}{\sqrt{(C\beta)^2+8}+3C\beta} = 0$$

Which proves that $\tilde{\beta} = \max_{\beta>0} \Pi_T(\beta)$ exists and is achieved in $(0, \infty)$.

We now show that $\tilde{\beta}$ is unique. We do this by contradiction.

From equation 21, we have that,

$$\frac{d\Pi_T(\beta)}{d\beta} = 2(1+\beta^2)e^{\frac{\beta^2}{2}}[1 - \Phi(C\beta)] - 2C\beta\phi(C\beta)e^{\frac{\beta^2}{2}}.$$

Suppose there are two maximizers, namely β_1 and β_2. Then they both satisfy the equation:

$$\frac{1 - \Phi(C\beta)}{\phi(C\beta)} = \frac{C\beta}{1+\beta^2}.$$

Thus,

$$log[1-\Phi(C\beta_i)] = log\phi(C\beta_i) + log(\beta_i) + log(C) - log(1+\beta_i^2), \quad \text{for i = 1, 2.} \tag{23}$$

Since both β_1 and β_2 are maximizers of $\Pi_T(\beta)$, then $\Pi_T(\beta_1) = \Pi_T(\beta_2)$ with $\beta_1 \neq \beta_2$.

Then from equation 22, we have

$$log[1-\Phi(C\beta_1)] - log[1-\Phi(C\beta_2)] = log\left(\frac{\beta_2}{\beta_1}\right) + \frac{1}{2}(\beta_2^2 - \beta_1^2). \tag{24}$$

Equations 23 and 24 together implies:

$$log\left(\frac{\phi(C\beta_1)}{\phi(C\beta_2)}\right) + log\left(\frac{1+\beta_2^2}{1+\beta_1^2}\right) + \frac{1}{2}(\beta_1^2 - \beta_2^2) + 2log\left(\frac{\beta_1}{\beta_2}\right) = 0$$

This is equivalent to

$$\frac{\beta_1^2}{1+\beta_1^2} e^{-\frac{1}{2}(C^2-1)\beta_1^2} = \frac{\beta_2^2}{1+\beta_2^2} e^{-\frac{1}{2}(C^2-1)\beta_2^2}$$

Consider the function:

$$l(x) = \frac{xe^{-\frac{1}{2}(C^2-1)x}}{1+x} \text{for} \quad x > 0$$

The derivative of $l(x)$ is given by:

$$l'(x) = \frac{[1-\frac{1}{2}(C^2-1)x(x+1)]}{(x+1)^2} e^{-\frac{1}{2}(C^2-1)x}$$

We note that the equation $1-\frac{1}{2}(C^2-1)x(x+1) = 0$ has two roots: One is negative and the other is positive, which is given by

$$\tilde{x} = \frac{-1+\sqrt{1+\frac{8}{C^2-1}}}{2}.$$

When $x > \tilde{x}, l'(x) < 0$, which implies $l(x)$ is strictly decreasing on $(\tilde{x}, \infty)$. When $0 < x < \tilde{x}, l'(x) > 0$, which implies $l(x)$ is strictly increasing on $(0, \tilde{x})$. Since $l(\beta_1^2) = l(\beta_2^2)$ and $\beta_1 \neq \beta_2$, by Rolle's theorem, there exists at least one r between $\beta_1^2 \text{and} \beta_2^2$ such that $l'(r) = 0$. W.L.O.G., we assume $\beta_1 < \beta_2$. Then $r \in (\beta_1^2, \beta_2^2)$. Using the fact that $l'(x) = 0$ has only one positive root $\tilde{x}$, we conclude that $r = \tilde{x}$. Since β_2 is a maximizer of $\Pi_T(\beta)$, we obtain

$$\Pi_T(\beta_2) = \frac{2C}{\sqrt{2\pi}} l(\beta_2^2).$$

But $\beta_2^2 > r$ and $l(x)$ is a strictly decreasing function of x on $x > r$. Thus we conclude that

$$\Pi_T(\beta_2) < \frac{2C}{\sqrt{2\pi}} l(r) = \Pi_T(r).$$

Which contradicts the assumption that β is a maximizer of $\Pi_T(\beta)$. Thus it follows that the maximizer of $\Pi_T(\beta)$ is unique, which completes the proof. □

LEMMA 4.7. *Given $0 < T_s < T < \infty$, there exists some $\tilde{\beta} \in (0, \infty)$ which maximizes $e^{-\beta} P(T_s < \tau < T)$, and it is unique for sufficiently large T.*

PROOF. To show the existence, let $\Pi_{T_s,T}(\beta) = e^{\beta} P(T_s < \tau < T)$, where $\beta > 0$.

Thus

$$\Pi_{T_s,T}(\beta) = e^{\beta} \int_{T_s}^{T} \sqrt{\frac{2}{\pi}} \frac{\beta e^{2t}}{\sqrt{(e^{2t}-1)^3}} e^{-\frac{\beta^2}{2(e^{2t}-1)}} dt.$$

By change of variable with $u = \frac{1}{e^{2t}-1}$, we obtain

$$\Pi_{T_s,T}(\beta) = \frac{\beta e^{-\beta}}{\sqrt{2\pi}} \int_{(e^{2T}-1)^{-1}}^{(e^{2T_s}-1)^{-1}} u^{-\frac{1}{2}} e^{-\frac{\beta^2}{2}u} du.$$

By another change of variable with $y = \sqrt{u}$, we get

$$\begin{aligned} \Pi_{T_s,T}(\beta) &= \frac{2\beta e^{-\beta}}{\sqrt{2\pi}} \int_{C_1}^{C_2} e^{-\frac{(\beta y)^2}{2}} dy, \text{ where } \quad C_1 = \sqrt{\frac{1}{e^{2T}-1}}, C_2 = \sqrt{\frac{1}{e^{2T_s}-1}} \\ &= 2e^{-\beta} \left[\Phi(\beta C_2) - \Phi(\beta C_1)\right], \end{aligned}$$

where Φ is the standard normal CDF.

It follows from the above equation that

$$lim_{\beta \to 0} \Pi_{T_s,T}(\beta) = 0 \text{ and } lim_{\beta \to \infty} \Pi_{T_s,T}(\beta) = 0$$

To show uniqueness, we consider the function,

$$l(\beta) = log(\Pi_{T_s,T}(\beta))$$

Thus,

(25)

$$l'(\beta) = -1 + \frac{\phi(\beta C_2) C_2 - \phi(\beta C_1) C_1}{\Phi(\beta C_2) - \Phi(\beta C_1)}, \text{where } \phi \text{ is the standard normal PDF}$$

and

(26)
$$l''(\beta) = \frac{[-\beta C_2\phi(\beta C_2)C_2^2 + \beta C_1\phi(\beta C_1)C_1^2][\Phi(\beta C_2) - \Phi(\beta C_1)] - [\phi(\beta C_2)C_2 - \phi(\beta C_1)C_1]^2}{[\Phi(\beta C_2) - \Phi(\beta C_1)]^2}.$$

Since $C_2 > C_1$, $\Phi(\beta C_2) - \Phi(\beta C_1) > 0$

For sufficiently large T, $C_1 \to 0$, which implies $\phi(\beta C_1)C_1^3 - \phi(\beta C_2)C_2^3 < 0$

Thus $l''(\beta) < 0$ Hence for sufficiently large T, the maximizer of $l(\beta)$ is unique. □

To cover uniqueness for small values of T, a more delicate mathematical analysis is needed. This has been a challenge. However, let us consider the function $\xi_{T_s,T}(\beta)$, defined by:

$$\xi_{T_s,T}(\beta) = 2e^{-\beta}\left[\beta\Delta - \frac{1}{2}\beta^3 C_1\Delta^2\right]\phi(\beta C_1), \text{where} \Delta = C_2 - C_1$$

This function is an accurate approximation of $\Pi_{T_s,T}(\beta)$. More importantly, their maximizers are approximately the same. We show in figure 4.2 a plot of these two functions for the case of $T_s = 1.25$ and $T = 5$. In fact for smaller values of T, it is impossible to distinguish between the two curves as they overlay each other perfectly.

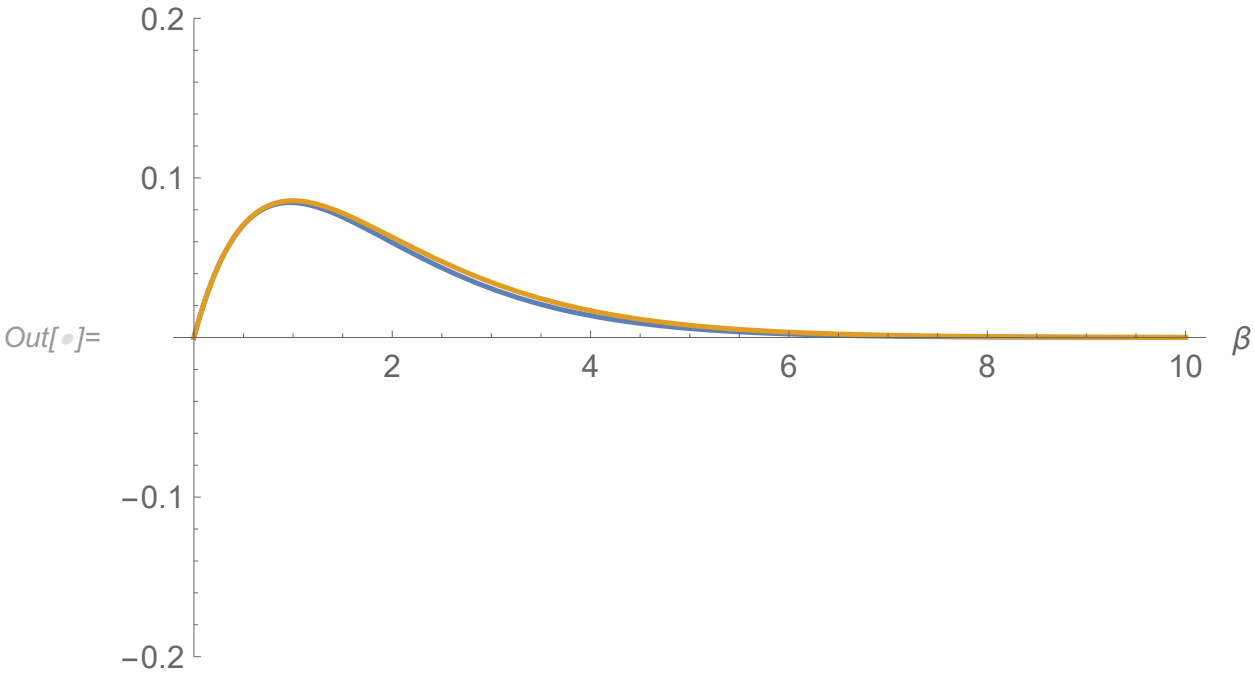

FIGURE 4.2. Graph of $\xi_{T_s,T}(\beta)$(Yellow) and $\Pi_{T_s,T}(\beta)$(Blue), with $T_s = 1.25$ and $T = 5$

Note that at this point, we only need to show uniqueness of the maximizer of $\Pi_{T_s,T}(\beta)$ for small values of T (i.e. when T is close to T_s), in order to obtain a complete proof of theorem 4.5. Nonetheless, we will prove a general statement for $\xi_{T_s,T}(\beta)$.

LEMMA 4.8. *For all $0 < T_s < T < \infty$, there exists some $\tilde{\beta} \in (0, \infty)$ which maximizes $\xi_{T_s,T}(\beta)$, and it is unique.*

PROOF. Let $\xi_{T_s,T}(\beta) = 2e^{-\beta}\left[\beta\Delta - \frac{1}{2}\beta^3 C_1\Delta^2\right]\phi(\beta C_1)$, where $\Delta = C_2 - C_1$.
Now,

$$\begin{aligned} e^{-\beta}\phi(C_1\beta) &= e^{-\beta}\frac{1}{\sqrt{2\pi}}e^{-\frac{(C_1\beta)^2}{2}} \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{(C_1\beta)^2+2\beta}{2}} \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{C_1^2\left(\beta^2+\frac{2\beta}{C_1^2}+\frac{1}{C_1^4}-\frac{1}{C_1^4}\right)}{2}} \\ &= e^{\frac{1}{2C_1^2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(C_1\beta+\frac{1}{C_1}\right)^2}{2}} \end{aligned}$$

Hence,

$$\xi_{T_s,T}(\beta) = 2\Delta\left[\beta - \frac{1}{2}\beta^3 C_1\Delta\right] e^{\frac{1}{2C_1^2}}\phi\left(C_1\beta + \frac{1}{C_1}\right)$$

Let $\chi_{T_s,T}(\beta) = log\xi_{T_s,T}(\beta)$.

Thus,

$$\chi_{T_s,T}(\beta) = log(2\Delta) + log(\beta - \frac{1}{2}\beta^3 C_1\Delta) + \frac{1}{2C_1^2} + log\left[\phi\left(C_1\beta + \frac{1}{C_1}\right)\right]$$

So,

$$\chi_{T_s,T}{}'(\beta) = \frac{1 - \frac{3}{2}C_1\beta^2\Delta}{\beta - \frac{1}{2}C_1\beta^3\Delta} - (C_1^2\beta + 1)$$

and

$$\chi_{T_s,T}{}''(\beta) = \frac{-3C_1\beta\Delta(\beta - \frac{1}{2}C_1\beta^3\Delta) - (1 - \frac{3}{2}C_1\beta^2\Delta)^2}{(\beta - \frac{1}{2}C_1\beta^3\Delta)^2} - C_1^2$$

Now,

$$\begin{aligned}
&-3C_1\beta\Delta(\beta - \frac{1}{2}C_1\beta^3\Delta) - (1 - \frac{3}{2}C_1\beta^2\Delta)^2 \\
&\quad = -3C_1\beta(C_2 - C_1)(\beta - \frac{1}{2}C_1\beta^3(C_2 - C_1)) - (1 - \frac{3}{2}C_1\beta^2(C_2 - C_1))^2 \\
&\quad = -3\beta^2 C_1 C_2 + \frac{3}{2}\beta^4 C_1^2 C_2^2 - \frac{3}{2}\beta^4 C_1^3 C_2 + 3\beta^2 C_1^2 - \frac{3}{2}\beta^4 C_1^3 C_2 + \frac{3}{2}\beta^4 C_1^4 \\
&\qquad - \left(1 + \frac{9}{2}\beta^4 C_1^2 C_2^2 + \frac{9}{2}\beta^4 C_1^4 - 3\beta^2 C_1 C_2 + 3\beta^2 C_1^2 - 2\left(\frac{3}{2}\beta^2 C_1 C_2\right)\left(\frac{3}{2}\beta^2 C_1^2\right)\right) \\
&\quad = -1 - 3\beta^4 C_1^2 C_2^2 - \frac{3}{4}\beta^4 C_1^4 \\
&\quad < 0,
\end{aligned}$$

which implies that $\chi_{T_s,T}{}''(\beta) < 0$.

Also, we may rewrite $\xi_{T_s,T}(\beta)$ as $\xi_{T_s,T}(\beta) = 2e^{-\beta}\beta\left[\Delta - \frac{1}{2}\beta^2 C_1\Delta^2\right]\phi(\beta C_1)$.

Then, it is then clear that $lim_{\beta\to 0}\xi_{T_s,T}(\beta) = 0$, and $lim_{\beta\to\infty}\xi_{T_s,T}(\beta) = 0$.

Thus we have proven both the existence and uniqueness of a maximizer. □

Theorem 4.5 thus follows from lemma 4.6, lemma 4.7 and lemma 4.8.

By setting the first derivative of the objective function equal to zero, we find that for $0 < T \leq 1.25$, the unique $\tilde{\beta}$ can be found by solving the equation

$$(1 + \beta^2)[1 - \Phi(C\beta)] = C\beta\phi(C\beta)$$

and for $1.25 < T < \infty$, it can be found by solving the equation

$$C_2\phi(C_2\beta) - C_1\phi(C_1\beta) = \Phi(C_2\beta) - \Phi(C_1\beta),$$

where Φ, ϕ, C, C_1 and C_2 are as defined above.

4.5.2. Case 2: $\gamma \geq 0$ and $\rho = 1$

The dimensionless system threshold for this case is

$$\tilde{g}(\tilde{t}) = \beta e^{-\tilde{t}} - \gamma e^{\rho\tilde{t}}$$

The objective function is

$$h(\beta, \gamma) = \begin{cases} E\left[(\beta e^{-\tau} - \gamma e^{\tau}) I_{(0<\tau\leq T)}\right], & 0 < T \leq 1.25 \\ \gamma e^{-\beta} P(1.25 < \tau < T), & 1.25 < T < \infty \end{cases}$$

We claim that for any $0 < T < \infty$, there exists some $0 < \tilde{\gamma} < \tilde{\beta} < \infty$ which maximizes $h(\beta, \gamma)$ and $(\tilde{\beta}, \tilde{\gamma})$ is unique.

The mathematical proof of the above claim is part of an ongoing project and is not yet complete. However this claim is backed by several numerical examples that we performed with the numerical maximization function in Mathematica [38]. An example of the first piece with $T = 0.5$ is shown in figure 4.3, where the maximum is $h(\tilde{\beta}, \tilde{\gamma}) = 0.3632$ and occurs at $(\tilde{\beta}, \tilde{\gamma}) = (2.2621, 0.6944)$. Also, an example of the second piece with $T = 5$ is shown in figure 4.4, where the maximum is $h(\tilde{\beta}, \tilde{\gamma}) = 0.006564$ and occurs at $(\tilde{\beta}, \tilde{\gamma}) = (1.3351, 0.1856)$

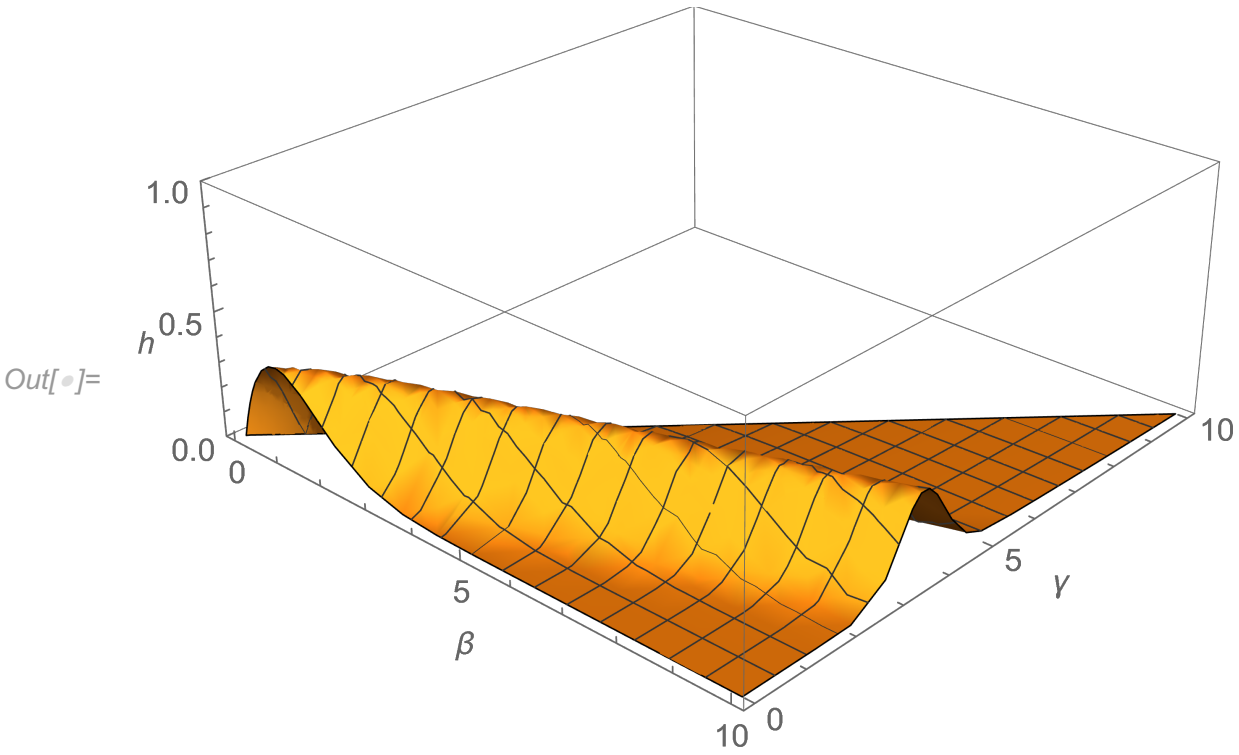

FIGURE 4.3. Graph of $h(\beta, \gamma)$ for $T = 0.5$, with maximum value $h(\tilde{\beta}, \tilde{\gamma}) = 0.3632$ at $\tilde{\beta} = 2.2621$ and $\tilde{\gamma} = 0.6944$

4.6. Certain Moments of First Passage Time of Brownian Motions to Some Class of Boundaries

THEOREM 4.9. *Let* $\zeta(s) = -a + b(s+c)^d$, *where* $a > 0$, $b > 0$, $d > 0$, $bc^d < a$ *and* $(W_s)_{s \geq 0}$ *be a standard Brownian motion. We define the first passage time (from above) of* W_s *to* $\zeta(s)$ *by* $\tau := inf\{s \geq 0 : W_s \leq \zeta(s)\}$. *Let* $F(s) = P(\tau \leq s)$ *be its distribution function, and* $f(s)$ *be the probability density function. Then* $E\left[(\tau + c)^d\right] = \frac{a}{b}$, *if* $(s+c)^d$ *is defined on* $0 \leq s \leq \infty$.

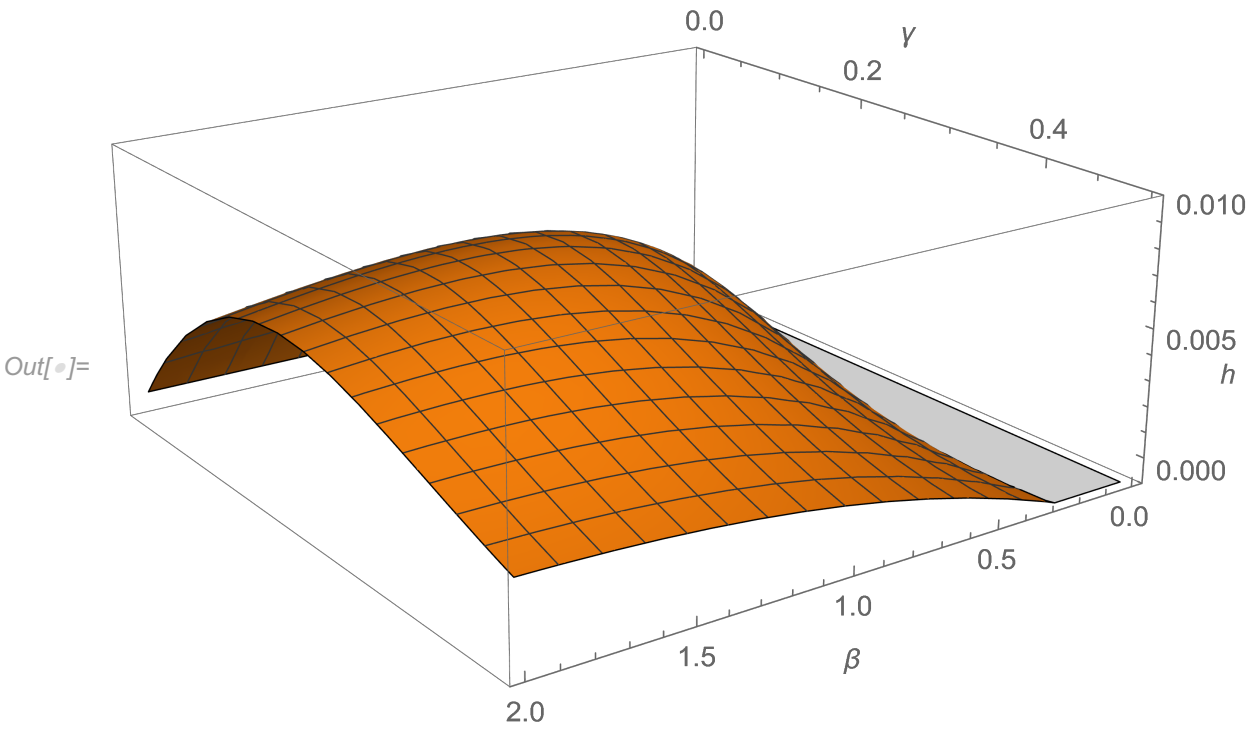

FIGURE 4.4. Graph of $h(\beta, \gamma)$ for $T = 5$, with maximum value $h(\tilde{\beta}, \tilde{\gamma}) = 0.006564$ at $\tilde{\beta} = 1.3351$ and $\tilde{\gamma} = 0.1856$

PROOF. Let $\tau_\alpha = inf\{s \geq 0 : W_s \leq \zeta(s) + \alpha s\}, \quad \alpha \in [0, \infty)$.

Let F_α be the distribution function of τ_α

Then,

$$\begin{aligned} \zeta(s) + \alpha s = & -a + b(s+c)^d + \alpha s \\ \geq & -a + b(s+c)^d \\ \geq & -a + min\{bc^d, 0\} \end{aligned}$$

Let $l = -a + min\{bc^d, 0\}$

Since $\tau_l := inf\{s \geq 0 : W_s \leq l\}$ is almost surely finite, then so is τ_α.

Thus $F_\alpha(\infty) = 1$.

Now, let $B_s = W_s - \alpha s$

Thus $\tau_\alpha = inf\{s \geq 0 : B_s \leq \zeta(s)\}$

Define $Z := e^{-\frac{\alpha^2 s}{2} + \alpha W_s}$

Define a probability measure $Q(A) := \int_A z(w)dP(w)$.

By Girsanov's theorem, A.4, $B_s = W_s - \alpha s$ is a standard Brownian motion under Q.

Thus τ_α has the distribution F under Q.

So, $E_Q\left[1_{\tau_\alpha \le s}\right] = E_P\left[1_{\tau_\alpha \le t} Z\right]$.

$$\begin{aligned}
F(s) &= Q(\tau_\alpha \le s) \\
&= E_P\left[1_{\tau_\alpha \le s} Z\right] \\
&= E_P\left[1_{\tau_\alpha \le s} e^{-\frac{\alpha^2 s}{2} + \alpha W_s}\right] \\
&= E_P\left[E_P\left[1_{\tau_\alpha \le s} e^{-\frac{\alpha^2 s}{2} + \alpha W_s} | \tau_\alpha = \nu\right]\right] \\
&= \int_0^s E_P\left[e^{-\frac{\alpha^2 s}{2} + \alpha W_s} | \tau_\alpha = \nu\right] F_\alpha(d\nu) \\
&= \int_0^s E_P\left[e^{-\frac{\alpha^2 s}{2} + \alpha(W_s - W_{\tau_\alpha} + W_{\tau_\alpha})} | \tau_\alpha = \nu\right] F_\alpha(d\nu) \\
&= \int_0^s E_P\left[e^{-\frac{\alpha^2 s}{2} + \alpha(W_s - W_\nu)} e^{\alpha W_{\tau_\alpha}} | \tau_\alpha = \nu\right] F_\alpha(d\nu).
\end{aligned}$$

Since $W_{\tau_\alpha} = \zeta(\tau_\alpha) + \alpha\tau_\alpha$ almost surely,

$$\begin{aligned}
F(s) &= \int_0^s e^{\alpha(\zeta(\nu) + \alpha\nu)} E_P\left[e^{-\frac{\alpha^2 s}{2} + \alpha(W_s - W_\nu)} | \tau_\alpha = \nu\right] F_\alpha(d\nu) \\
&= \int_0^s e^{\alpha\zeta(\nu) + \alpha^2\nu} E_P\left[e^{-\frac{\alpha^2 s}{2} + \alpha(W_s - W_\nu)}\right] F_\alpha(d\nu) \\
&= \int_0^s e^{\alpha\zeta(\nu) + \alpha^2\nu} e^{-\frac{\alpha^2 s}{2}} E_P\left[e^{\alpha(W_s - W_\nu)}\right] F_\alpha(d\nu) \\
&= \int_0^s e^{\alpha\zeta(\nu) + \alpha^2\nu} e^{-\frac{\alpha^2 s}{2}} e^{\frac{\alpha^2}{2}(s-\nu)} F_\alpha(d\nu) \\
&= \int_0^s e^{\alpha\zeta(\nu) + \alpha^2\nu - \frac{\alpha^2\nu}{2}} F_\alpha(d\nu) \\
&= \int_0^s e^{\alpha\zeta(\nu) + \frac{\alpha^2\nu}{2}} F_\alpha(d\nu)
\end{aligned}$$

This implies,

$$F(ds) = e^{\alpha\zeta(s) + \frac{\alpha^2 s}{2}} F_\alpha(ds)$$

$$e^{-\alpha\zeta(s)-\frac{\alpha^2 s}{2}}F(ds) = F_\alpha(ds)$$

$$\int_0^\infty e^{-\alpha\zeta(s)-\frac{\alpha^2 s}{2}}F(ds) = \int_0^\infty F_\alpha(ds)$$

$$= F_\alpha(\infty)$$

$$= 1$$

We differentiate both sides with respect to α

$$\frac{d}{d\alpha}\int_0^\infty e^{-\alpha\zeta(s)-\frac{\alpha^2 s}{2}}F(ds) = 0$$

$$\int_0^\infty \frac{d}{d\alpha}e^{-\alpha\zeta(s)-\frac{\alpha^2 s}{2}}F(ds) = 0$$

$$\int_0^\infty (-\zeta(s)-\alpha s)e^{-\alpha\zeta(s)-\frac{\alpha^2 s}{2}}F(ds) = 0$$

Let $\alpha \to 0$.

Then,

$$\int_0^\infty -\zeta(s)F(ds) = 0$$

$$\int_0^\infty -(-a+b(s+c)^d)F(ds) = 0$$

$$b\int_0^\infty (s+c)^d F(ds) = \int_0^\infty aF(ds)$$

$$\int_0^\infty (s+c)^d F(ds) = \frac{a}{b}$$

Thus $E\left[(\tau+c)^d\right] = \frac{a}{b}$ □

THEOREM 4.10. *Let $\zeta(s) = a - b(s+c)^d$, where $a > 0$, $b > 0$, $d > 0$, $bc^d < a$ and $(W_s)_{s\geq 0}$ be a standard Brownian motion. We define the first passage time (from below) of W_s to $\zeta(s)$ by $\tau := inf\{s \geq 0 : W_s \geq \zeta(s)\}$. Let $F(s) = P(\tau \leq s)$ be its distribution function, and $f(s)$ be the probability density function. Then $E\left[(\tau+c)^d\right] = \frac{a}{b}$, if $(s+c)^d$ is defined on $0 \leq s \leq \infty$.*

PROOF. Let $\tau_\alpha = inf\{s \geq 0 : W_s \geq \zeta(s) - \alpha s\}, \quad \alpha \in [0, \infty)$.

Let F_α be the distribution function of τ_α

Then,

$$\begin{aligned}\zeta(s) - \alpha s &= a - b(s+c)^d - \alpha s \\ &\leq a - b(s+c)^d \\ &\leq a - min\{bc^d, 0\}\end{aligned}$$

Let $l = a - min\{bc^d, 0\}$

Since $\tau_l := inf\{s \geq 0 : W_s \geq l\}$ is almost surely finite, then so is τ_α.

Thus $F_\alpha(\infty) = 1$.

Now, let $B_s = W_s + \alpha s$

Thus $\tau_\alpha = inf\{s \geq 0 : B_s \geq \zeta(s)\}$

Define $Z := e^{-\frac{\alpha^2 s}{2} - \alpha W_s}$

Define a probability measure $Q(A) := \int_A z(w) dP(w)$.

By Girsanov's theorem, A.4, $B_s = W_s + \alpha s$ is a standard Brownian motion under Q.

Thus τ_α has the distribution F under Q.

So, $E_Q[1_{\tau_\alpha \leq s}] = E_P[1_{\tau_\alpha \leq t} Z]$.

Thus,

$$\begin{aligned}F(s) &= Q(\tau_\alpha \leq s) \\ &= E_P[1_{\tau_\alpha \leq s} Z] \\ &= E_P\left[1_{\tau_\alpha \leq s} e^{-\frac{\alpha^2 s}{2} - \alpha W_s}\right] \\ &= E_P\left[E_P\left[1_{\tau_\alpha \leq s} e^{-\frac{\alpha^2 s}{2} - \alpha W_s} | \tau_\alpha = \nu\right]\right] \\ &= \int_0^s E_P\left[e^{-\frac{\alpha^2 s}{2} - \alpha W_s} | \tau_\alpha = \nu\right] F_\alpha(d\nu) \\ &= \int_0^s E_P\left[e^{-\frac{\alpha^2 s}{2} - \alpha(W_s - W_{\tau_\alpha} + W_{\tau_\alpha})} | \tau_\alpha = \nu\right] F_\alpha(d\nu)\end{aligned}$$

$$= \int_0^s E_P \left[e^{-\frac{\alpha^2 s}{2} - \alpha(W_s - W_\nu)} e^{-\alpha W_{\tau_\alpha}} | \tau_\alpha = \nu \right] F_\alpha(d\nu).$$

Since $W_{\tau_\alpha} = \zeta(\tau_\alpha) - \alpha\tau_\alpha$ almost surely,

$$\begin{aligned}
F(s) &= \int_0^s e^{-\alpha(\zeta(\nu) - \alpha\nu)} E_P \left[e^{-\frac{\alpha^2 s}{2} - \alpha(W_s - W_\nu)} | \tau_\alpha = \nu \right] F_\alpha(d\nu) \\
&= \int_0^s e^{-\alpha\zeta(\nu) + \alpha^2\nu} E_P \left[e^{-\frac{\alpha^2 s}{2} - \alpha(W_s - W_\nu)} \right] F_\alpha(d\nu) \\
&= \int_0^s e^{-\alpha\zeta(\nu) + \alpha^2\nu} e^{-\frac{\alpha^2 s}{2}} E_P \left[e^{-\alpha(W_s - W_\nu)} \right] F_\alpha(d\nu) \\
&= \int_0^s e^{-\alpha\zeta(\nu) + \alpha^2\nu} e^{-\frac{\alpha^2 s}{2}} e^{\frac{\alpha^2}{2}(s-\nu)} F_\alpha(d\nu) \\
&= \int_0^s e^{-\alpha\zeta(\nu) + \alpha^2\nu - \frac{\alpha^2\nu}{2}} F_\alpha(d\nu) \\
&= \int_0^s e^{-\alpha\zeta(\nu) + \frac{\alpha^2\nu}{2}} F_\alpha(d\nu)
\end{aligned}$$

This implies,

$$\begin{aligned}
F(ds) &= e^{-\alpha\zeta(s) + \frac{\alpha^2 s}{2}} F_\alpha(ds) \\
e^{\alpha\zeta(s) - \frac{\alpha^2 s}{2}} F(ds) &= F_\alpha(ds) \\
\int_0^\infty e^{\alpha\zeta(s) - \frac{\alpha^2 s}{2}} F(ds) &= \int_0^\infty F_\alpha(ds) \\
&= F_\alpha(\infty) \\
&= 1
\end{aligned}$$

We differentiate both sides with respect to α

$$\begin{aligned}
\frac{d}{d\alpha} \int_0^\infty e^{\alpha\zeta(s) - \frac{\alpha^2 s}{2}} F(ds) &= 0 \\
\int_0^\infty \frac{d}{d\alpha} e^{\alpha\zeta(s) - \frac{\alpha^2 s}{2}} F(ds) &= 0 \\
\int_0^\infty (\zeta(s) - \alpha s) e^{\alpha\zeta(s) - \frac{\alpha^2 s}{2}} F(ds) &= 0
\end{aligned}$$

Let $\alpha \to 0$.

Then,

$$\int_0^\infty \zeta(s)F(ds) = 0$$
$$\int_0^\infty (a - b(s+c)^d)F(ds) = 0$$
$$b\int_0^\infty (s+c)^d F(ds) = \int_0^\infty aF(ds)$$
$$\int_0^\infty (s+c)^d F(ds) = \frac{a}{b}$$

Thus $E\left[(\tau+c)^d\right] = \frac{a}{b}$ □

Let us now consider some examples.

4.6.1. Example 1

Let us consider the linear boundary $\zeta(s) = -a + bs$, where $a > 0$, $b > 0$ and $(W_s)_{s\geq 0}$ be a standard Brownian motion.

We define the first passage time (from above) of W_s to $\zeta(s)$ by $\tau := inf\{s \geq 0 : W_s \leq \zeta(s)\}$.

Then by theorem 4.9, we have:

$$E[\tau] = \frac{a}{b}$$

To verify this result, we consider that according to [24], τ has the following density:

$$f(\tau) = \frac{|-a|}{\sqrt{2\pi\tau^3}} e^{-\frac{(a-b\tau)^2}{2\tau}}$$

which we may rewrite as follows,

$$f(\tau) = \frac{a}{\sqrt{2\pi\tau^3}} e^{-\frac{b^2(\tau-\frac{a}{b})^2}{2\tau}} \tag{27}$$
$$= \frac{a}{\sqrt{2\pi\tau^3}} e^{-\frac{a^2(\tau-\frac{a}{b})^2}{2(\frac{a}{b})^2\tau}} \sim Inverse\ Gaussian(\frac{a}{b}, a^2) \tag{28}$$

Thus, if follows that

$$E[\tau] = \frac{a}{b},$$

which matches our result.

Similarly, for the boundary $\zeta(s) = a - bs$, where $a > 0$, $b > 0$ and the standard Brownian motion $(W_s)_{s\geq 0}$, we consider the first passage time (from below) of W_s to $\zeta(s)$, $\tau := inf\{s \geq 0 : W_s \geq \zeta(s)\}$.

It follows from theorem 4.10 that

$$E[\tau] = \frac{a}{b},$$

and one can verify this similarly as shown above, due to symmetry of Brownian motions.

4.6.2. Example 2

Let us consider the square root boundary $\zeta(s) = -a + b\sqrt{(s+c)}$, where $a > 0$, $b > 0$, $b\sqrt{c} < a$ and $(W_s)_{s\geq 0}$ be a standard Brownian motion. We define the first passage time,τ, (from above) of W_s to $\zeta(s)$ as before. Then by theorem 4.9, we have:

$$E\left[(\tau + c)^{\frac{1}{2}}\right] = \frac{a}{b}$$

To check this claim, let us consider the following result, which is due to Novikov.

Let $\tau = inf\{s \geq 0 : W_s \leq -a + b\sqrt{s+c}\}$, where $a \geq 0, c \geq 0$ and $b\sqrt{c} < a$. Let $D_\nu(z)$ be the parabolic cylinder function of the parameter ν, and let z_ν be the largest root of the equation $D_\nu(z) = 0$ (If $\nu \leq 0$, then $D_\nu(z) > 0$ and $z_\nu = -\infty$).

THEOREM 4.11. [28], *Let $-\infty < \nu < \infty$. Then,*

$$E\left[(\tau + c)^{\frac{\nu}{2}}\right] = \begin{cases} \frac{c^{\nu/2} e^{a^2/4c - b^2/4c} D_\nu(a/\sqrt{c})}{D_\nu(b)}, & b > z_\nu \\ \infty, & b \leq z_\nu \end{cases}$$

For $\nu = n = 1, 2, ...$ and $c = 0$, this formula becomes

$$E\left[(\tau)^{\frac{\nu}{2}}\right] = \begin{cases} \frac{a^n}{He_n(b)}, & b > z_\nu \\ \infty, & b \leq z_\nu \end{cases}$$

where $He_n(z) = e^{z^2/4}D_n(z)$ *is the Hermite polynomial of order* n.

Our case corresponds to $\nu = 1$ in Novikov's formula. This gives the value of largest root as $Z_1 = -58.446$, which is less than zero. Since we are only considering $b > 0$, our result should match Novikov's formula.

Now, by Novikov's formula,

$$E\left[(\tau + c)^{\frac{1}{2}}\right] = \frac{c^{1/2}e^{a^2/4c - b^2/4c}D_1(a/\sqrt{c})}{D_1(b)}$$

Since $\nu = 1$ is a nonnegative integer, then $e^{z^2/4}D_1(z) = He_1(z)$, the Hermite polynomial of order 1.

Thus,

$$\begin{aligned} E\left[(\tau + c)^{\frac{1}{2}}\right] &= \frac{c^{1/2}He_1(a/\sqrt{c})}{He_1(b)} \\ &= \frac{c^{1/2}(-1)e^{(\frac{a}{\sqrt{c}})^2/2}\frac{d}{d(a/\sqrt{c})}e^{-(\frac{a}{\sqrt{c}})^2/2}}{(-1)e^{b^2/2}\frac{d}{db}e^{-b^2/2}} \\ &= \frac{c^{1/2}e^{(\frac{a}{\sqrt{c}})^2/2}e^{-(\frac{a}{\sqrt{c}})^2/2}(-\frac{2\frac{a}{\sqrt{c}}}{2})}{e^{b^2/2}e^{-b^2/2}(-\frac{2b}{2})} \\ &= \frac{a}{b}, \end{aligned}$$

which matches our result.

In the special case where $c = 0$, we simply have:

$$\begin{aligned} E\left[(\tau)^{\frac{1}{2}}\right] &= \frac{a}{He_1(b)} \\ &= \frac{a}{b}. \end{aligned}$$

4.7. Expectation of a Functional of the First Passage Time of Ornstein Uhlenbeck Processes to Some Class of Boundaries

We now extend our results in section 4.6 to the OU process.

THEOREM 4.12. *Given the boundary*

$$g(t) = \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\rho\lambda t}) + \mu,$$

of the OU process X_t, where, $\lambda > 0$, $\sigma > 0$, $\beta > \frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu) + \gamma \geq 0$, $\gamma > 0$ and $\rho > -1$, and the first passage time τ of X_t from x_0 to the $g(t)$,

$$E\left[e^{\tau(\rho+1)}\right] = \frac{\frac{\sqrt{2\lambda}}{\sigma}(-x_0 + \mu) + \beta}{\gamma} \tag{29}$$

PROOF. The equivalent boundary to $g(t)$ for the standardized OU process is $\tilde{g}(\tilde{t}) = \beta e^{-\lambda \tilde{t}} - \gamma e^{\rho \lambda \tilde{t}}$. The condition that $\beta > \frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu) + \gamma$ implies $g(0) > 0$.

By 9, the equivalent boundary for the standard Brownian motion is $\zeta(s) = \frac{\sqrt{2\lambda}}{\sigma}(-x_0 + \mu) + \beta - \gamma(1+s)^{\frac{\rho+1}{2}}$, where $s(t) = e^{2\tilde{t}} - 1$.

The conditions $\beta > \frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu) + \gamma$ and $\gamma > 0$ implies $\zeta(s) > 0$ and $\frac{\sqrt{2\lambda}}{\sigma}(-x_0 + \mu) + \beta > 0$.

Let θ be the first passage time of the standard Brownian motion to the boundary $\zeta(s)$. Then $P(\tau \leq t) = P(\theta \leq s)$.

Therefore,

$$\begin{aligned}
E\left[e^{\tau(\rho+1)}\right] &= E\left[(e^{2\tau})^{\frac{\rho+1}{2}}\right] \\
&= \int_0^\infty (e^{2\tau})^{\frac{\rho+1}{2}} P(d\tau) \\
&= \int_0^\infty (\theta+1)^{\frac{\rho+1}{2}} P(d\theta) \\
&= \frac{\frac{\sqrt{2\lambda}}{\sigma}(-x_0 + \mu) + \beta}{\gamma}, \quad \text{by theorem 4.10}
\end{aligned}$$

□

By a similar argument as above, an application of theorem 4.9 here, can be use to show that for the boundary

$$g(t) = \frac{\sigma}{\sqrt{2\lambda}}(-\beta e^{-\lambda t} + \gamma e^{\rho \lambda t}) + \mu$$

of X_t, where $\lambda > 0$, $\sigma > 0$, $\beta > \frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu) + \gamma \geq 0$, $\gamma > 0$ and $\rho > -1$,

$$E\left[e^{\tau(\rho+1)}\right] = \frac{\frac{\sqrt{2\lambda}}{\sigma}(x_0 - \mu) + \beta}{\gamma} \tag{30}$$

CHAPTER 5

SIMULATIONS

5.1. Zero Trend Ornstein-Uhlenbeck Process

5.1.1. The Process

The zero trend mean-reverting Ornstein-Uhlenbeck process satisfies the following system:

$$dX_t = \lambda(\mu - X_t)dt + \sigma dB_t, \quad t \geq 0 \tag{31}$$

$$X_0 = x_0, \quad \lambda > 0, \quad \sigma > 0$$

The solution to the above system can be obtained using Ito's lemma. The lemma can be found in various literature. See for instance [27].

Let $Y_t = e^{\lambda t} X_t$

Then by Ito's lemma A.3,

$$\begin{aligned} dY_t &= (\lambda e^{\lambda t} X_t + \lambda e^{\lambda t}(\mu - X_t))dt + \sigma e^{\lambda t} dB_t \\ &= \lambda \mu e^{\lambda t} dt + \sigma e^{\lambda t} dB_t \end{aligned}$$

Integrating both sides on the interval $0 \leq t \leq s$, we get,

$$\begin{aligned} \int_0^s dY_t &= \lambda\mu \int_0^s e^{\lambda t} dt + \sigma \int_0^s e^{\lambda t} dB_t \\ Y_s - Y_0 &= \left[\lambda\mu \frac{1}{\lambda} e^{\lambda t}\right]_0^s + \sigma \int_0^s e^{\lambda t} dB_t \end{aligned}$$

Thus,

$$\begin{aligned} e^{\lambda t} X_t - X_0 &= \mu e^{\lambda t} - \mu + \sigma \int_0^t e^{\lambda \nu} dB_\nu \\ X_t &= e^{-\lambda t} x_0 + \mu(1 - e^{-\lambda t}) + \sigma \int_0^t e^{-\lambda(t-\nu)} dB_\nu \end{aligned}$$

We may similarly integrate over the time interval $t-1 \leq \nu \leq t$, where $\Delta t = t-(t-1)$ is a unit change in time, to obtain:

$$\begin{aligned}
\int_{t-1}^{t} dY_\nu &= \lambda\mu \int_{t-1}^{t} e^{\lambda\nu} d\nu + \sigma \int_{t-1}^{t} e^{\lambda\nu} dB_\nu \\
Y_t - Y_{t-1} &= \left[\lambda\mu \frac{1}{\lambda} e^{\lambda\nu}\right]_{t-1}^{t} + \sigma \int_{t-1}^{t} e^{\lambda\nu} dB_\nu \\
e^{\lambda t} X_t - e^{\lambda(t-1)} X_{t-1} &= \mu e^{\lambda t} - \mu e^{\lambda(t-1)} + \sigma \int_{t-1}^{t} e^{\lambda\nu} dB_\nu \\
X_t &= e^{-\lambda(t-(t-1))} X_{t-1} + \mu - \mu e^{-\lambda(t-(t-1))} + \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu \\
&= e^{-\lambda\Delta t} X_{t-1} + \mu(1 - e^{-\lambda\Delta t}) + \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu \qquad (32)
\end{aligned}$$

We consider the Ito integral $\int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu$ in equation 32 above. Since $e^{-\lambda(t-\nu)}$ is a deterministic function of ν which does not depend on B_ν, it follows that the Ito integral, $\int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu$ follows a normal distribution with zero mean and variance obtained as follow:

$$\begin{aligned}
Var\left[\int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu\right] &= E\left[\left(\int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu\right)^2\right] \\
&= E\left[\int_{t-1}^{t} e^{-2\lambda(t-\nu)} d\nu\right] \qquad (33) \\
&= \left[\frac{1}{2\lambda} e^{-2\lambda(t-\nu)}\right]_{t-1}^{t} \\
&= \frac{1}{2\lambda}(1 - e^{-2\lambda\Delta t}), \qquad (34)
\end{aligned}$$

where we have used the isometry property of Ito integral in equation 33.

It then follows from equations 32 and 34 that:

$$X_t | X_{t-1} \sim N\left(e^{-\lambda\Delta t} X_{t-1} + \mu(1 - e^{-\lambda\Delta t}), \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t})\right). \qquad (35)$$

A discretized form of the OU process X_t can be obtain from the distribution in 35 above.

We consider an example of this process with parameters $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$. We also pick an initial value of $X_0 = 2.4$, which is not too far from the long term

mean of 2.3. Using the discretized solution above, we form a path of length ten thousand. A plot of this path is presented in figure (5.1). As one would notice, the initial value of X_t, thus $X_0 = 2.4$, deviates from the long term mean of the process, $\mu = 2.3$. As such the process quickly reverts to the long term mean. There are also instances along the path that the process deviates from the long term mean, and whenever this happens, we see a reversion back to μ.

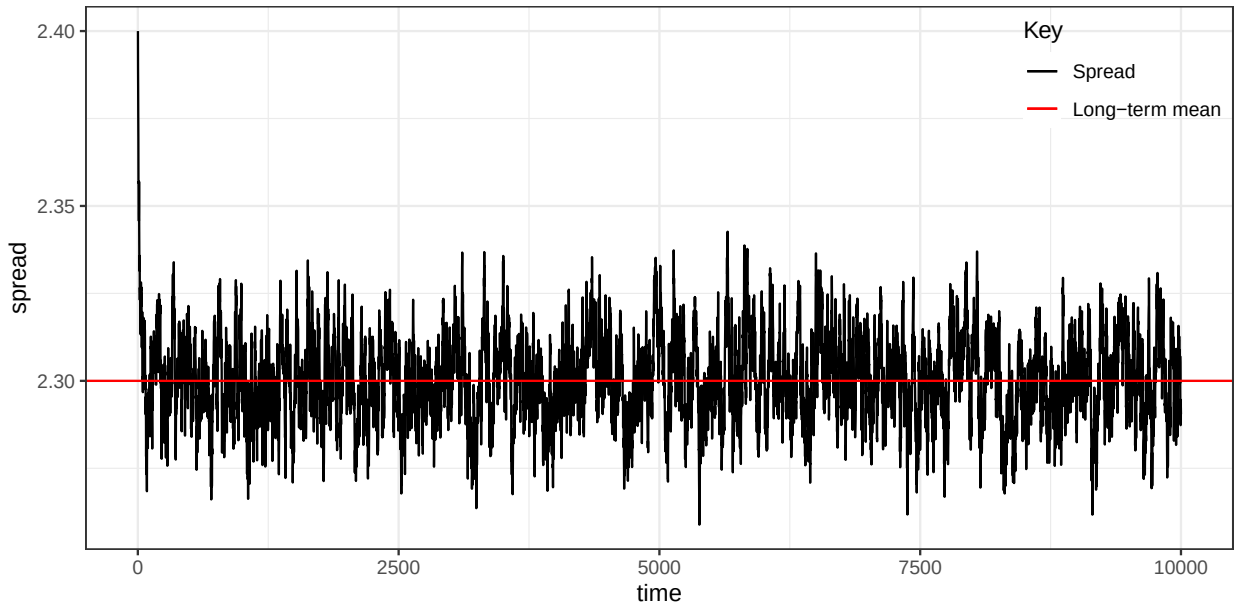

FIGURE 5.1. Discretized Ornstein-Uhlenbeck Process of length 10000, with parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$, starting from $x_0 = 2.4$

5.1.2. Parameter Estimation

In this subsection, we consider how to estimate the parameters of the model from a given data set. We will use two methods, namely the maximum likelihood and the least squares methods.

5.1.2.1. Maximum Likelihood Estimator

The likelihood function can be obtained from 35 as follows:

$$L(\mu,\lambda,\sigma|X_t)=\prod_{t=1}^{N}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})}}e^{-\frac{(X_t-(e^{-\lambda\Delta t}X_{t-1}+\mu(1-e^{-\lambda\Delta t})))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}}\right)$$

We then obtain the log-likelihood as:

$$l(\mu,\lambda,\sigma|X_t)=log\prod_{t=1}^{N}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})}}e^{-\frac{(X_t-(e^{-\lambda\Delta t}X_{t-1}+\mu(1-e^{-\lambda\Delta t})))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}}\right)$$

(36)

$$=-\frac{N}{2}log(2\pi)-\frac{N}{2}log(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))-\frac{\lambda}{\sigma^2(1-e^{-2\lambda\Delta t})}\sum_{t=1}^{N}(X_t-e^{-\lambda\Delta t}X_{t-1}-\mu(1-e^{-\lambda\Delta t}))^2$$

The maximum likelihood estimates for μ, λ and σ will be the values $\hat{\mu}$, $\hat{\lambda}$ and $\hat{\sigma}$ respectively, that maximize the above expression. We will solve this numerically from equation 36, using the maxLik package in R.

5.1.2.2. Method of Least Squares

While we acknowledge that this is a nonlinear least squares problem, from several numerical examples, we observed that the linear least squares estimate approximate the nonlinear least squares estimates up to eight decimal places. Thus we will approximate the optimization solution with linear least squares method throughout this work. This also saves us from having to deal with complicated hessian matrices that result from the nonlinear least squares optimization problem, particularly in proving its positive definiteness.

We formulate the least squares regression equation from equation 32 as follows:

(37) $$X_t=\zeta+\psi X_{t-1}+\epsilon_t,\quad \epsilon_t\sim i.i.d.N(0,\sigma_\epsilon)$$

where

(38) $$\zeta=\mu(1-e^{-\lambda\Delta t})$$

(39) $$\psi=e^{-\lambda\Delta t}$$

$$\sigma_\epsilon^2 = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t}) \tag{40}$$

Substituting the least squares estimates of equation 37, $\hat{\zeta}$, $\hat{\psi}$ and $\hat{\sigma}_\epsilon$, into equations 38, 39 and 40, and solving the system gives the least squares parameter estimates for the OU model as:

$$\hat{\lambda} = -\frac{log\hat{\psi}}{\Delta t}$$

$$\hat{\mu} = \frac{\hat{\zeta}}{1 - \hat{\psi}}$$

$$\hat{\sigma} = \frac{\sqrt{2\hat{\lambda}}\hat{\sigma}_\epsilon}{\sqrt{1 - \hat{\psi}^2}}$$

5.1.3. Long Path

In this subsection, we will look into estimating the OU process parameters from a long dicretized path. We also perform parameter estimation using both maximum likelihood estimation and the method of least squares. We then perform Monte Carlo simulations to test the robustness of the estimation methods discussed in the previous subsection.

5.1.3.1. Generating the Spread

We use the last 1260 points of the example presented at the beginning of the section and in figure 5.1 in our study. This corresponds to 5 years of trading data in the US financial market, if we consider each time step as a day. A graph of this path can be found in figure 5.2. We observe that the process always oscillates around the long term mean, and whenever there is a deviation, it reverts back to the mean.

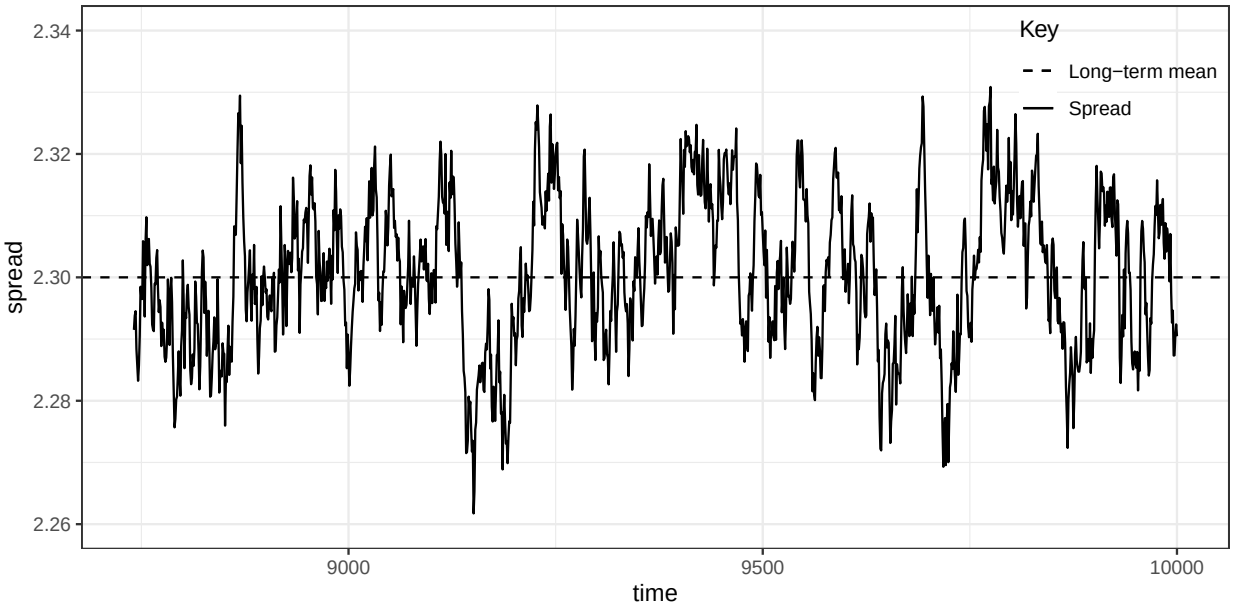

FIGURE 5.2. Discretized Ornstein-Uhlenbeck Process of length 1260, with parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

5.1.3.2. Parameter Estimation

From the path generated above, we can estimate the parameters of the OU process to see how good the model performs. The result of this is shown in table 5.1. We see that given enough data points, both the maximum likelihood and the least squares methods return good estimates for all three parameters.

Figure 5.3 presents a comparison of the true path of the OU process with the parameters $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$, with the paths produced from the estimates of the maximum likelihood and least squares methods. One can hardly distinguish among the paths, which shows how well the estimates perform.

5.1.3.3. Monte Carlo Simulation

We perform a Monte Carlo simulation with of the OU process with the parameters $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$, using both the maximum likelihood estimator and the method of least squares. The result of this simulation can be seen in table 5.2. Both methods estimate all three parameters very well, and are robust.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	2.30000000	MLE	2.30048740	0.00048740	0.00155000	0.00000264
		LS	2.30048494	0.00048494		
λ	0.08000000	MLE	0.09484300	0.01484300	0.01288570	0.00038636
		LS	0.09486643	0.01486643		
σ	0.00500000	MLE	0.00521410	0.00021410	0.00010890	0.00000006
		LS	0.00521647	0.00021647		

TABLE 5.1. Parameter estimation for the discretized OU process of length 1260 by maximum likelihood and method of least squares, with true parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

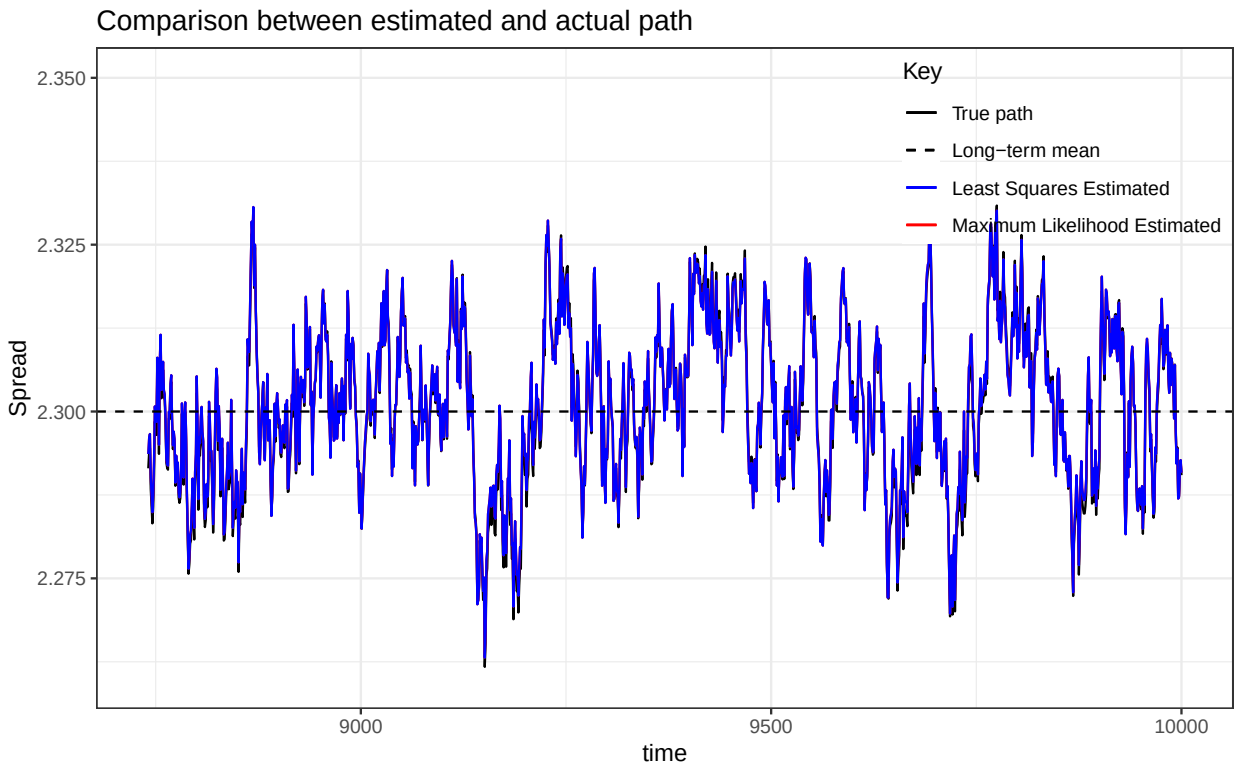

FIGURE 5.3. Paths of length 1260, generated from one realization of parameter estimates from maximum likelihood and least squares methods, with true parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	2.30000000	MLE	2.30021012	0.00021012	0.00144191	0.0000021
		LS	2.30000084	0.00000084	0.00178205	0.00000318
λ	0.08000000	MLE	0.08256713	0.00256713	0.01220114	0.00015546
		LS	0.08349471	0.00349471	0.01231287	0.00016382
σ	0.00500000	MLE	0.00500196	0.00000196	0.00010335	0.00000001
		LS	0.00500611	0.00000611	0.00010417	0.00000001

TABLE 5.2. Performance of Maximum Likelihood and Least Squares methods for estimating the OU parameters, by 10,000 Monte Carlo simulations, using a path of length 1260

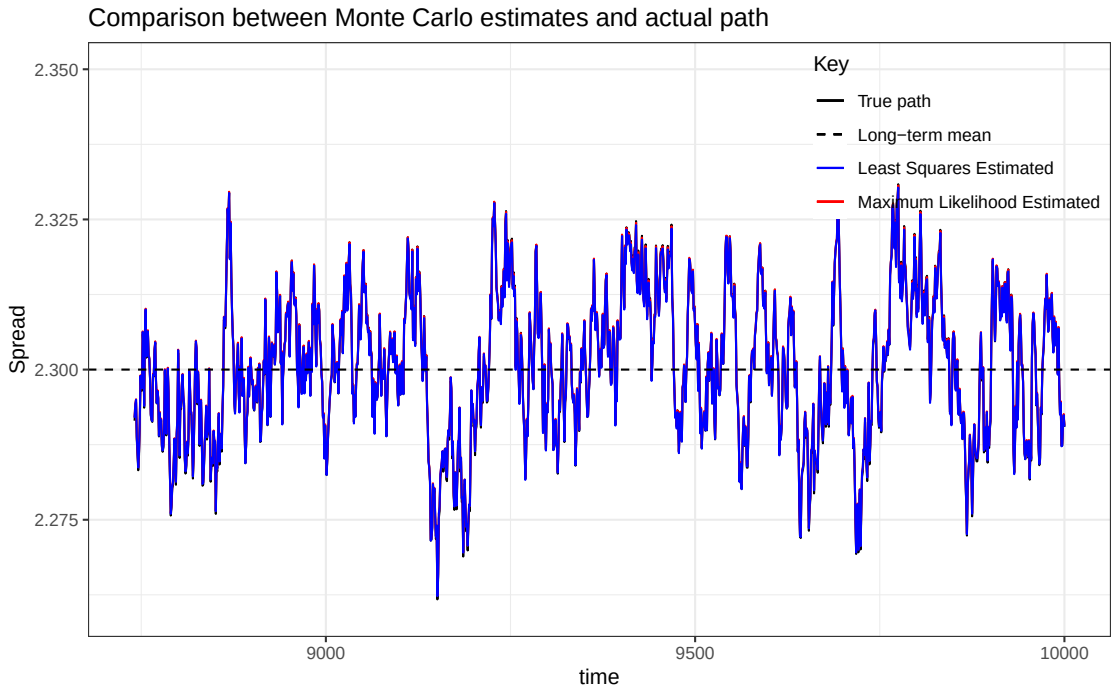

FIGURE 5.4. Paths generated from parameter estimates by Maximum Likelihood and Method of Least Squares, using Monte Carlo averages with 10,000 replications of paths of length 1260. True parameter values are $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

5.1.4. Short Path

We also consider a short path. Similar to subsection 5.1.3, we look into the parameter estimation of the OU process from such paths, using both the maximum likelihood and the least squares methods, and then perform Monte Carlo simulations to determine the robustness of the estimators.

5.1.4.1. Generating the Spread

In this case, we consider the last 126 steps of the path generated at the beginning of the section. A graph of this path is provided in figure 5.5. Again we see that the process oscillates around the long term mean.

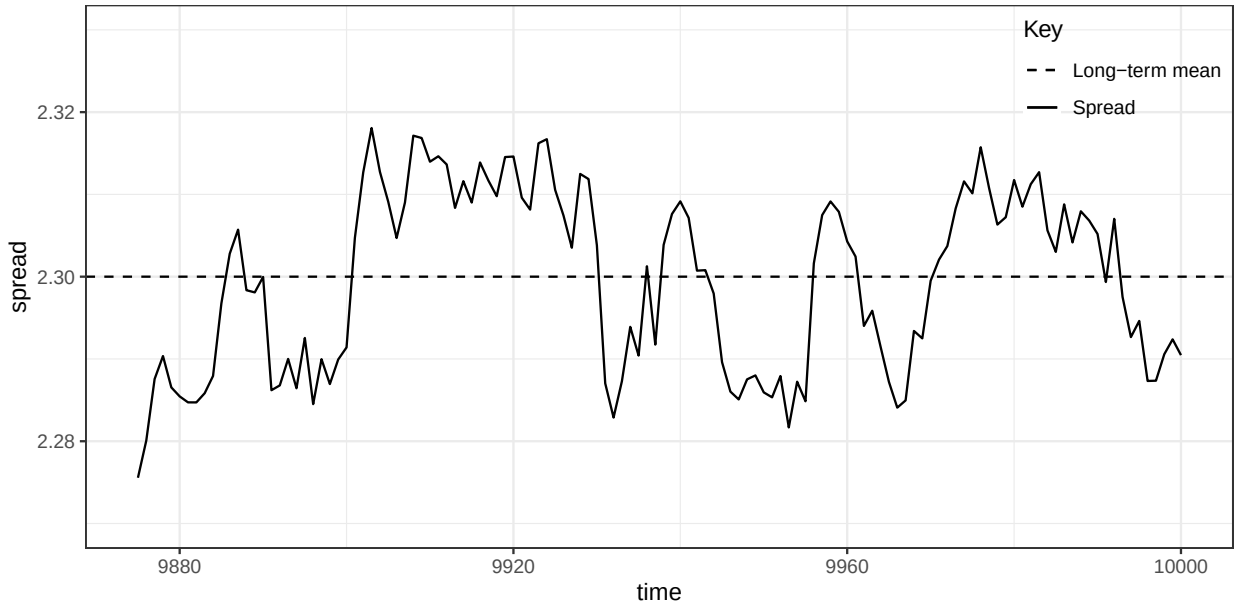

FIGURE 5.5. Discretized Ornstein-Uhlenbeck Process of length 126, with parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

5.1.4.2. Parameter Estimation

Both the maximum likelihood estimators and the method of least squares are employed to estimate the parameters μ, λ and σ of the OU process. From table 5.3, we see that both methods give good estimates of the μ and σ, but do not do very well for λ. A graphical comparison of the true path and the paths produced from the estimates is shown in figure

5.6. While one can hardly distinguish between the estimated paths, we notice however that they both differ slightly from the true path of the process.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	2.30000000	MLE	2.30049290	0.00049290	0.00329950	0.00001113
		LS	2.30048086	0.00048086		
λ	0.08000000	MLE	0.15094150	0.07094150	0.04952930	0.00748585
		LS	0.15091469	0.07091469		
σ	0.00500000	MLE	0.00554470	0.00054470	0.00037440	0.00000044
		LS	0.00556660	0.00056660		

TABLE 5.3. Parameter estimation for the discretized OU process of length 126 by maximum likelihood and method of least squares, with true parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

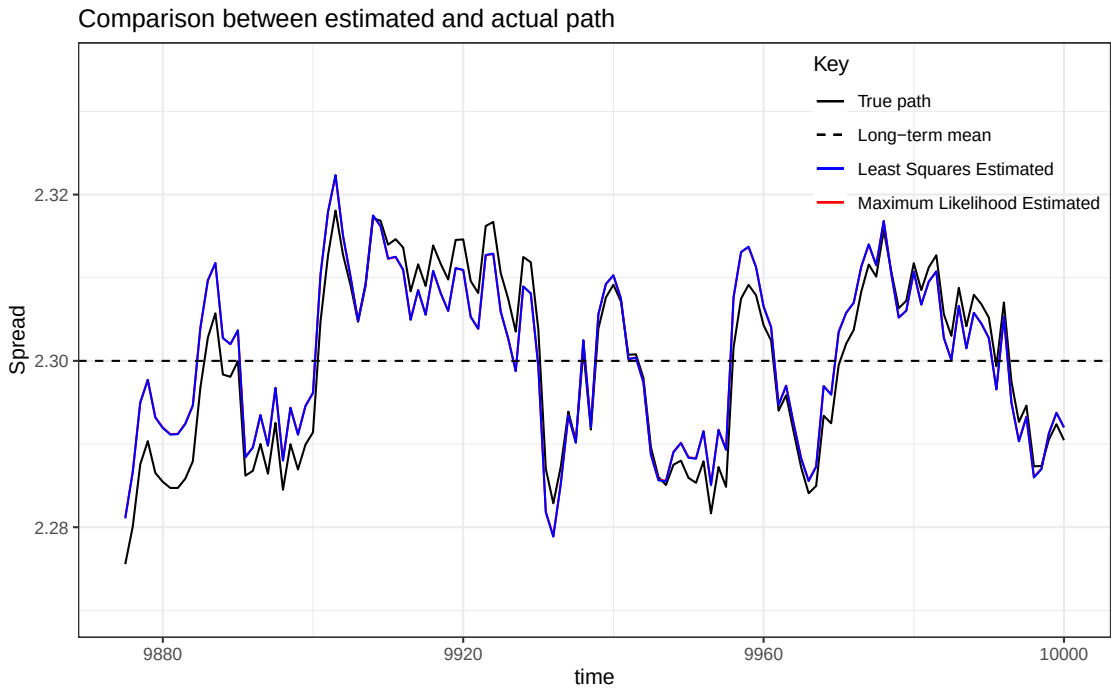

FIGURE 5.6. Paths of length 126, generated from one realization of estimates from maximum likelihood and least squares methods, with true parameter values $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$.

5.1.4.3. Monte Carlo Simulation

Like we did with the long path, we similarly perform Monte Carlo simulations for the short path using both the maximum likelihood estimator and the least squares method. The results are presented in table 5.4. We notice here that both methods are robust in estimating the parameters μ and σ, but on the other hand, they do not perform very well in estimating λ. However, this does not pose much of a problem, since our goal is not to predict the path perfectly, but rather obtain good enough estimates to determine an optimal threshold.

5.1.5. Artificial Stocks

In this subsection, we will test our methods on an artificial pair. In finance, stock price processes are often assumed to follow a geometric Brownian motion (GBM), [34]. So, to test the performance of our model and parameter estimation techniques in a pseudo-practical situation, we will assume we have a stock price time series Q_t that follows the GBM, and then assuming X_t is an OU process, we will derive the appropriate time series P_t for the price of a second stock P, such that P_t and Q_t are cointegrated. We will then perform our parameter estimation on the resulting data sets and also employ Monte Carlo simulation to assess the average performance of our estimators.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	2.30000000	MLE	2.30020981	0.00020981	0.01982779	0.00039319
		LS	2.29999153	-0.00000847	0.00562541	0.00003165
λ	0.08000000	MLE	0.11421728	0.03421728	0.05234302	0.00391061
		LS	0.11714017	0.03714017	0.05267428	0.00415397
σ	0.00500000	MLE	0.00503053	0.00003053	0.00033744	0.00000011
		LS	0.00505308	0.00005308	0.00033910	0.00000012

TABLE 5.4. Performance of Maximum Likelihood and Least Squares methods for estimating the OU parameters, by 10,000 Monte Carlo simulations, using a path of length 126

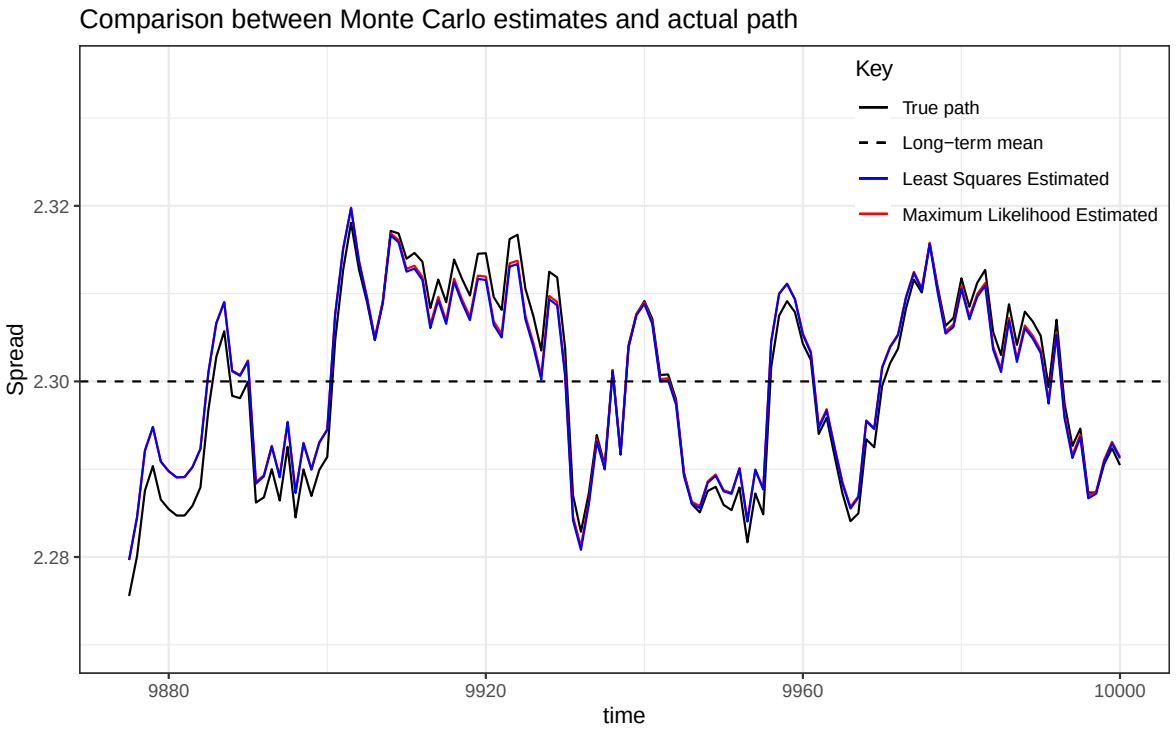

FIGURE 5.7. OU paths from parameter estimation by Maximum Likelihood and Method of Least Squares, using Monte Carlo averages with 10,000 replications of paths of length 126. The true parameter values are $\mu = 2.3$, $\lambda = 0.08$ and $\sigma = 0.005$

First, Q_t satisfies the following stochastic differential equation.

$$dQ_t = \kappa Q_t dt + \sigma Q_t dB_t, \tag{41}$$

where κ is the percentage drift and σ is the percentage volatility.

Let $Y_t = lnQ_t$.

By Ito's lemma A.3,

$$dY_t = \left(\kappa Q_t \frac{1}{Q_t} + \frac{\sigma^2 {Q_t}^2}{2}\left(-\frac{1}{{Q_t}^2}\right)\right) dt + \sigma Q_t \frac{1}{Q_t} dB_t$$

Thus,

$$\int_{t-1}^{t} dY_\nu = \int_{t-1}^{t} \left(\kappa - \frac{\sigma^2}{2}\right) d\nu + \sigma \int_{t-1}^{t} dB_\nu$$

$$Y_t - Y_{t-1} = \left[\left(\kappa - \frac{\sigma^2}{2}\right)\nu\right]_{t-1}^{t} + \sigma(B_t - B_{t-1})$$

$$lnQ_t - lnQ_{t-1} = \left(\kappa - \frac{\sigma^2}{2}\right)(t-(t-1)) + \sigma(B_t - B_{t-1})$$

$$lnQ_t = lnQ_{t-1} + \left(\kappa - \frac{\sigma^2}{2}\right)(t-(t-1)) + \sigma(B_t - B_{t-1}),$$

$$B_t - B_{t-1} \sim \text{ i.i.d. } N(0, \Delta t) \tag{42}$$

5.1.5.1. Generating the Pair

A discretized form of the the GBM can be obtained from equation 42. In our example, we use the parameters $\kappa = 0.001$ and $\sigma = 0.02$ to generate a logarithm of the price time series Q_t starting at $ln(Q_0) = 1.5$ and including 10,000 steps. The plot is shown in figure 5.8. We also use relation 35 to generate a 10,000-step OU process that starts at $x_0 = 2.35$, with parameters $\mu = 2.5$, $\lambda = 0.09$ and $\sigma = 0.013$. The graph can be found in figure 5.9. We then combine these two processes using equation 1.7, with $\eta = 0.3$, to obtain the logarithm of the prices time series P_t of a second stock P, such that $ln(P_t)$ and $ln(Q_t)$ are cointegrated. $ln(P_t)$ is shown in figure 5.10. We will take the time unit Δt to be a day.

We know from equation 32 that:

$$X_t = e^{-\lambda\Delta t}X_{t-1} + \mu(1 - e^{-\lambda\Delta t}) + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu,$$

and by equation 1.7, it follows that:

$$ln(P_t) - \eta ln(Q_t) = e^{-\lambda\Delta t}(ln(P_{t-1}) - \eta ln(Q_{t-1})) + \mu(1 - e^{-\lambda\Delta t}) + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu$$

$$ln(P_t) = \mu(1-e^{-\lambda\Delta t}) + \eta ln(Q_t) + e^{-\lambda\Delta t}ln(P_{t-1}) - \eta e^{-\lambda\Delta t}ln(Q_{t-1}) + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu \tag{43}$$

Thus,

(44)

$$ln(P_t)|ln(P_{t-1}) \sim N\left(\mu(1 - e^{-\lambda\Delta t}) + \eta ln(Q_t) + e^{-\lambda\Delta t}ln(P_{t-1}) - \eta e^{-\lambda\Delta t}ln(Q_{t-1}), \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t})\right),$$

which gives the following log-likelihood:

$$l(\eta,\mu,\lambda,\sigma|lnP_t,lnQ_t,lnP_{t-1},lnQ_{t-1})$$
$$= log\prod_{t=1}^{N}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})}}e^{-\frac{(ln(P_t)-\mu(1-e^{-\lambda\Delta t})-\eta ln(Q_t)-e^{-\lambda\Delta t}ln(P_{t-1})+\eta e^{-\lambda\Delta t}ln(Q_{t-1})))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}}\right),$$

$$(45)\quad = -\frac{N}{2}log(2\pi)-\frac{N}{2}log(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))$$
$$-\frac{\lambda}{\sigma^2(1-e^{-2\lambda\Delta t})}\sum_{t=1}^{N}(ln(P_t)-\mu(1-e^{-\lambda\Delta t})-\eta ln(Q_t)-e^{-\lambda\Delta t}ln(P_{t-1})+\eta e^{-\lambda\Delta t}ln(Q_{t-1})))^2$$

The maxLik package in R will be used to numerically solve for optimal values of the parameters η, μ, λ and σ for a given data set.

We may also obtain the least squares estimates as follows:

First, we formulate the least squares regression equation from equation 43

$$(46)\qquad ln(P_t)=\zeta+\theta ln(P_{t-1}+\eta ln(Q_t)+\psi Q_{t-1}+\epsilon_t,\quad \epsilon_t\sim i.i.d.N(0,\sigma_\epsilon)$$

where

$$(47)\qquad \zeta=\mu(1-e^{-\lambda\Delta t})$$

$$(48)\qquad \theta=e^{-\lambda\Delta t}$$

$$(49)\qquad \sigma_\epsilon^2=\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})$$

Substituting the least squares estimates of equation 46, $\hat{\zeta}$, $\hat{\alpha}$, $\hat{\theta}$, $\hat{\psi}$ and $\hat{\sigma_\epsilon}$, into equations 47, 48 and 49, and solving the system gives the least squares parameter estimates for the OU model as:

$$\hat{\lambda}=-\frac{log\hat{\theta}}{\Delta t}$$

$$\hat{\mu} = \frac{\hat{\zeta}}{1-\hat{\theta}}$$

$$\hat{\sigma} = \frac{\sqrt{2\hat{\lambda}}\hat{\sigma}_\epsilon}{\sqrt{1-\hat{\theta}^2}}$$

To see the performance of our parameter estimation methods, we will apply them to both a long path and a short one.

FIGURE 5.8. $\ln(Q_t)$ generated from Geometric Brownian Motion of length 10,000, with parameters $\kappa = 0.001$ and $\sigma = 0.02$, starting at 1.5

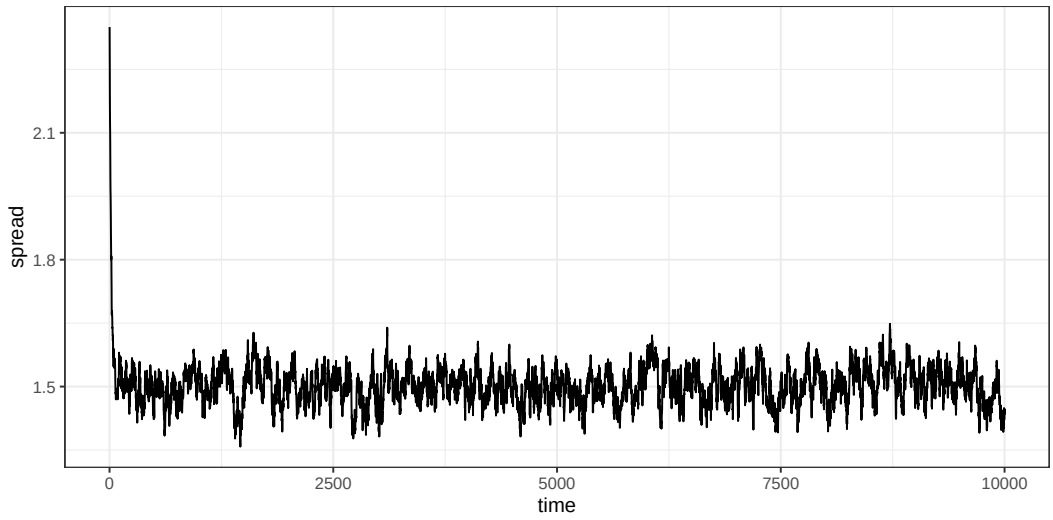

FIGURE 5.9. Spread of length 10,000 from an OU process, with parameters $\mu = 2.5$, $\lambda = 0.09$ and $\sigma = 0.013$, start at $x_0 = 2.35$

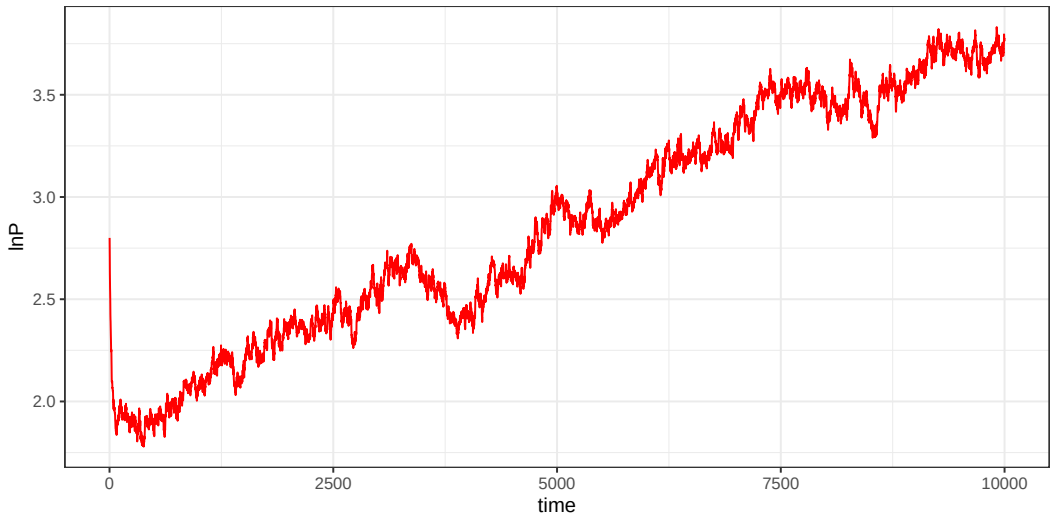

FIGURE 5.10. $ln(P_t)$ obtained from $ln(Q_t)$ and the OU process spread by the cointegration equation.

5.1.5.2. Long Path

Let us first consider the last 1260 days of the process generated in subsection 5.1.5.1. The OU process and the logarithm of the price time series are presented in figures 5.11 and 5.12 respectively.

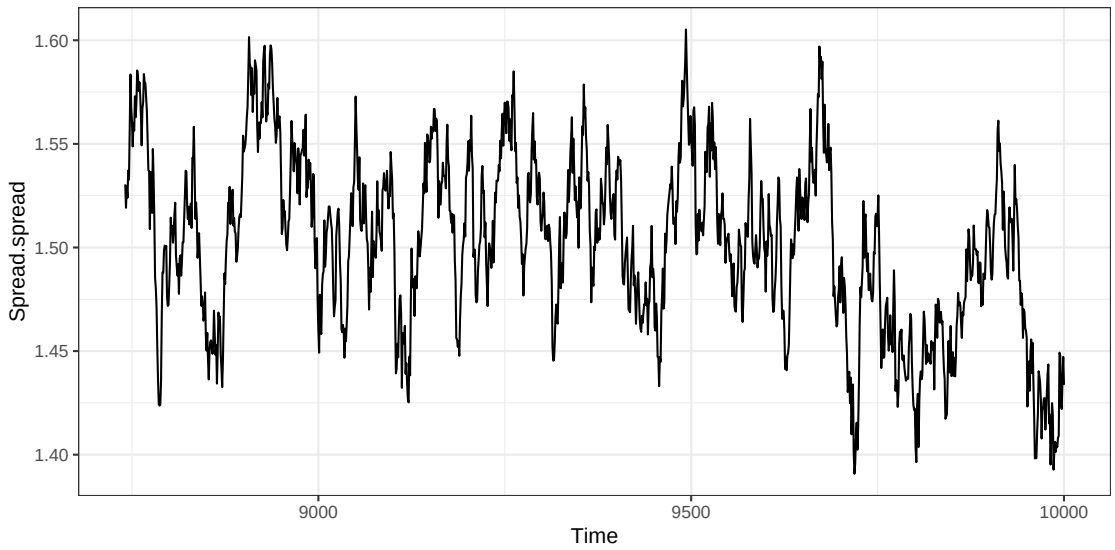

FIGURE 5.11. Last 1,260 steps of the OU process in figure 5.9

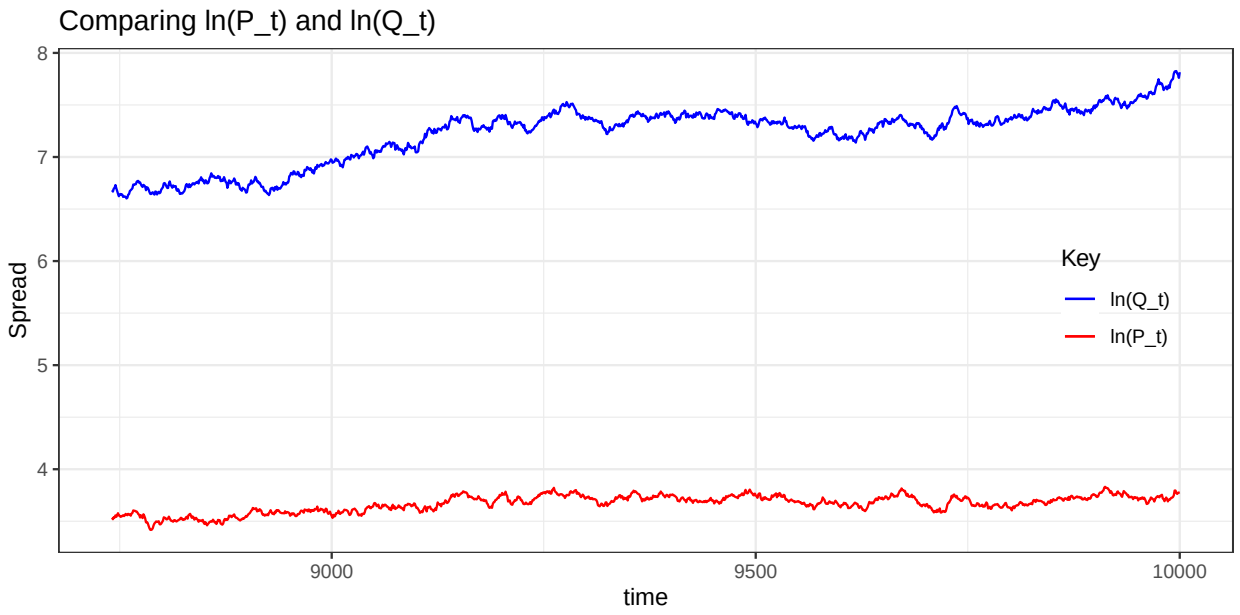

FIGURE 5.12. Last 1,260 steps of $ln(P_t)$ vs. $ln(Q_t)$ from figures 5.8 and 5.10

Linear model type	Lag	ADF	p-value
Type 1: no drift, no trend	0	-5.81	≤ 0.01
Type 2: with drift, no trend	0	-5.81	≤ 0.01
Type 3: with drift and trend	0	-5.91	≤ 0.01

TABLE 5.5. Augmented Dickey-Fulller test for stationarity of the cointegration between logarithmic returns of the artificial pair P_t and Q_t.

Note 1: Alternate hypothesis: Stationary

5.1.5.3. Parameter Estimation

By the augmented Dickey-Fuller test result presented in table 5.5, it is reasonable to assume the processes $ln(P_t)$ and $ln(Q_t)$ are cointegrated. Thus $X_t = ln(P_t) - \eta ln(Q_t)$ is a stationary process. So we estimate the parameters η, μ, λ and σ by maximum likelihood method based on equation 45 and also by the method of least squares. A comparison of the outputs is presented in table 5.6. Based on these parameter values, we reproduce the OU process and plot them in figure 5.13. We observe here that both the maximum likelihood and

least squares methods overestimate the μ, the level of the long term mean, but the maximum likelihood method performed better. This is not very surprising, given that we only picked one instance of the process. To better understand the performance of the model, we perform Monte Carlo simulations with 10,000 sample paths and obtain the average estimates for the parameters, using both maximum likelihood and least squares methods. Table 5.7 shows the result. We again reproduce the paths based on these averages, and they are plotted in figure 5.14. We observe here that the estimates perform very well and it is almost impossible to distinguish among the true path, the path from the maximum likelihood estimates and the path from the least squares estimates.

5.1.5.4. Short Path

We now turn our attention to the last 126 days of the processes discussed in 5.1.5.1. Similar to subsection 5.1.5.2, we show the OU process and the logarithm of the price time series of the two artificial stocks in figures 5.15 and 5.16 respectively.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.28454530	-0.01545470	0.01557660	0.00048148
		LS	0.29773916	-0.00226084		
μ	1.50000000	MLE	1.61252270	0.11252270	0.11299430	0.02542907
		LS	1.82803207	0.32803207		
λ	0.05000000	MLE	0.05121930	0.00121930	0.00966090	0.00009482
		LS	0.05324758	0.00324758		
σ	0.01300000	MLE	0.01318950	0.00018950	0.00027020	0.00000011
		LS	0.01319774	0.00019774		

TABLE 5.6. Parameter estimation for the OU process representing the spread between lnP_t and lnQ_t, by maximum likelihood and method of least squares.

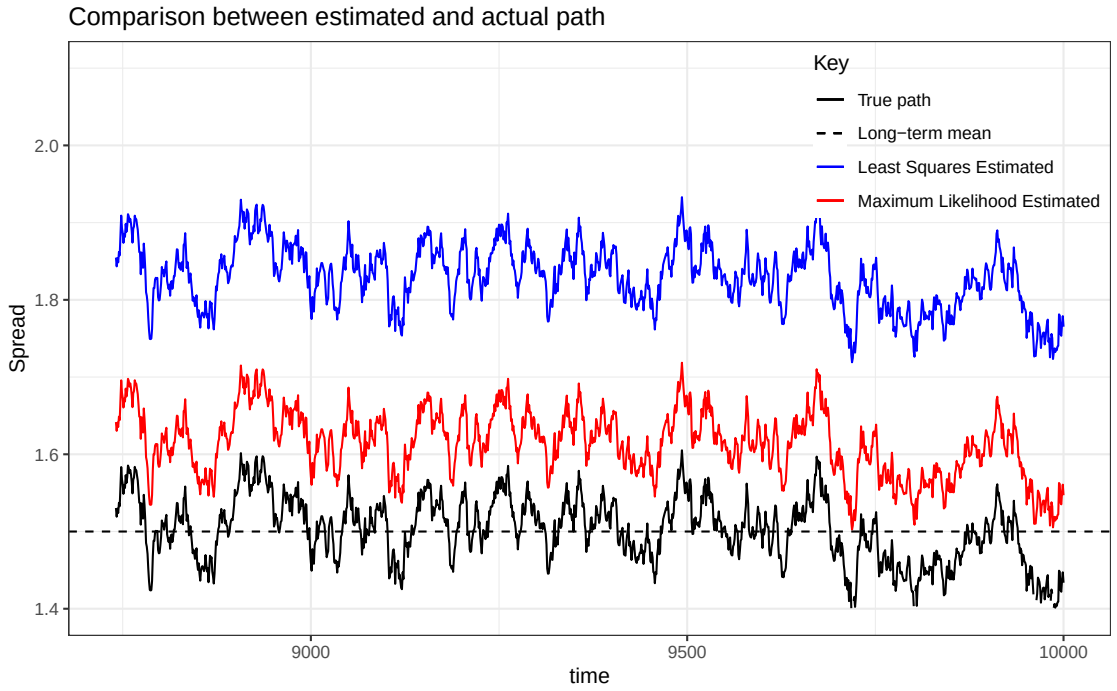

FIGURE 5.13. Spreads of length 1260 each, generated from one realization of parameter estimates for the OU process, from maximum likelihood and least squares methods, using $ln(Q_t)$ and $ln(P_t)$.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.30009965	0.00009965	0.01477949	0.00021844
		LS	0.29997765	-0.00002235	0.01752722	0.00030720
μ	2.50000000	MLE	1.49889455	-0.00110545	0.10779558	0.01162111
		LS	1.50040114	0.00040114	0.18946277	0.03589630
λ	0.09000000	MLE	0.05400583	0.00400583	0.01032903	0.00012274
		LS	0.05508819	0.00508819	0.01011371	0.00012818
σ	0.01300000	MLE	0.01299313	-0.00000687	0.00026786	0.00000007
		LS	0.01301512	0.00001512	0.00026522	0.00000007

TABLE 5.7. Maximum Likelihood and Least Squares estimates of the OU parameters for the path of length 1260, by 10,000 Monte Carlo simulations

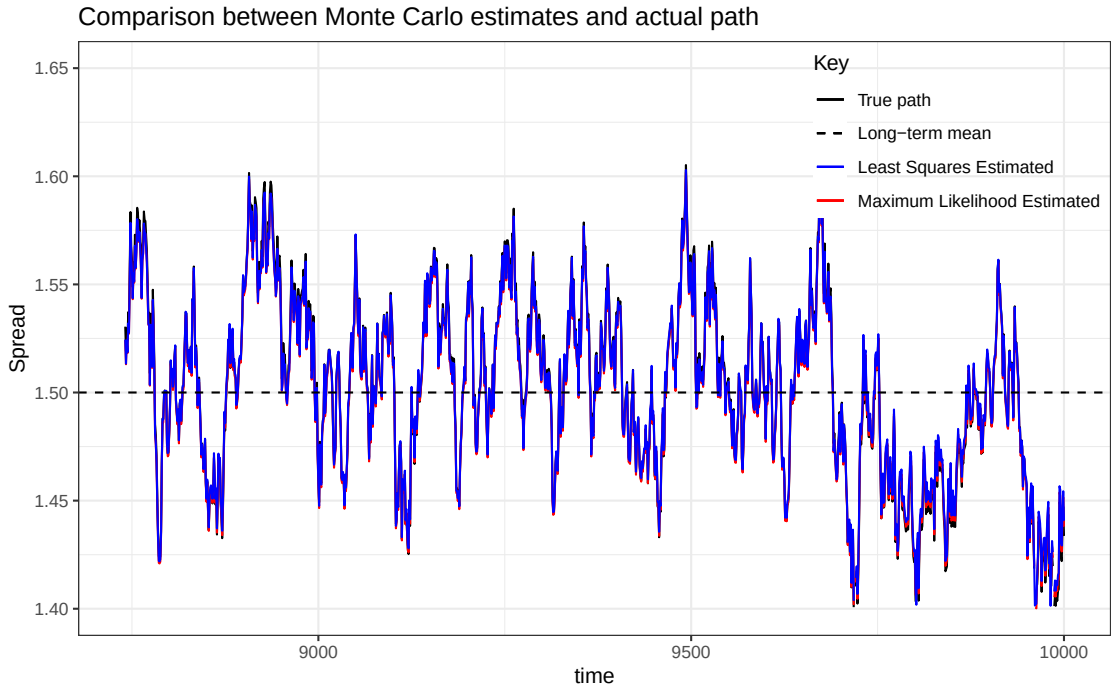

FIGURE 5.14. Spreads of length 1260 each, generated from parameter estimates by maximum likelihood and least squares methods from $ln(Q_t)$ and $ln(P_t)$, based on Monte Carlo averages.

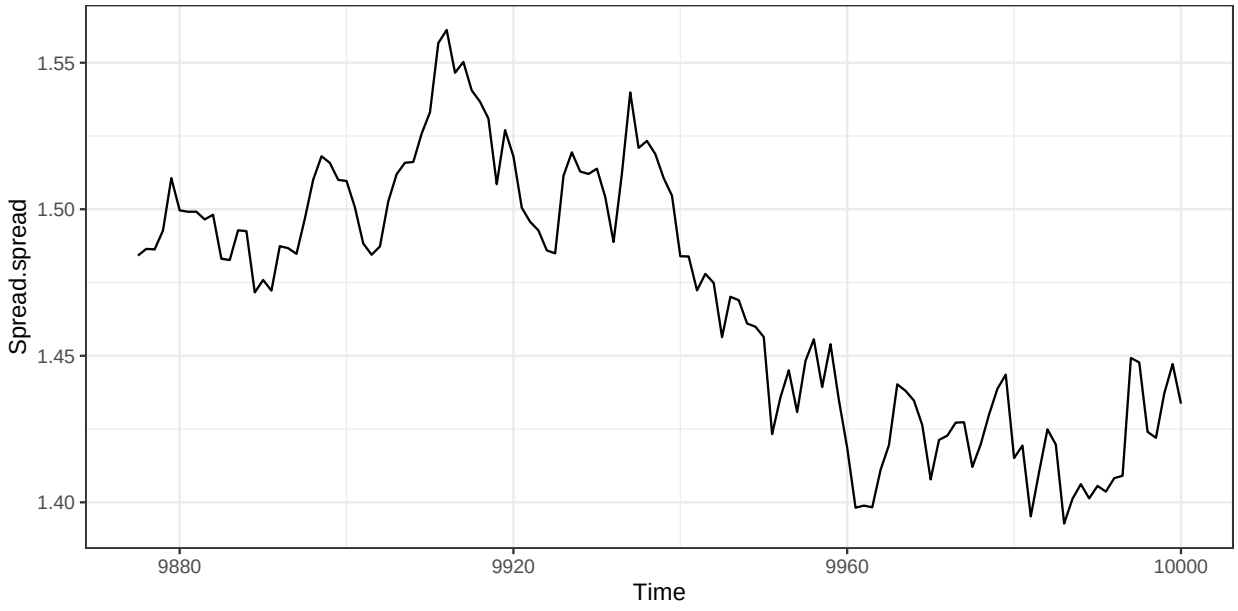

FIGURE 5.15. Last 126 steps of the OU process, with parameter values $\mu = 2.5$, $\lambda = 0.09$ and $\sigma = 0.013$

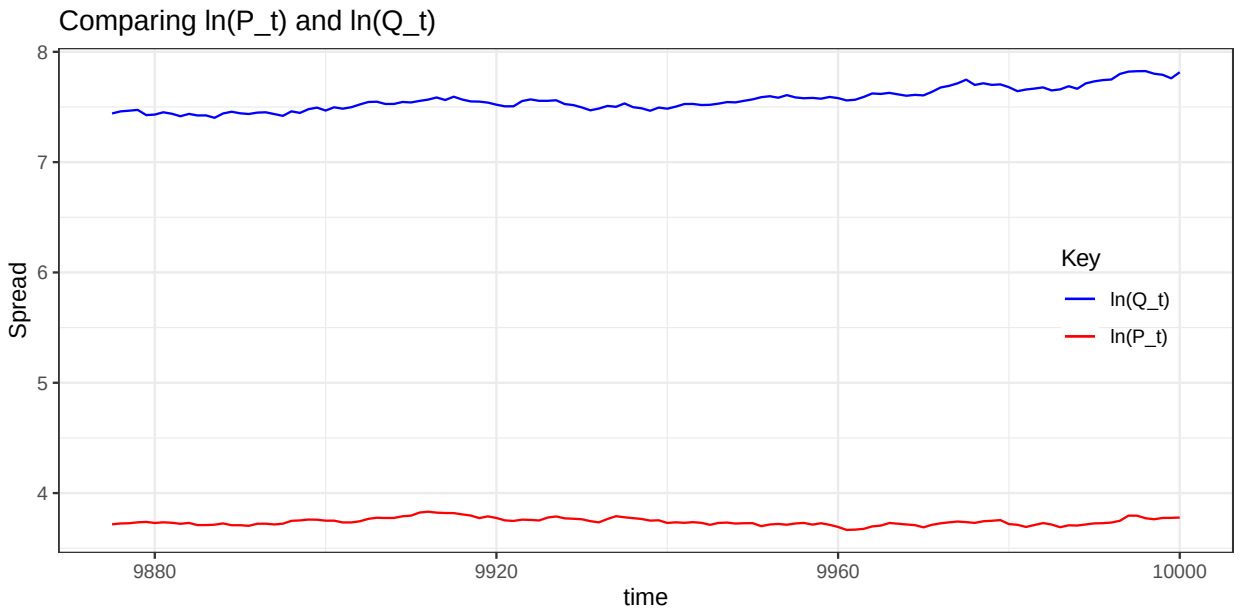

FIGURE 5.16. Last 126 steps of $ln(P_t)$ vs. $ln(Q_t)$ from figures 5.8 and 5.10

Linear model type	Lag	ADF	p-value
Type 1: no drift, no trend	0	-2.32	0.0218
Type 2: with drift, no trend	0	-2.30	0.21
Type 3: with drift and trend	0	-2.35	0.428

TABLE 5.8. Augmented Dickey-Fulller test for stationarity of the cointegration between logarithmic returns of the artificial pair P_t and Q_t.

Note 1: Alternate hypothesis: Stationary

5.1.5.5. Parameter Estimation

The augmented Dickey-Fuller test result for the process $X_t = ln(P_t) - \eta ln(Q_t)$ is shown in 5.8. It is reasonable to assume $ln(P_t)$ and $ln(Q_t)$ are cointegrated.

Thus we estimate the parameters η, μ, λ and σ by maximum likelihood method based on equation 45 and also by the method of least squares. A comparison of the outputs is presented in table 5.9. Although both methods do not perform well in this case, the maximum likelihood estimates are better than the method of least squares estimates. A

visualization of this is also shown in figure 5.17.

To better understand the performance of the model, we perform Monte Carlo simulations with 10,000 sample paths and obtain the average estimates for the parameters, using both maximum likelihood and least squares methods. The results are shown in table 5.10. We again reproduce the paths based on these averages, and they are plotted in figure 5.18. We observe here that the estimates do not perform very well. However, the maximum likelihood estimate is closer to the true path than the path from the least squares estimates.

5.1.6. Optimal Threshold

In this subsection, we obtain optimal thresholds for the process discussed in subsection 5.1.5. We will let Δt correspond to a day. Thus the 126 trading days correspond to half of a year of trading days and 1260 corresponds to five years of trade data in the US financial market.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.26151180	-0.03848820	0.05445690	0.00444670
		LS	0.26272021	-0.03727979		
μ	2.50000000	MLE	1.75560820	0.25560820	0.41761910	0.23974127
		LS	2.66869415	1.16869142		
λ	0.09000000	MLE	0.04647160	-0.00352840	0.02996400	0.00091029
		LS	0.05683273	0.00683273		
σ	0.01300000	MLE	0.01272400	-0.00027600	0.00082630	0.00000076
		LS	0.01282800	-0.00017200		

TABLE 5.9. Parameter estimates by one realization, for the OU representation of the spread between lnP_t and lnQ_t for 126 steps, by maximum likelihood and method of least squares.

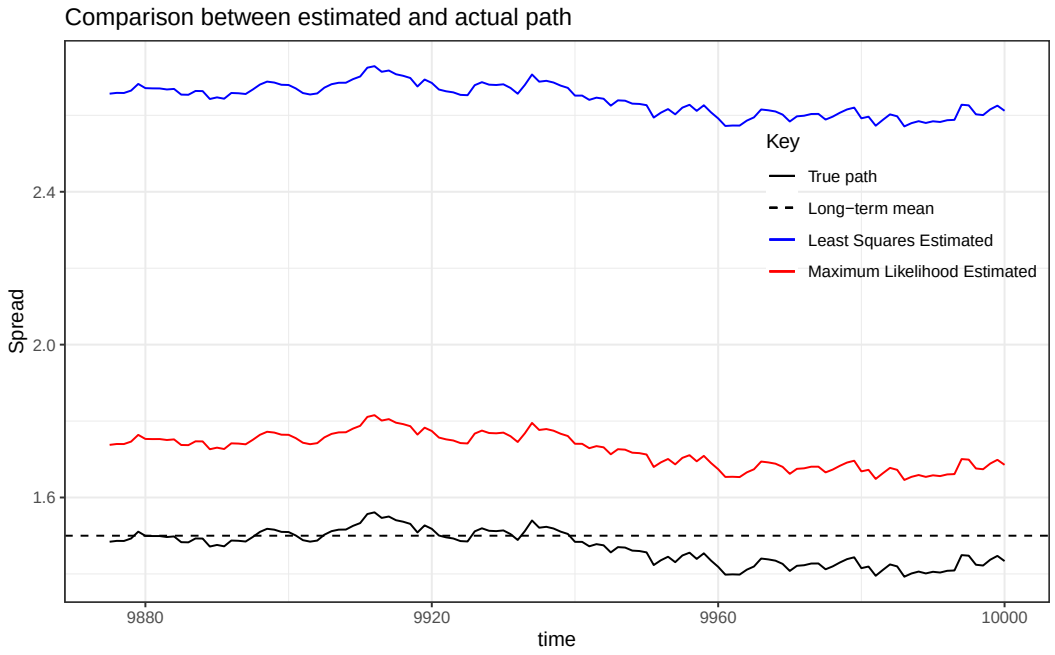

FIGURE 5.17. Spreads of length 126 each, generated from one realization of parameter estimates from maximum likelihood and least squares methods, using $ln(Q_t)$ and $ln(P_t)$

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.29927062	-0.00072938	0.05609920	0.00314765
		LS	0.29571629	-0.00428371	0.04870552	0.00239058
μ	1.50000000	MLE	1.54209289	0.04209289	0.51689144	0.26894857
		LS	1.43436118	-0.06563882	1.46113841	2.13923392
λ	0.05000000	MLE	0.08370286	0.03370286	0.05781343	0.00447828
		LS	0.10151337	0.05151337	0.04976264	0.00512995
σ	0.01300000	MLE	0.01303405	0.00003405	0.00088106	0.00000078
		LS	0.01321786	0.00021786	0.00092366	0.00000090

TABLE 5.10. Parameter estimates by Maximum Likelihood and Least Squares methods the 126 steps OU process, by Monte Carlo simulations

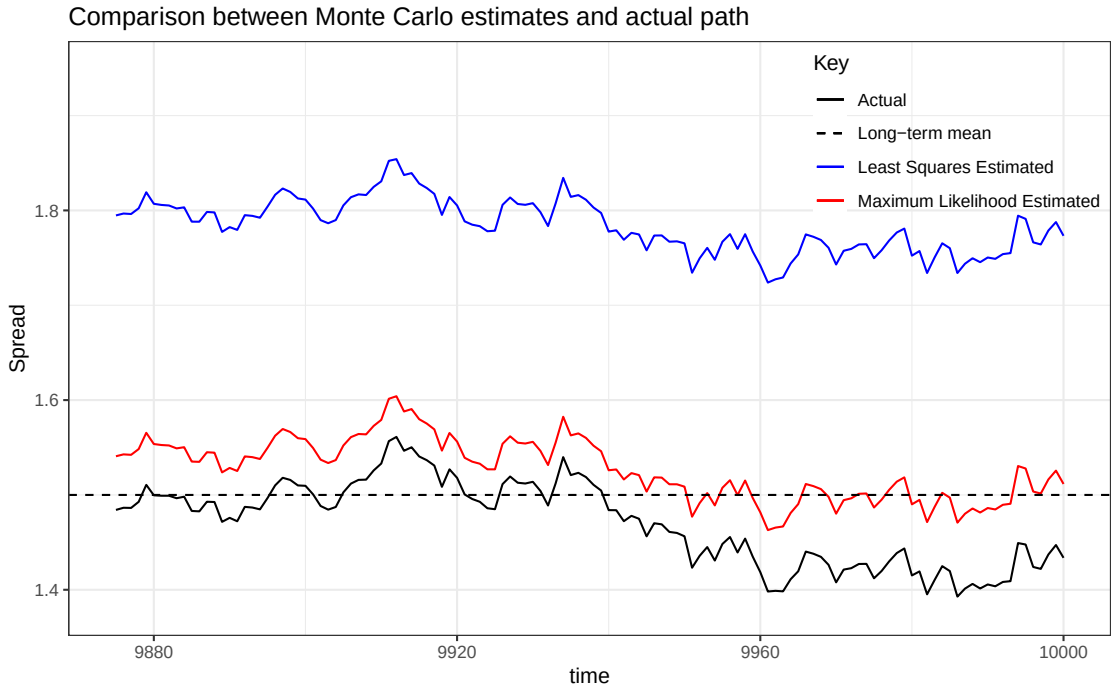

FIGURE 5.18. Spreads of length 126 each, generated from parameter estimates by maximum likelihood and least squares methods from $ln(Q_t)$ and $ln(P_t)$, based on Monte Carlo averages.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	1.50000000	MLE	1.50195030	0.00195030	0.00633370	0.00004392
		LS	1.50197867	0.00197867		
λ	0.05000000	MLE	0.06021400	0.01021400	0.01006890	0.00020571
		LS	0.06025913	0.01025913		
σ	0.01300000	MLE	0.01352990	0.00052990	0.00027780	0.00000039
		LS	0.01353488	0.00053488		

TABLE 5.11. One realization of parameter estimates by maximum likelihood and least squares methods, for the OU representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

Ideally, it is best to estimate all the four parameters η, μ, λ and σ together. But, con-

sidering how the one realization estimates perform, particularly in estimating the long-term mean μ when all four parameters are estimated together, and the standard error obtained for μ, we recommend estimating η separately by simple linear regression of P_t on Q_t and then estimate the remaining three parameters using maximum likelihood estimates as shown in subsections 5.1.3 and 5.1.4. This is in line with the methods used by [40] and [17].

5.1.6.1. Long Path

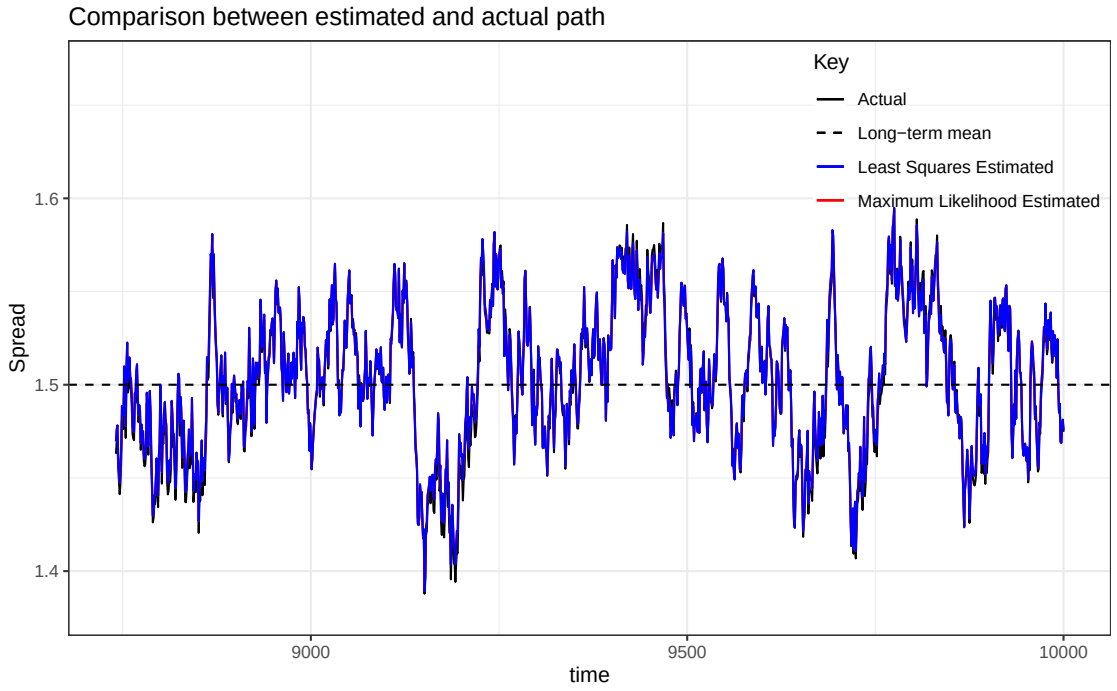

FIGURE 5.19. Paths generated from one realization of parameter estimates by maximum likelihood and least squares methods, for the OU representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

5.1.6.2. Case 1: $\gamma = 0$

The thresholds for this case are given by

$$g(t) = \mu \pm \frac{\sigma}{\sqrt{2\lambda}}\beta e^{-\lambda t}. \tag{50}$$

Substituting the time horizon $T = 5$ into the corresponding $h(\beta)$ in chapter 4 and solving gives the optimizer as $\tilde{\beta} \approx 0.9715$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{\mu} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(0.9715)e^{-\tilde{\lambda}t}.$$

From maximum likelihood estimates in table 5.11, we take the values of the parameters as $\tilde{\mu} = 1.5020$, $\tilde{\lambda} = 0.0602$ and $\tilde{\sigma} = 0.0135$. Hence we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &= 1.5020 \pm \frac{0.0135}{\sqrt{2(0.0602)}}(0.9715)e^{-0.0602t} \\ &= 1.5020 \pm 0.0378e^{-0.0602t} \end{aligned} \tag{51}$$

A graphical representation of this result is shown in figure 5.20.

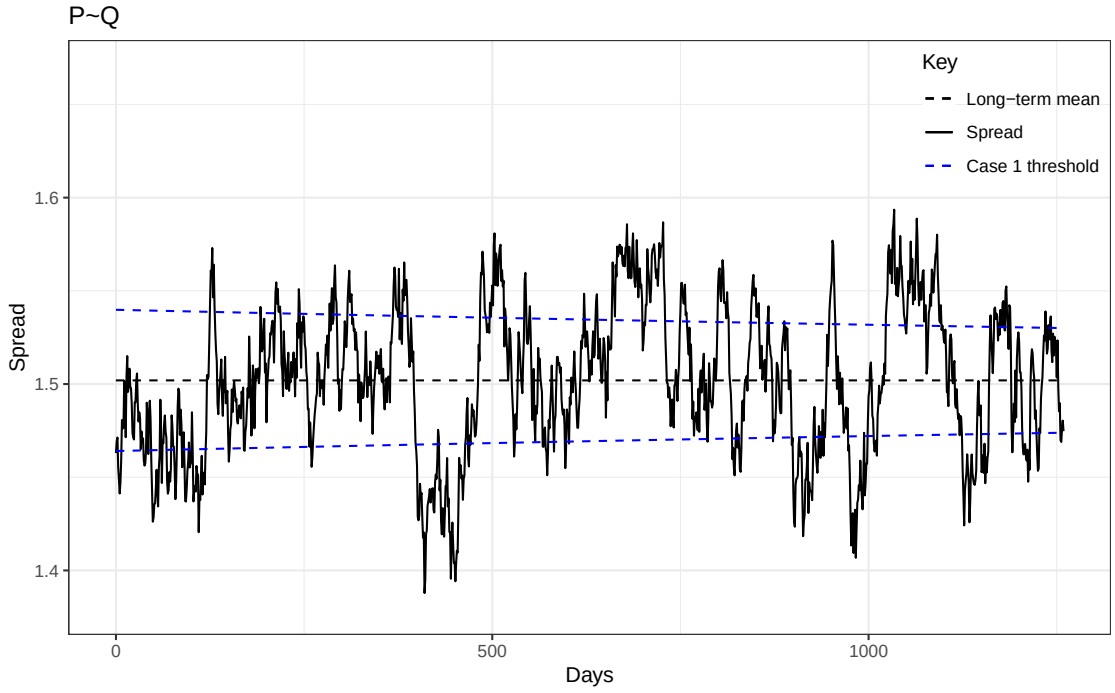

FIGURE 5.20. New threshold case 1 on the OU representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

5.1.6.3. Case 2: $\gamma \geq 0$ and $\rho = 1$

For this, the threshold is given by:

$$g(t) = \mu \pm \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\lambda t}). \tag{52}$$

Substituting the time horizon $T = 5$ into the corresponding $h(\beta, \gamma)$ in chapter 4 and solving gives the optimizer as $(\tilde{\beta}, \tilde{\gamma}) \approx (1.3351, 0.1856)$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{\mu} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(1.3351e^{-\tilde{\lambda}t} - 0.1856e^{\tilde{\lambda}t}).$$

Taking the estimates of the OU process parameters to be $\tilde{\mu} = 1.5020$, $\tilde{\lambda} = 0.0602$ and $\tilde{\sigma} = 0.0135$, as in case 1, we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &\approx 1.5020 \pm \frac{0.0135}{\sqrt{2(0.0602)}}(1.3351e^{-0.0602t} - 0.1856e^{0.0602t}) \\ &= 1.5020 \pm 0.0389(1.3351e^{-0.0602t} - 0.1856e^{0.0602t}) \end{aligned} \tag{53}$$

A graphical representation of this result is shown in figure 5.21.

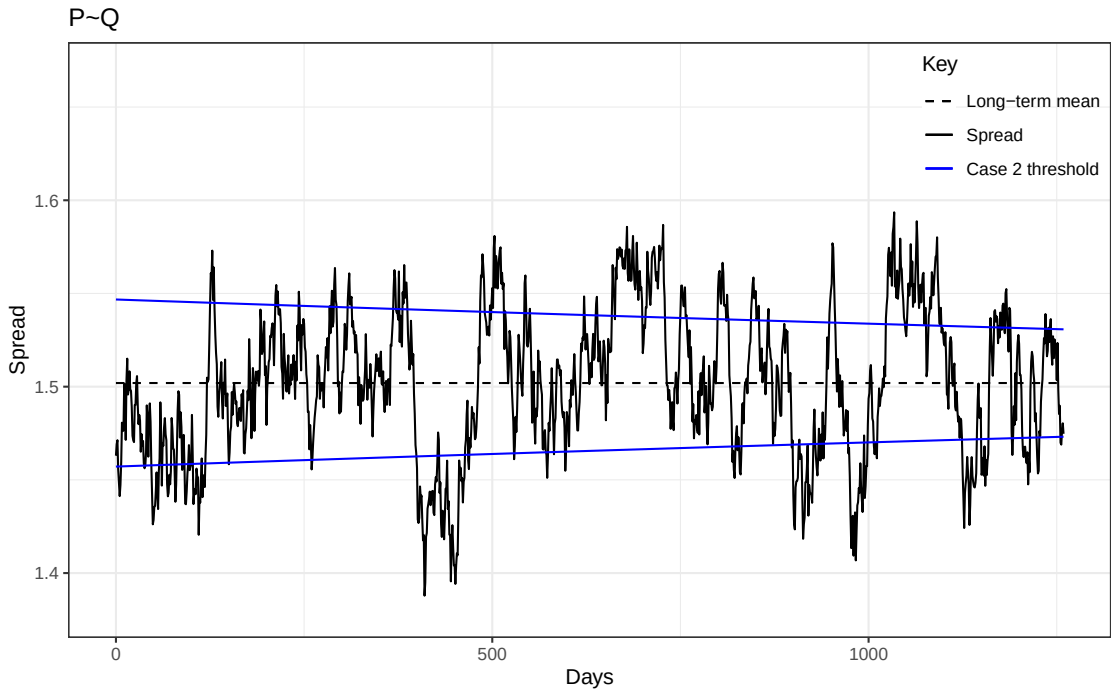

FIGURE 5.21. New threshold case 2 on the OU representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

In contrast, the optimal thresholds provided by case 1 has a narrower band in comparison with that of case 2 and as such is crossed more often by the path and hence generates more trades, while the optimal threshold presented in case 2 has a broader band, and as a result may have fewer number of crossings than case 1 and therefore fewer number of trades, but due to the wide band, the return per trade cycle is greater for case 2 than case one.

5.1.6.4. Short Path

Let us now look at the situation for short trade time horizons. Specifically, we consider 126 days or half a year. From table 5.12 , we see that apart from λ, the other two parameters were quite well estimated. Nonetheless, a look at the graph in figure 5.22 shows that the true path is quite well approximated by our estimated path.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
μ	1.50000000	MLE	1.50027460	0.00027460	0.01057400	0.00011189
		LS	1.50028069	0.00028069		
λ	0.05000000	MLE	0.12241420	0.07241420	0.04302880	0.00709529
		LS	0.12250859	0.07250859		
σ	0.01300000	MLE	0.01438200	0.00138200	0.00095680	0.00000283
		LS	0.01444111	0.00144111		

TABLE 5.12. One realization of parameter estimates by maximum likelihood and least squares methods, for the OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

5.1.6.5. Case 1: $\gamma = 0$

Substituting the time horizon $T = \frac{1}{2}$ into the corresponding $h(\beta)$ in chapter 4 and solving gives the optimizer of the expected return as $\tilde{\beta} \approx 0.8735$, approximated to four decimal places.

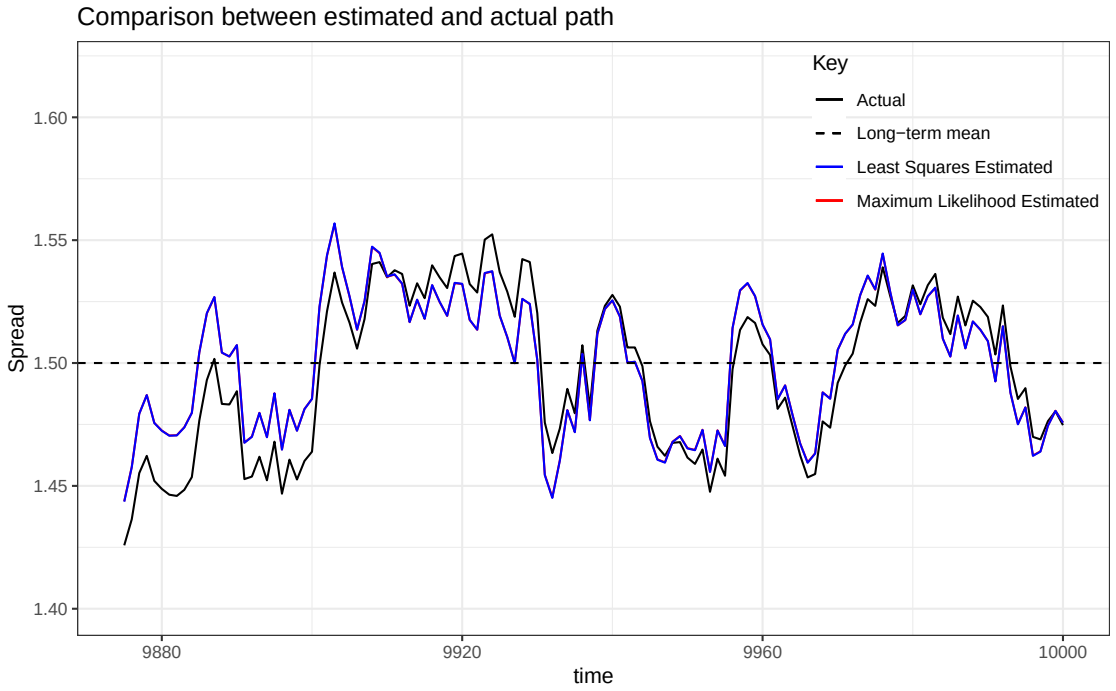

FIGURE 5.22. Paths generated from one realization of parameter estimates by maximum likelihood and least squares methods, for the OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$ based on equation ...

Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{\mu} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(0.8735)e^{-\tilde{\lambda}t}.$$

From the maximum likelihood estimates in table 5.12, we take the values of the parameters as $\tilde{\mu} = 1.5003$, $\tilde{\lambda} = 0.1224$ and $\tilde{\sigma} = 0.0144$. Hence we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &= 1.5003 \pm \frac{0.0144}{\sqrt{2(0.1224)}}(0.8735)e^{-0.1224t} \\ &= 1.5003 \pm 0.0254e^{-0.1224t} \end{aligned} \tag{54}$$

A graphical representation of this result is shown in figure 5.23.

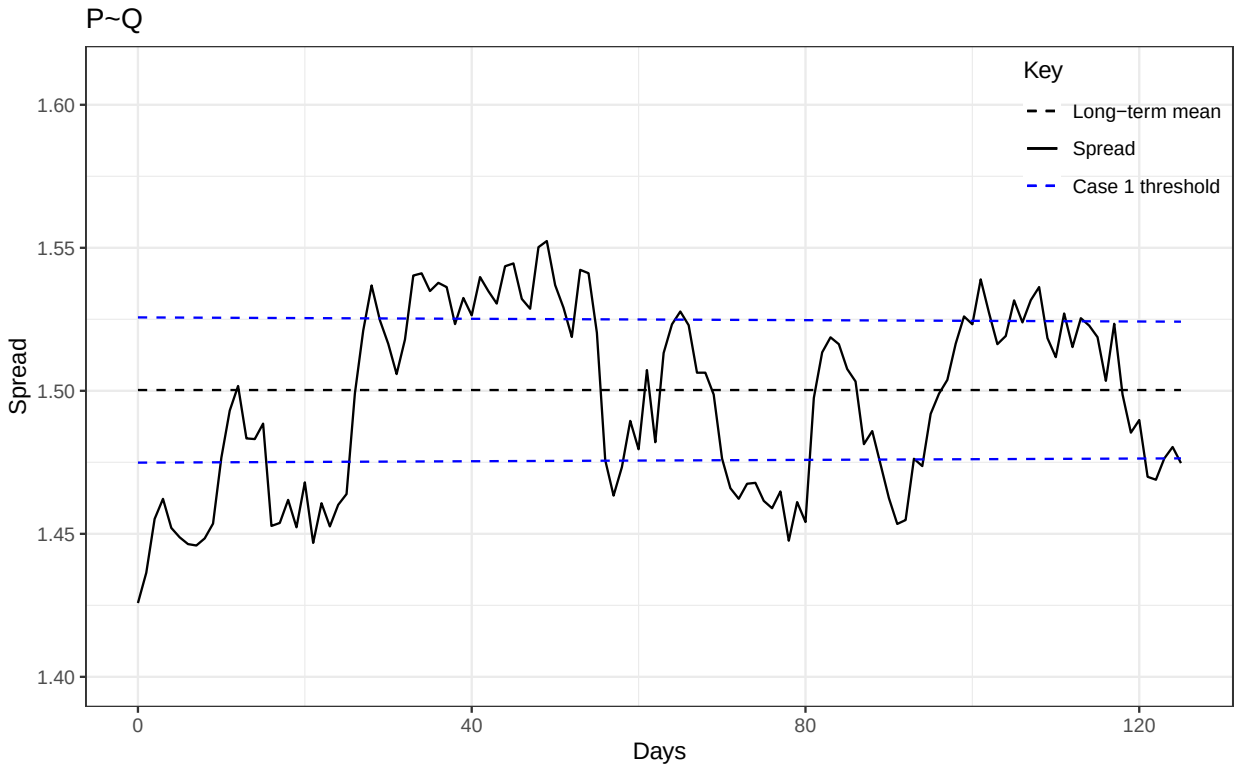

FIGURE 5.23. New threshold case 1 on the OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

5.1.6.6. Case 2: $\gamma \geq 0$ and $\rho = 1$

Substituting the time horizon $T = \frac{1}{2}$ into the corresponding $h(\beta, \gamma)$ in chapter 4 and solving gives the optimizer of the expected return as $(\tilde{\beta}, \tilde{\gamma}) \approx (2.2621, 0.6944)$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{\mu} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(2.2621e^{-\tilde{\lambda}t} - 0.6944e^{\tilde{\lambda}t}).$$

Taking the estimates of the OU process parameters to be $\tilde{\mu} = 1.5003$, $\tilde{\lambda} = 0.1224$ and $\tilde{\sigma} = 0.0144$, as in case 1, we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &\approx 1.5003 \pm \frac{0.0144}{\sqrt{2(0.1224)}}(2.2621e^{-0.1224t} - 0.6944e^{0.1224t}) \\ &= 1.5003 \pm 0.0291(2.2621e^{-0.1224t} - 0.6944e^{0.1224t}) \end{aligned} \tag{55}$$

A graphical representation of this result is shown in figure 5.24.

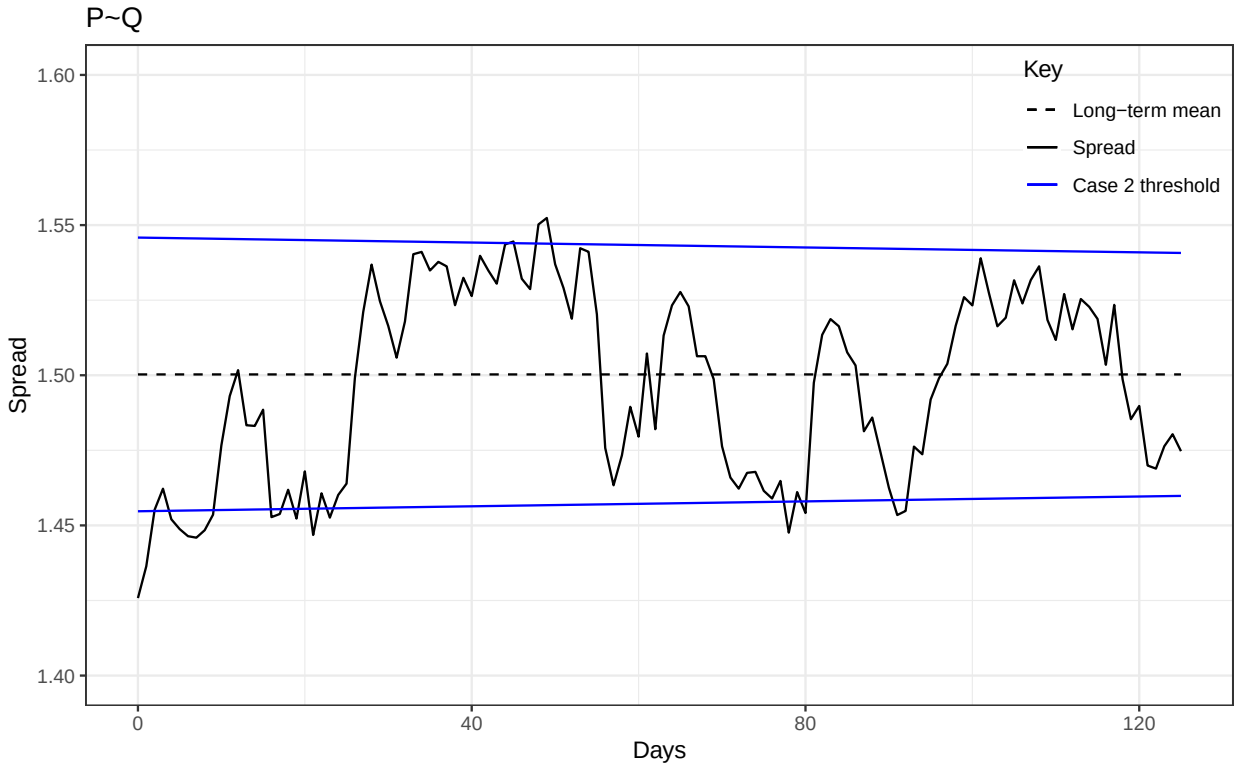

FIGURE 5.24. New threshold case 2 on the OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

Again, the optimal thresholds provided by case 1 has a narrower band and as such is crossed more often by the path, while the optimal threshold presented in case 2 has a broader band, and as a result may have fewer number of crossings than case 1, but may generate higher return per trade cycle.

5.2. Nonzero Trend Generalized Ornstein-Uhlenbeck Process

5.2.1. The Process

Let us now consider the generalized version of the Ornstein-Uhlenbeck process:

$$d(X_t - (at+b)) = -\lambda(X_t - (at+b))dt + \sigma dB_t, \quad t \geq 0 \tag{56}$$

$$X_0 = x_0, \quad \lambda > 0, \quad \sigma > 0$$

As noted in chapter four, this process is trend-stationary, in that it pulls back to the linear trend $at + b$ whenever it deviates from it. We apply Ito's lemma A.3 once again to solve the system.

First, we rewrite the stochastic differential equation as

$$dX_t = (a - \lambda(X_t - (at + b)))dt + \sigma dB$$

Let $Y_t = e^{\lambda t} X_t$

By Ito's lemma A.3,

$$\begin{aligned} dY_t &= (\lambda e^{\lambda t} X_t + e^{\lambda t}(a - \lambda(X_t - (at + b))))dt + \sigma e^{\lambda t} dB_t \\ &= (a + \lambda at + \lambda b)e^{\lambda t} dt + \sigma e^{\lambda t} dB_t \end{aligned} \tag{57}$$

Integrating both sides on the interval $0 \leq t \leq s$, we get,

$$\begin{aligned} \int_0^s dY_t &= \int_0^s (a + \lambda at + \lambda b)e^{\lambda t} dt + \sigma \int_0^s e^{\lambda t} dB_t \\ Y_s - Y_0 &= \int_0^s (a + \lambda at + \lambda b)e^{\lambda t} dt + \sigma \int_0^s e^{\lambda t} dB_t \end{aligned}$$

Thus,

$$e^{\lambda t} X_s - X_0 = \int_0^s (a + \lambda at + \lambda b)e^{\lambda t} dt + \sigma \int_0^s e^{\lambda t} dB_t$$

Applying integration by parts to the first term on the right hand side gives,

$$\begin{aligned} e^{\lambda s} X_s &= X_0 + \left[\frac{(a + \lambda at + \lambda b)e^{\lambda t}}{\lambda}\right]_0^s - \int_0^s a e^{\lambda t} dt + \sigma \int_0^s e^{\lambda t} dB_t \\ &= X_0 + \frac{(a + \lambda as + \lambda b)e^{\lambda s}}{\lambda} - \frac{(a + \lambda b)}{\lambda} - \left[\frac{1}{\lambda} a e^{\lambda t}\right]_0^s + \sigma \int_0^s e^{\lambda t} dB_t \\ &= X_0 + \frac{(a + \lambda as + \lambda b)e^{\lambda s}}{\lambda} - \frac{(a + \lambda b)}{\lambda} - \frac{a e^{\lambda t}}{\lambda} + \frac{a}{\lambda} + \sigma \int_0^s e^{\lambda t} dB_t \\ &= X_0 + (as + b)e^{\lambda s} - b + \sigma \int_0^s e^{\lambda t} dB_t \end{aligned}$$

Dividing both sides by $e^{\lambda s}$, we get

$$X_s = e^{-\lambda s}X_0 + as + b - be^{-\lambda s} + \sigma \int_0^s e^{-\lambda(s-t)}dB_t$$

Similarly, if we integrate 57 over the time interval $t-1 \leq \nu \leq t$, where $\Delta t = t-(t-1)$ is a unit change in time, we obtain:

$$\begin{aligned}
\int_{t-1}^{t} dY_\nu &= \int_{t-1}^{t} (a + \lambda a\nu + \lambda b)e^{\lambda\nu}d\nu + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu \\
Y_t - Y_{t-1} &= \int_{t-1}^{t} (a + \lambda a\nu + \lambda b)e^{\lambda\nu}d\nu + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu \\
e^{\lambda t}X_t - e^{\lambda(t-1)}X_{t-1} &= \int_{t-1}^{t} (a + \lambda a\nu + \lambda b)e^{\lambda\nu}d\nu + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu
\end{aligned}$$

Integrating the first term on the right hand side by parts gives

$$\begin{aligned}
e^{\lambda t}X_t &= e^{\lambda(t-1)}X_{t-1} + \left[\frac{(a + \lambda a\nu + \lambda b)e^{\lambda\nu}}{\lambda}\right]_{t-1}^{t} - \int_{t-1}^{t} ae^{\lambda\nu}d\nu + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu \\
&= e^{\lambda(t-1)}X_{t-1} + \frac{(a + \lambda at + \lambda b)e^{\lambda t}}{\lambda} - \frac{(a + \lambda a(t-1) + \lambda b)e^{\lambda(t-1)}}{\lambda} \\
&\quad - \left[\frac{1}{\lambda}ae^{\lambda\nu}\right]_{t-1}^{t} + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu \\
&= e^{\lambda(t-1)}X_{t-1} + \frac{(a + \lambda at + \lambda b)e^{\lambda t}}{\lambda} - \frac{(a + \lambda a(t-1) + \lambda b)e^{\lambda(t-1)}}{\lambda} \\
&\quad - \frac{ae^{\lambda t}}{\lambda} + \frac{ae^{\lambda(t-1)}}{\lambda} + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu
\end{aligned}$$

Dividing both sides by $e^{\lambda t}$, we get

$$\begin{aligned}
X_t &= e^{-\lambda(t-(t-1))}X_{t-1} + \frac{a}{\lambda} + at + b - \frac{a}{\lambda}e^{-\lambda(t-(t-1))} - a(t-1)e^{-\lambda(t-(t-1))} \\
&\quad - be^{-\lambda(t-(t-1))} - \frac{a}{\lambda} + \frac{a}{\lambda}e^{-\lambda(t-(t-1))} + \sigma \int_{t-1}^{t} e^{\lambda\nu}dB_\nu \\
&= e^{-\lambda\Delta t}X_{t-1} + at + b - a(t-1)e^{-\lambda\Delta t} - be^{-\lambda\Delta t} + \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu \\
&= e^{-\lambda\Delta t}X_{t-1} + b(1 - e^{-\lambda\Delta t}) + a(t - (t-1)e^{-\lambda\Delta t}) + \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu \\
&= e^{-\lambda\Delta t}X_{t-1} + b(1 - e^{-\lambda\Delta t}) + a(t - (t-\Delta t)e^{-\lambda\Delta t}) + \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu
\end{aligned}$$

$$= e^{-\lambda\Delta t}X_{t-1} + b(1 - e^{-\lambda\Delta t}) + a(1 - e^{-\lambda\Delta t})t + a\Delta t e^{-\lambda\Delta t}$$
$$+ \sigma \int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu$$

(58)

As discussed earlier, the Ito integral $\int_{t-1}^{t} e^{-\lambda(t-\nu)} dB_\nu$ in equation 58 above follows a normal distribution with mean zero and variance $\frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t})$.

So, it follows that:

(59)

$$X_t | X_{t-1} \sim N\left(e^{-\lambda\Delta t}X_{t-1} + b(1 - e^{-\lambda\Delta t}) + a(1 - e^{-\lambda\Delta t})t + a\Delta t e^{-\lambda\Delta t}, \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t})\right).$$

As before, a discretized form of the trending OU process X_t can be obtained from the distribution in 59 above.

Let us consider an example of this process with parameter values $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$, which starts at $x_0 = 0.4$. We generate a path of length 10,000 based on he model. This is shown in figure 5.25. We see here that the process oscillates around the line $X_t = 0.0002t + 0.02$, and always reverts to it whenever there is a deviation.

5.2.2. Parameter Estimation

We show in this subsection parameter estimation for the trend-stationary OU process both with maximum likelihood and least squares methods.

5.2.2.1. Maximum Likelihood Estimator

From relation 59, the likelihood function for this model is:

$$L(a, b, \lambda, \sigma | X_t) = \prod_{t=1}^{N} \left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda\Delta t})}} e^{-\frac{(X_t - (e^{-\lambda\Delta t}X_{t-1} + b(1-e^{-\lambda\Delta t}) + a(1-e^{-\lambda\Delta t})t + a\Delta t e^{-\lambda\Delta t}))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}} \right)$$

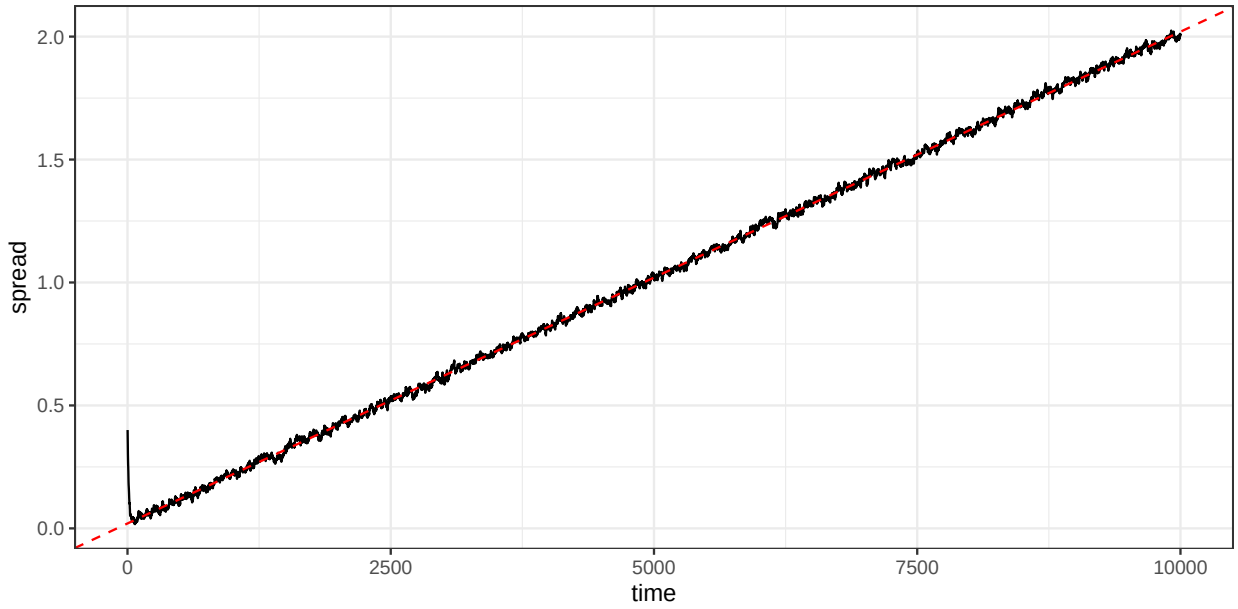

FIGURE 5.25. Discretized trend-stationary Ornstein-Uhlenbeck Process of length 10000, with parameters $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$, starting from $x_0 = 0.4$

We then obtain the log-likelihood as:

$$l(a,b,\lambda,\sigma|X_t) = log\prod_{t=1}^{N}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})}}e^{-\frac{(X_t-(e^{-\lambda\Delta t}X_{t-1}+b(1-e^{-\lambda\Delta t})+a(1-e^{-\lambda\Delta t})t+a\Delta te^{-\lambda\Delta t}))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}}\right)$$

$$= -\frac{N}{2}log(2\pi) - \frac{N}{2}log(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))$$

$$-\frac{\lambda}{\sigma^2(1-e^{-2\lambda\Delta t})}\sum_{t=1}^{N}(X_t - e^{-\lambda\Delta t}X_{t-1} - b(1-e^{-\lambda\Delta t}) - a(1-e^{-\lambda\Delta t})t - a\Delta te^{-\lambda\Delta t}))^2 \quad (60)$$

Thus the maximum likelihood estimates for a, b, λ and σ will be the values $\hat{a}$, $\hat{b}$, $\hat{\lambda}$ and $\hat{\sigma}$ respectively, that maximize the above expression. We will solve this numerically from equation 60.

5.2.2.2. Method of Least Squares

We formulate the least squares regression equation from equation 59 as follows:

$$X_t = \zeta + \alpha t + \psi X_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.N(0, \sigma_\epsilon) \tag{61}$$

where

$$\zeta = b(1 - e^{-\lambda \Delta t}) + a\Delta t e^{-\lambda \Delta t} \tag{62}$$

$$\alpha = a(1 - e^{-\lambda \Delta t}) \tag{63}$$

$$\psi = e^{-\lambda \Delta t} \tag{64}$$

$$\sigma_\epsilon^2 = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda \Delta t}) \tag{65}$$

Substituting the least squares estimates of equation 61, $\hat{\zeta}$, $\hat{\psi}$ and $\hat{\sigma_\epsilon}$, into equations 62, 63, 64 and 65, and solving the system gives the least squares parameter estimates for the OU model as:

$$\hat{\lambda} = -\frac{log\hat{\psi}}{\Delta t}$$

$$\hat{a} = \frac{\hat{\alpha}}{1 - \hat{\psi}}$$

$$\hat{b} = \frac{\hat{\zeta}(1 - \hat{\psi}) - \Delta \alpha \psi}{(1 - \hat{\psi})^2}$$

$$\hat{\sigma} = \frac{\sqrt{2\hat{\lambda}}\hat{\sigma_\epsilon}}{\sqrt{1 - \hat{\psi}^2}}$$

To see the performance of our parameter estimation methods, we will apply them to both a long path and a short one as before.

5.2.3. Long Path

5.2.3.1. Generating the Spread

We use the last 1260 points of the path presented in figure 5.25 in our study. A graph of this path can be found in figure 5.26.

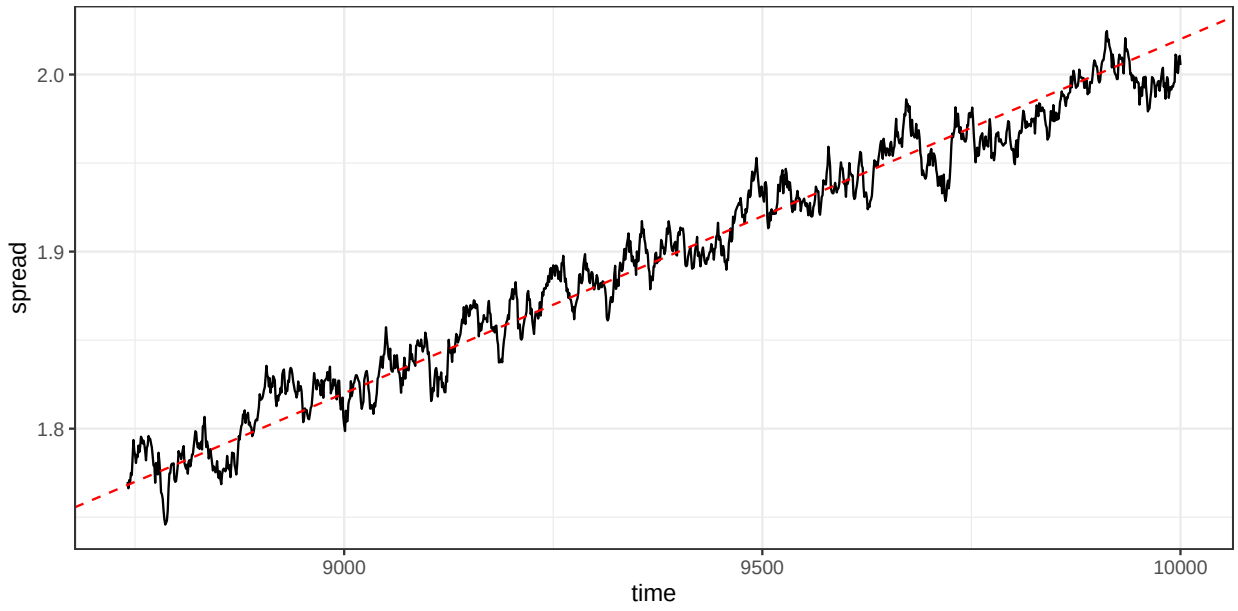

FIGURE 5.26. Discretized trend-stationary Ornstein-Uhlenbeck Process of length 1260, with parameter values $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

5.2.3.2. Parameter Estimation

We now estimate the parameters of the model by maximum likelihood and least squares methods. The result is shown in table 5.13. We notice that both methods estimate the parameters a, λ and σ quite well, but do not perform well for b. We generate paths from the estimates and compare them against the true path in figure 5.27. From this figure we can see that the model is quite resistant to the error in estimating b.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00018800	-0.00001200	0.00000017	0.00000000
		LS	0.00018798	-0.00001202		
b	0.02000000	MLE	0.13315400	0.11315400	0.00159200	0.01280636
		LS	0.13315080	0.11315080		
λ	0.08000000	MLE	0.08999000	0.00999000	0.01252300	0.00025663
		LS	0.09004500	0.01004500		
σ	0.00500000	MLE	0.00508300	0.00008300	0.00010200	0.00000002
		LS	0.00508510	0.00008510		

TABLE 5.13. Parameter estimation for the discretized trend-stationary OU process of length 1260 by maximum likelihood and method of least squares, with true parameter values $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

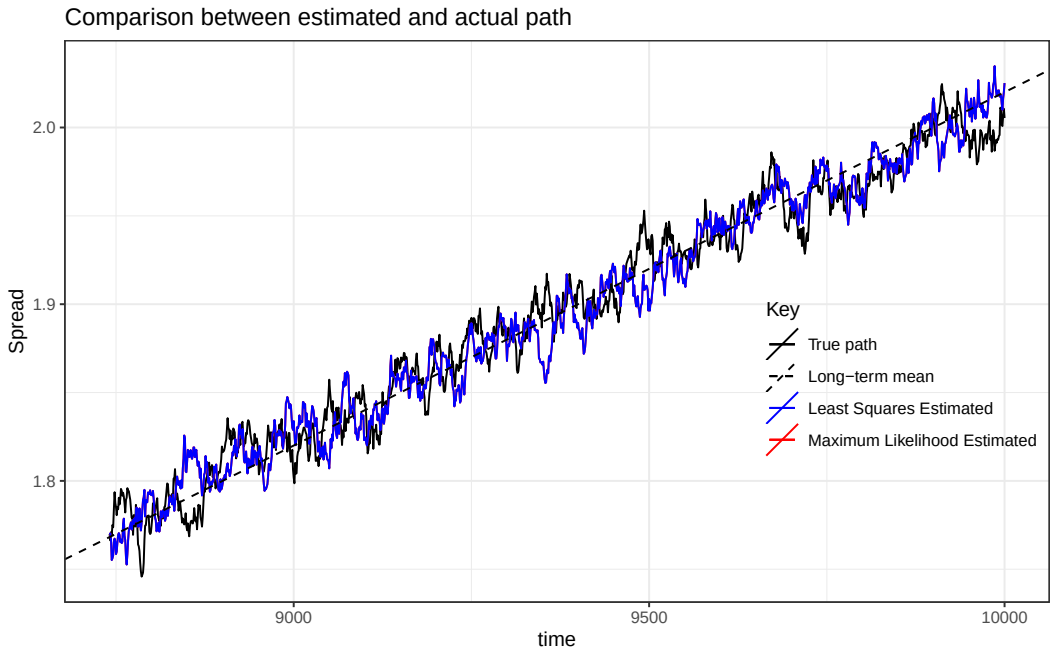

FIGURE 5.27. Paths of length 1260, generated from one realization of parameter estimates from maximum likelihood and least squares methods, with true parameter values $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$, for the trend-stationary OU process.

5.2.3.3. Monte Carlo Simulation

We also perform a Monte Carlo simulation with on the same parameter values as above to see the average estimate of the methods. From the results in table 5.14, we see that on average all the parameters are well estimated and there is significant improve in the estimates for b in comparison with the case of one realization. A plot of the paths generated from these estimates in comparison with the true path can be found in figure 5.28.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00019993	-0.00000007	0.00000290	0.00000000
		LS	0.00020000	-0.00000000	0.00000478	0.00000000
b	0.02000000	MLE	0.02083869	0.00083869	0.02720594	0.00074087
		LS	0.02017871	0.00017871	0.04493147	0.00201887
λ	0.08000000	MLE	0.08344855	0.00344855	0.01285059	0.00017703
		LS	0.08481821	0.00481821	0.01221014	0.00017230
σ	0.00500000	MLE	0.00500934	0.00000934	0.00010678	0.00000001
		LS	0.00500355	0.00000355	0.00010317	0.00000001

TABLE 5.14. Performance of Maximum Likelihood and Least Squares methods for estimating the trend-stationary OU parameters, by 10,000 Monte Carlo simulations, using a path of length 1260

5.2.4. Short Path

Let us now consider short term trade time horizon.

5.2.4.1. Generating the Spread

We will use the last 126 steps of the path presented in figure 5.25. A graph of this path can be found in figure 5.29.

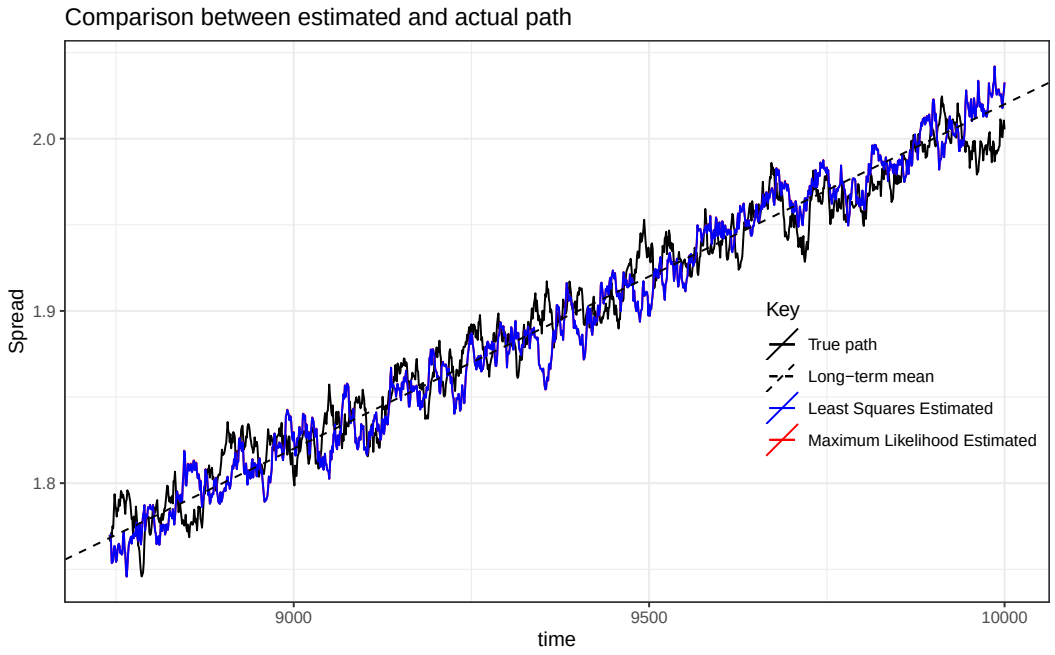

FIGURE 5.28. Paths generated from parameter estimates by Maximum Likelihood and Method of Least Squares, using Monte Carlo averages with 10,000 replications of paths of length 1260. True parameter values are $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

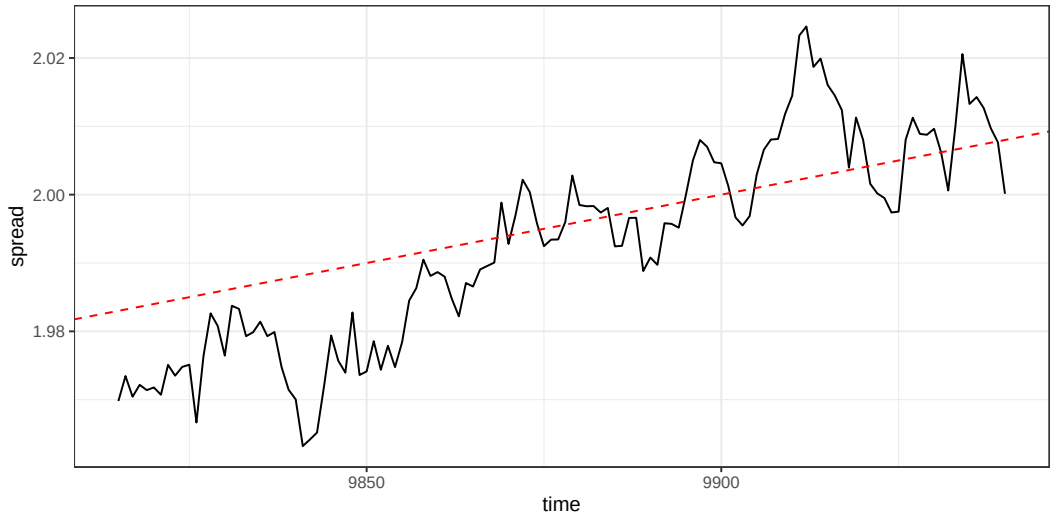

FIGURE 5.29. Discretized trend-stationary Ornstein-Uhlenbeck Process of length 126, with parameter values $a = 0.0002$, $b = 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

5.2.4.2. Parameter Estimation

In estimating the parameters from one realization, we notice from table 5.15 that a and σ are quite well estimated, while the estimates for b and λ are not very good, both for the maximum likelihood and least squares methods. However, the paths generated from these estimates approximate the true path well enough for our purpose. See figure 5.30.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00034270	0.00014270	0.00000020	0.00000002
		LS	0.00034272	0.00014272		
b	0.02000000	MLE	-1.39274270	-1.59274270	0.00187570	2.53683283
		LS	-1.39275034	-1.59275034		
λ	0.08000000	MLE	0.21820710	0.13820710	0.06940170	0.02391780
		LS	0.21808244	0.13808244		
σ	0.00500000	MLE	0.00455410	-0.00044590	0.00030570	0.00000029
		LS	0.00457163	-0.00042837		

TABLE 5.15. Parameter estimation for the discretized trend-stationary OU process of length 126 by maximum likelihood and method of least squares, with actual parameters $a = 0.0002$, $= 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

5.2.4.3. Monte Carlo Simulation

The average estimates from Monte Carlo simulations with 10,000 replications is shown if table 5.16. We notice an improvement over the one realization result for all the parameters, for both estimation methods. More importantly, we notice that the maximum likelihood estimates outperform the least squares estimates and provides better estimates for all the parameters, which is in consonance with known statistical results. A plot of the paths generated from these estimates is shown in figure 5.31.

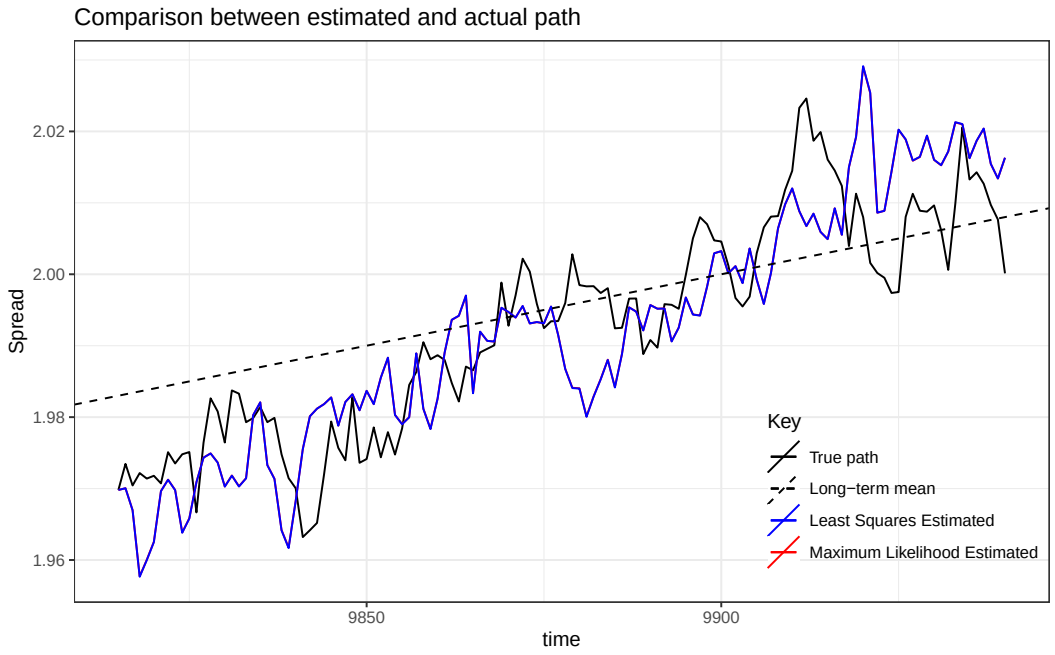

FIGURE 5.30. Paths of length 126 for the trend-stationary OU process, generated from one realization of estimates from maximum likelihood and least squares methods, with true parameter values $a = 0.0002$, $= 0.02$, $\lambda = 0.08$ and $\sigma = 0.005$

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00020047	0.00000047	0.00001548	0.00000000
		LS	0.00020751	0.00000751	0.00015013	0.00000002
b	0.02000000	MLE	0.01572380	-0.00427620	0.15276549	0.02335558
		LS	-0.05404180	-0.07404180	1.48428565	2.20858609
λ	0.08000000	MLE	0.11847379	0.03847379	0.05471677	0.00447416
		LS	0.14065392	0.06065392	0.05857387	0.00710980
σ	0.00500000	MLE	0.00505847	0.00005847	0.00033740	0.00000012
		LS	0.00508214	0.00008214	0.00034027	0.00000012

TABLE 5.16. Performance of Maximum Likelihood and Least Squares methods for estimating the trend-stationary OU parameters, by 10,000 Monte Carlo simulations, using a path of length 126

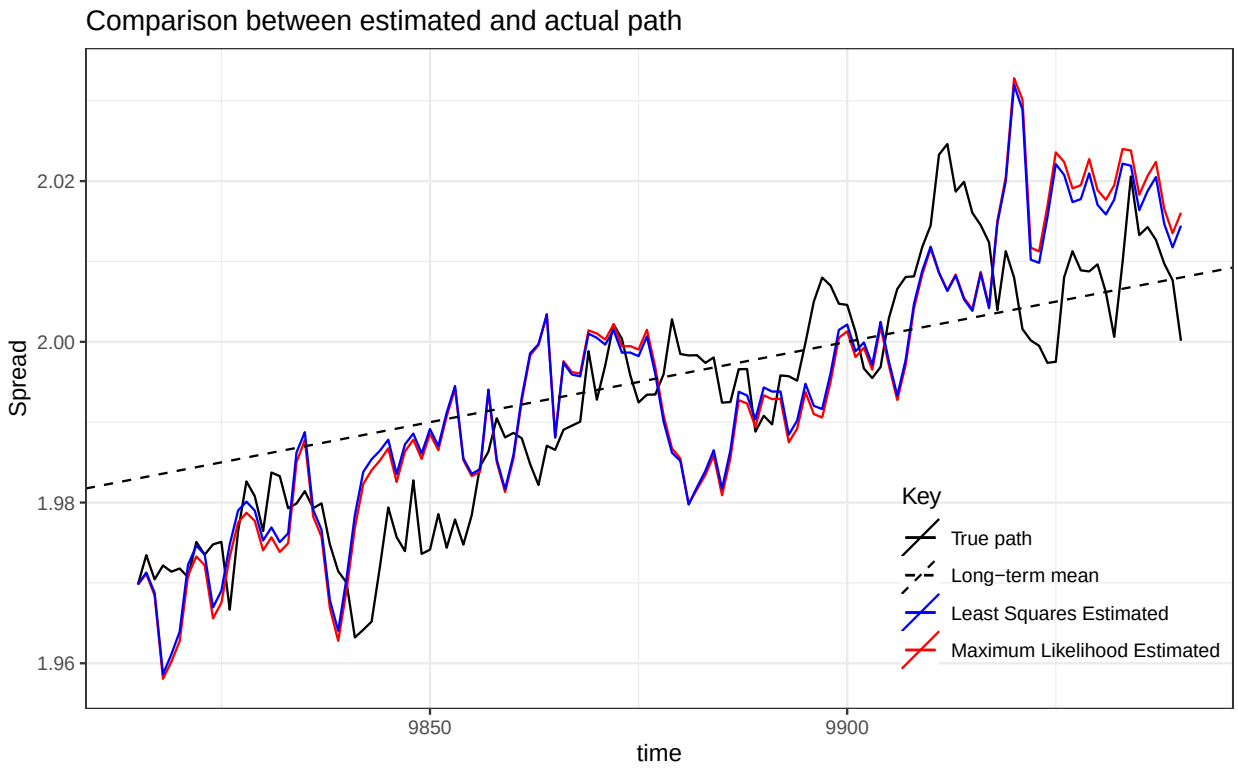

FIGURE 5.31. Trend-stationary OU paths from parameter estimation by Maximum Likelihood and Method of Least Squares, using Monte Carlo averages with 10,000 replications of paths of length 126. The true parameter values are

5.2.5. Artificial Stocks

We now generate a pair of artificial stock prices Q_t and P_t, such that the spread between their log-returns is trend-stationary and follows the trend-stationary OU process 5.2.1, and then perform parameter estimation for the pair.

5.2.5.1. Generating the Pair

For Q_t we will use the same GBM we generated in subsection 5.1.5. We use relation 59 to generate a 10,000 steps trend-stationary OU process starting at $x_0 = 1.4$, with parameter values $a = 0.0002$, $b = 0.8$, $\lambda = 0.08$ and $\sigma = 0.005$. We then combine these two processes using equation 1.7, with $\eta = 0.3$, to obtain the logarithm of the price time series P_t of a second stock P. The graphs of these are shown in figures 5.32, 5.33 and 5.34 respectively.

We know from equation 58 that:

$$X_t = e^{-\lambda\Delta t}X_{t-1} + b(1-e^{-\lambda\Delta t}) + a(1-e^{-\lambda\Delta t})t + a\Delta t e^{-\lambda\Delta t} + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu,$$

and by equation 1.7, it follows that:

$$\begin{aligned} ln(P_t) - \eta ln(Q_t) &= e^{-\lambda\Delta t}(ln(P_{t-1}) - \eta ln(Q_{t-1})) + b(1-e^{-\lambda\Delta t}) \\ &\quad + a(1-e^{-\lambda\Delta t})t + a\Delta t e^{-\lambda\Delta t} + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu \\ ln(P_t) &= b(1-e^{-\lambda\Delta t}) + a\Delta t e^{-\lambda\Delta t} + a(1-e^{-\lambda\Delta t})t + e^{-\lambda\Delta t}ln(P_{t-1}) \\ &\quad + \eta ln(Q_t) - \eta e^{-\lambda\Delta t}ln(Q_{t-1}) + \sigma\int_{t-1}^{t} e^{-\lambda(t-\nu)}dB_\nu \end{aligned} \tag{66}$$

Thus,

$$\begin{aligned} ln(P_t)|ln(P_{t-1}), ln(Q_t), ln(Q_{t-1}) \sim N\,(\, & b(1-e^{-\lambda\Delta t}) + a\Delta t e^{-\lambda\Delta t} + a(1-e^{-\lambda\Delta t})t + e^{-\lambda\Delta t}ln(P_{t-1}) \\ & + \eta ln(Q_t) - \eta e^{-\lambda\Delta t}ln(Q_{t-1}), \frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})\,)\,, \end{aligned} \tag{67}$$

which gives the following log-likelihood:

$$\begin{aligned} & l(\eta, a, b, \lambda, \sigma | lnP_t, lnQ_t, lnP_{t-1}, lnQ_{t-1}) \\ & = log\prod_{t=1}^{N}\left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})}}e^{-\frac{(ln(P_t)-b(1-e^{-\lambda\Delta t})-a\Delta t e^{-\lambda\Delta t}-a(1-e^{-\lambda\Delta t})t-e^{-\lambda\Delta t}ln(P_{t-1})-\eta ln(Q_t)+\eta e^{-\lambda\Delta t}ln(Q_{t-1}))^2}{2(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t}))}}\right), \end{aligned}$$

$$\begin{aligned} & = -\frac{N}{2}log(2\pi) - \frac{N}{2}log(\frac{\sigma^2}{2\lambda}(1-e^{-2\lambda\Delta t})) \\ & - \frac{\lambda}{\sigma^2(1-e^{-2\lambda\Delta t})}\sum_{t=1}^{N}(ln(P_t) - b(1-e^{-\lambda\Delta t}) - a\Delta t e^{-\lambda\Delta t} - a(1-e^{-\lambda\Delta t})t - e^{-\lambda\Delta t}ln(P_{t-1}) \\ & \qquad - \eta ln(Q_t) + \eta e^{-\lambda\Delta t}ln(Q_{t-1}))^2 \end{aligned} \tag{68}$$

The maxLik package in R will be used to numerically solve for optimal values of the parameters η, μ, λ and σ from given data set.

For the least squares estimates, we formulate the regression equation from equation 66 as follows:

$$ln(P_t) = \zeta + \alpha t + \theta ln(P_{t-1} + \eta ln(Q_t) + \psi Q_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.N(0, \sigma_\epsilon) \tag{69}$$

where

$$\zeta = b(1 - e^{-\lambda \Delta t}) + a\Delta t e^{-\lambda \Delta t} \tag{70}$$

$$\alpha = a(1 - e^{-\lambda \Delta t}) \tag{71}$$

$$\theta = e^{-\lambda \Delta t} \tag{72}$$

$$\sigma_\epsilon^2 = \frac{\sigma^2}{2\lambda}(1 - e^{-2\lambda \Delta t}) \tag{73}$$

Substituting the least squares estimates of equation 69, $\hat{\zeta}$, $\hat{\alpha}$, $\hat{\theta}$, $\hat{\psi}$ and $\hat{\sigma}_\epsilon$, into equations 70, 71, 72 and 73, and solving the system gives the least squares parameter estimates for the OU model as:

$$\hat{\lambda} = -\frac{log\hat{\theta}}{\Delta t}$$
$$\hat{a} = \frac{\hat{\alpha}}{1 - \hat{\theta}}$$
$$\hat{b} = \frac{\hat{\zeta}(1 - \hat{\theta}) - \Delta t \hat{\alpha} \hat{\theta}}{(1 - \hat{\theta})^2}$$
$$\hat{\sigma} = \frac{\sqrt{2\hat{\lambda}}\hat{\sigma}_\epsilon}{\sqrt{1 - \hat{\theta}^2}}$$

We again apply these parameter estimation methods to both a long-term trade horizon and a short-term trade horizon. The path $ln(Q_t)$ of length 10,000 from the GBM is shown in figure 5.32. The trend-stationary OU process, X_t is shown in figure 5.33, and the derived artificial time series $ln(P_t)$ is shown in figure 5.34.

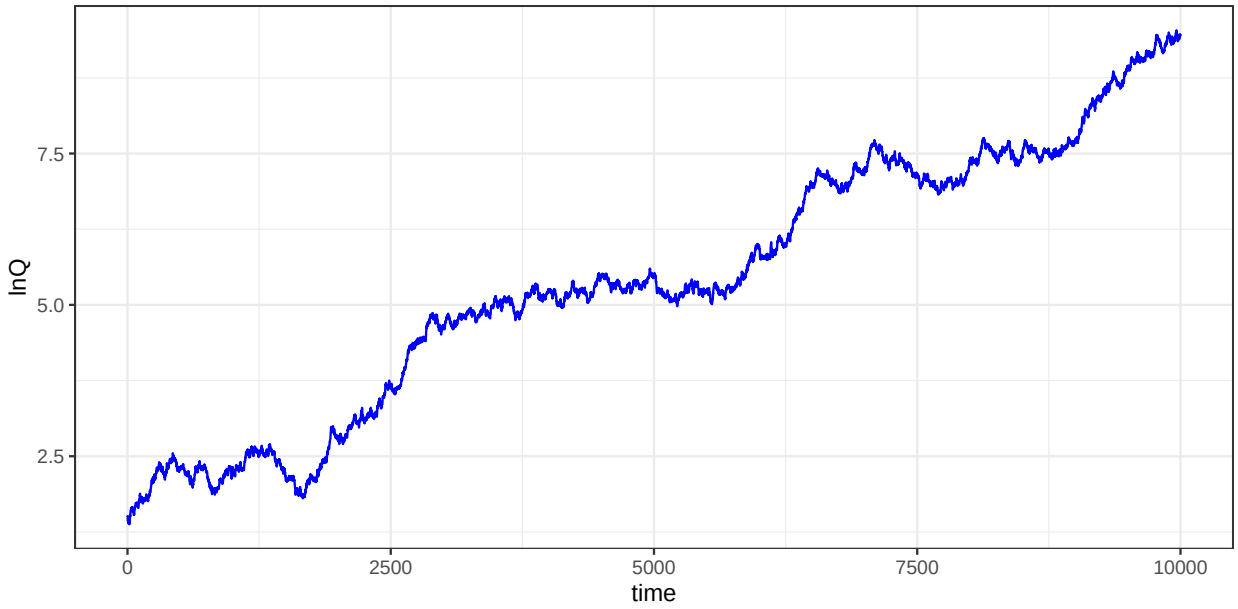

FIGURE 5.32. $\ln(Q_t)$ generated from Geometric Brownian Motion of length 10,000, with parameters $\kappa = 0.001$ and $\sigma = 0.02$), starting at 1.5

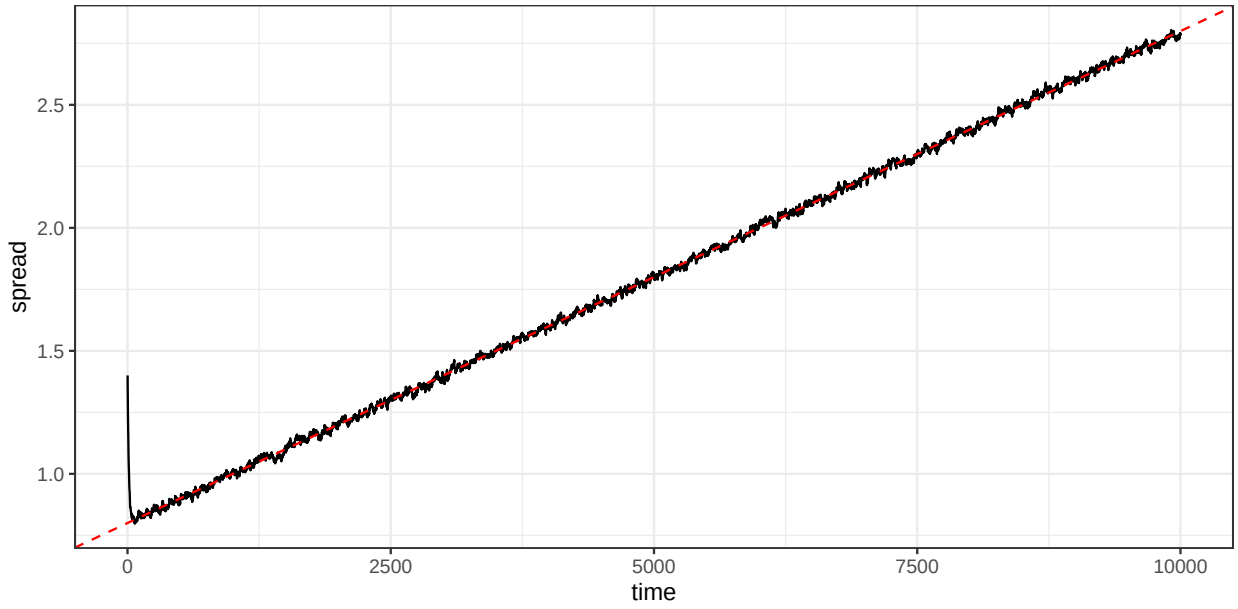

FIGURE 5.33. Spread of length 10,000 from a trend-stationary OU process, with parameters $a = 0.0002$, $b = 0.8$, $\lambda = 0.08$ and $\sigma = 0.005$

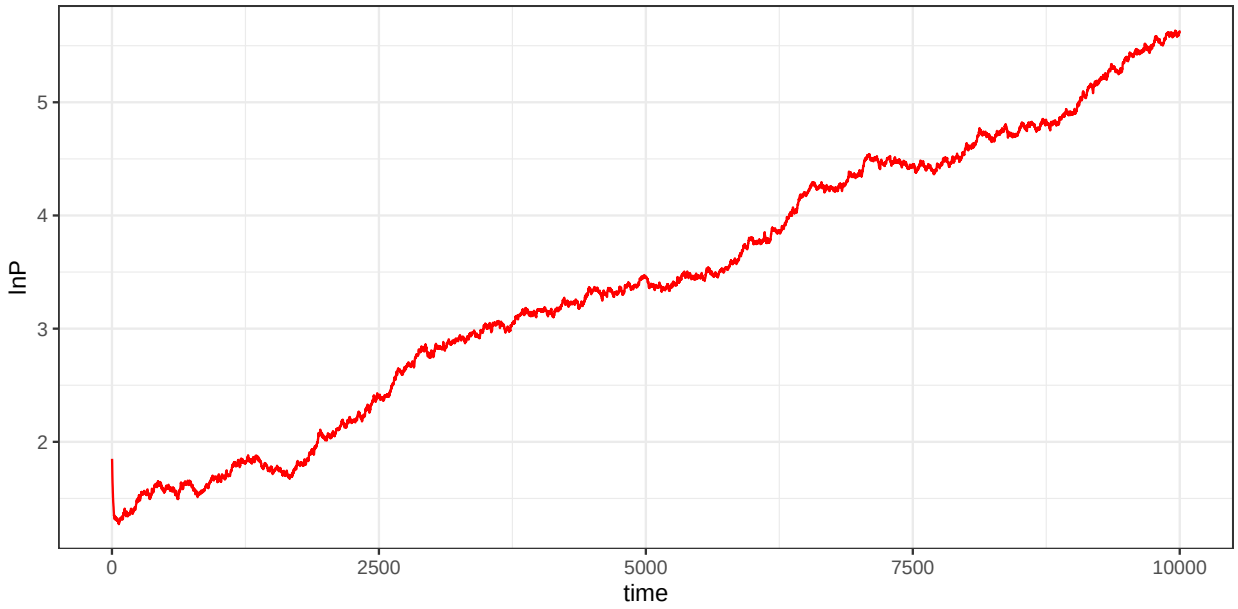

FIGURE 5.34. $ln(P_t)$ obtained from $ln(Q_t)$ and the trend-stationary OU process spread by the cointegration equation.

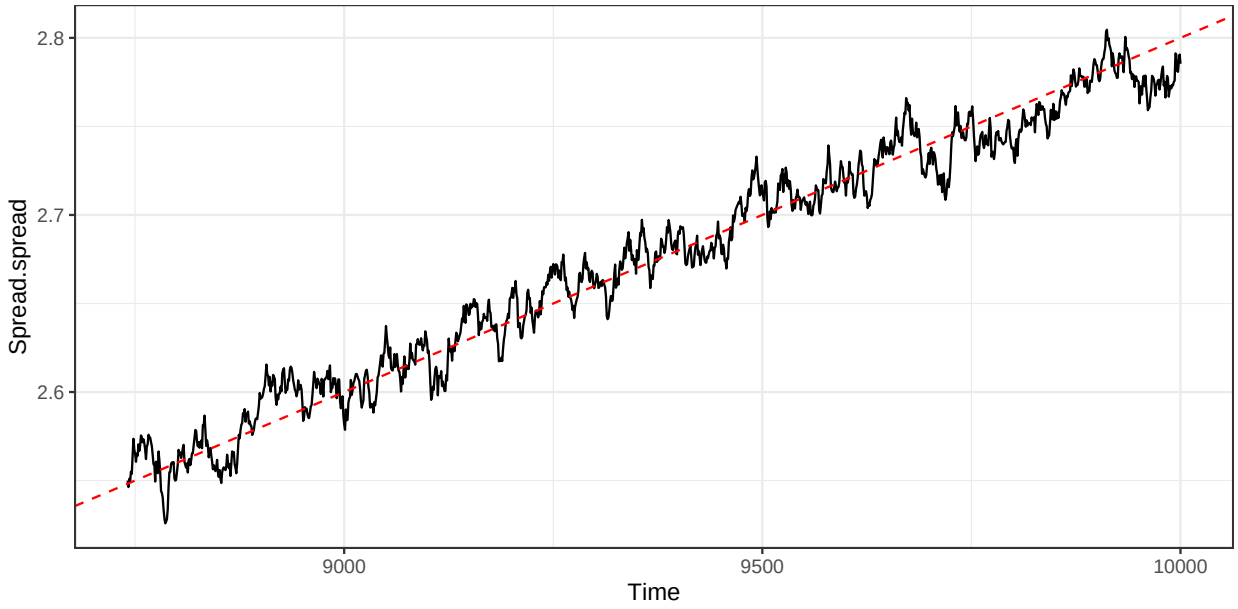

FIGURE 5.35. Last 1,260 steps of the trend-stationary OU process in figure 5.33

5.2.5.2. Long Path

Let us first consider the last 1260 days of the process generated in subsection 5.2.5.1. The OU process and the logarithm of the price time series are presented in figures 5.35 and 5.36 respectively. Notice that the augmented Dickey-Fuller test result 5.17 indicates that all the three linear model types are stationary. But observing the path in figure 5.35, it is more reasonable to consider the trend.

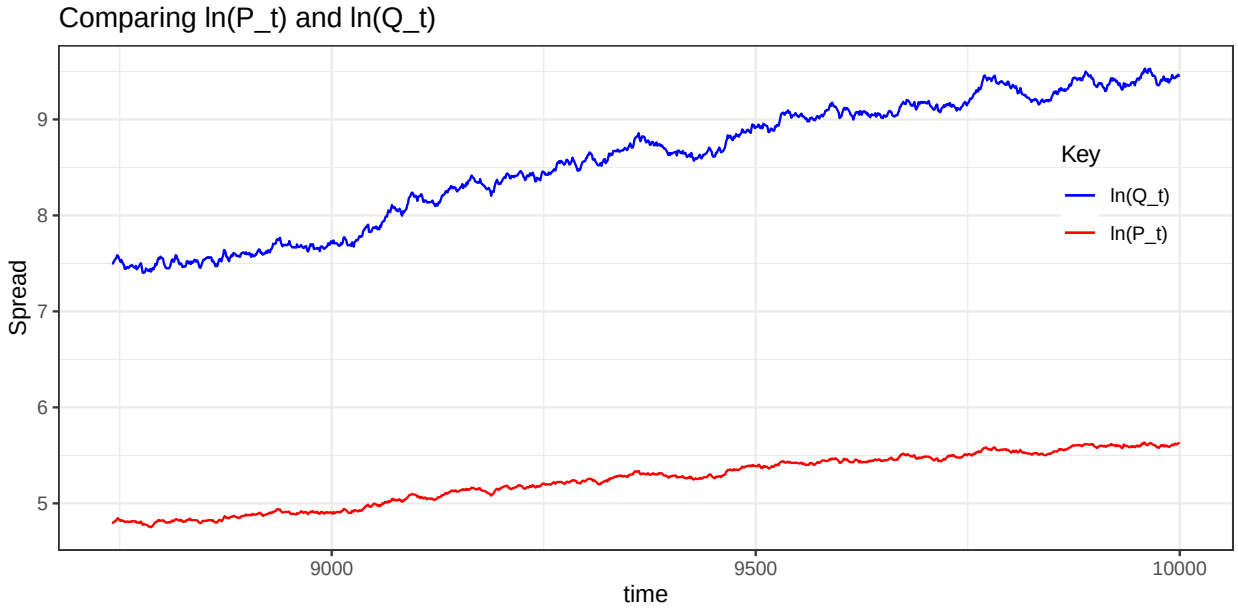

FIGURE 5.36. Last 1,260 steps of $ln(P_t)$ vs. $ln(Q_t)$ from figures 5.32 and 5.34

Linear model type	Lag	ADF	p-value
Type 1: no drift, no trend	0	-6.07	≤ 0.01
Type 2: with drift, no trend	0	-6.07	≤ 0.01
Type 3: with drift and trend	0	-6.12	≤ 0.01

TABLE 5.17. Augmented Dickey-Fulller test for stationarity of the cointegration between logarithmic returns of the artificial pair P_t and Q_t.

Note 1: Alternate hypothesis: Stationary

5.2.5.3. Parameter Estimation

We estimate the parameters η, a, b, λ and σ by maximum likelihood method based on equation 68 and also by the method of least squares using the system of equations obtained earlier in the section. A comparison of the outputs is presented in table 5.18. We observe from the table that the parameters are quite well estimated by both methods except for b for the least squares method, and λ is also slightly off. Based on these parameter values, we reproduce the trend-stationary OU process and plot them in figure 5.37. We see that again the maximum likelihood estimates are better than the least squares estimates. However these estimate for b in the least squares method is not good enough for the purpose of estimating optimal thresholds since the drift level plays a major role in the level of the thresholds.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.28905200	-0.01094800	0.00240500	0.00012564
		LS	0.30552645	0.00552645		
a	0.00020000	MLE	0.00021070	0.00001070	0.00000018	0.00000000
		LS	0.00015396	-0.00004604		
b	0.80000000	MLE	0.80031100	0.00031100	0.02068100	0.00042780
		LS	1.06789968	0.36789968		
λ	0.08000000	MLE	0.08562200	0.00562200	0.01212300	0.00017857
		LS	0.09469577	0.01469577		
σ	0.00500000	MLE	0.00508800	0.00008800	0.00010600	0.00000002
		LS	0.00509058	0.00009058		

TABLE 5.18. Parameter estimation for the trend-stationary OU process representing the spread between lnP_t and lnQ_t, by maximum likelihood and method of least squares.

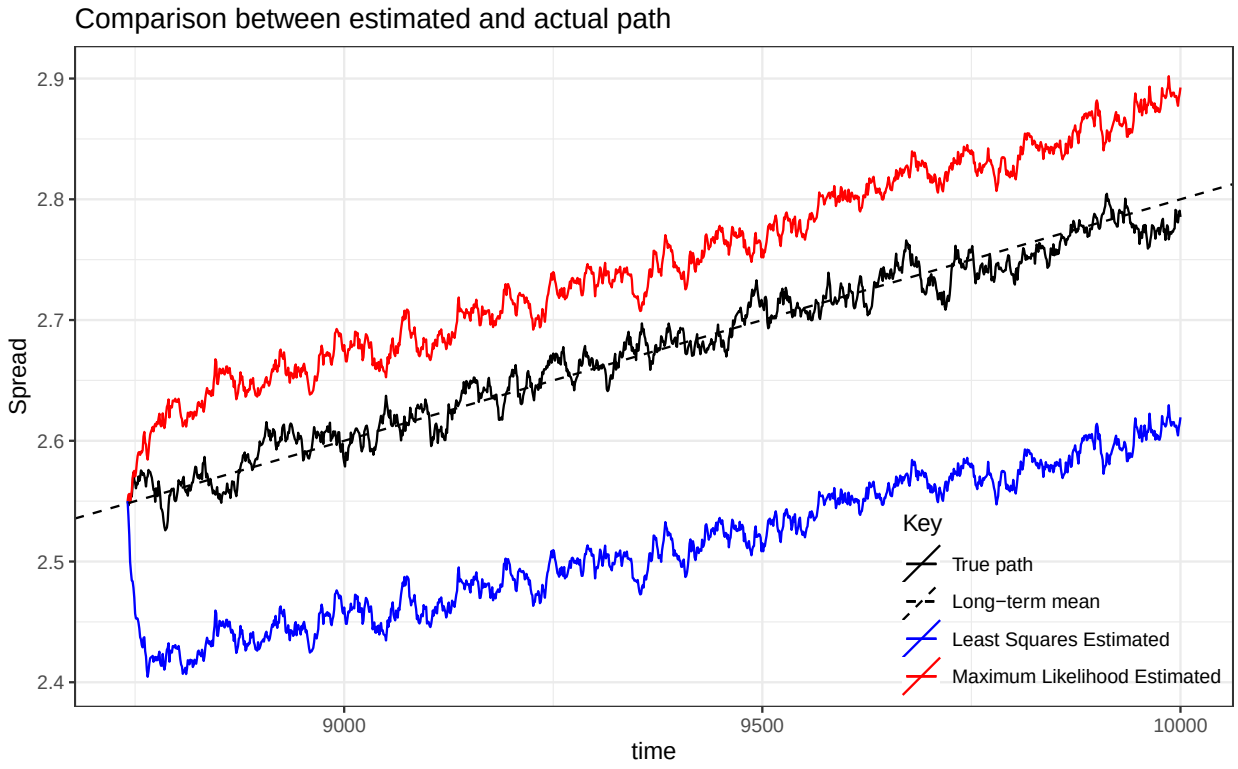

FIGURE 5.37. Spreads of length 1260 each, generated from one realization of parameter estimates for the trend-stationary OU process, from maximum likelihood and least squares methods, using $ln(Q_t)$ and $ln(P_t)$.

We also perform Monte Carlo simulations with 10,000 sample paths and obtain the average estimates for the parameters, using both maximum likelihood and least squares methods as before. Table 5.19 shows the result. We again reproduce the paths based on these averages, and they are plotted in figure 5.38. We observe that the estimators perform very well here.

5.2.5.4. Short Path

we now look at the last 126 days of the processes discussed in 5.2.5.1. Similar to subsection 5.1.5.2, we show the OU process and the logarithm of the price time series of the two artificial stocks in figures 5.39 and 5.40 respectively.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.30049995	0.00049995	0.00000013	0.00000025
		LS	0.29996349	-0.00003651	0.00656867	0.00004315
a	0.00020000	MLE	0.00020167	0.00000167	0.00000018	0.00000000
		LS	0.00020101	0.00000101	0.00002581	0.00000000
b	0.80000000	MLE	0.78000000	-0.02000000	0.00000001	0.00040000
		LS	0.79533500	-0.00466500	0.12626155	0.01596374
λ	0.08000000	MLE	0.08599982	0.00599982	0.00000028	0.00003600
		LS	0.08646885	0.00646885	0.01239505	0.00019548
σ	0.00500000	MLE	0.00500550	0.00000550	0.00008399	0.00000001
		LS	0.00500336	0.00000336	0.00010325	0.00000001

TABLE 5.19. Maximum Likelihood and Least Squares estimates of the trend-stationary OU process parameters for the path of length 1260, by 10,000 Monte Carlo simulations

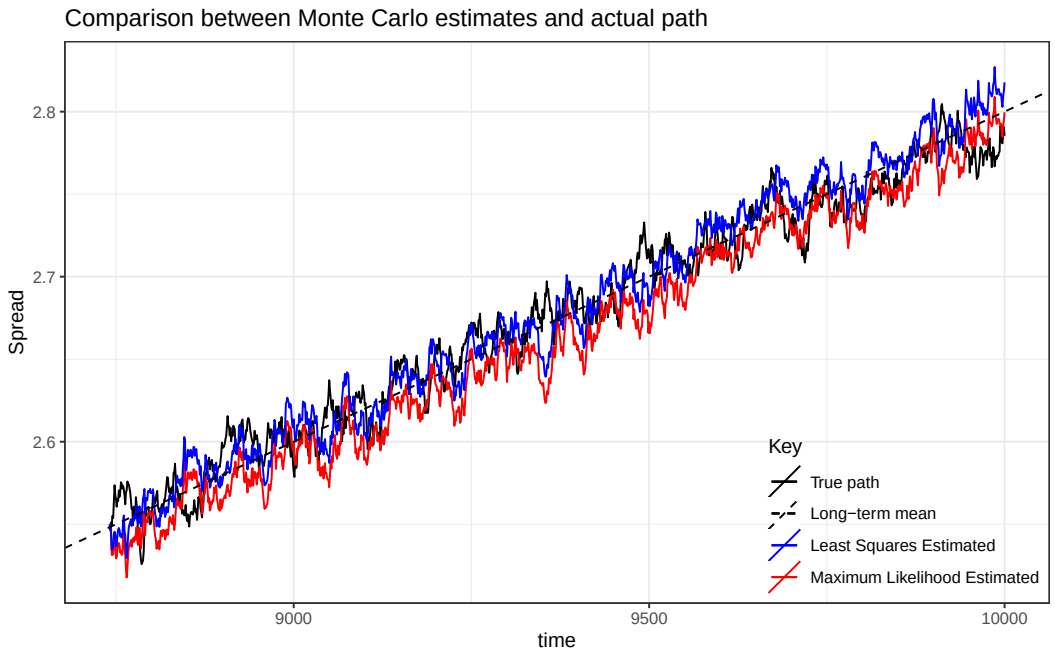

FIGURE 5.38. Spreads of length 1,260 each, generated from parameter estimates by maximum likelihood and least squares methods from $ln(Q_t)$ and $ln(P_t)$, based on Monte Carlo averages. See table 5.19

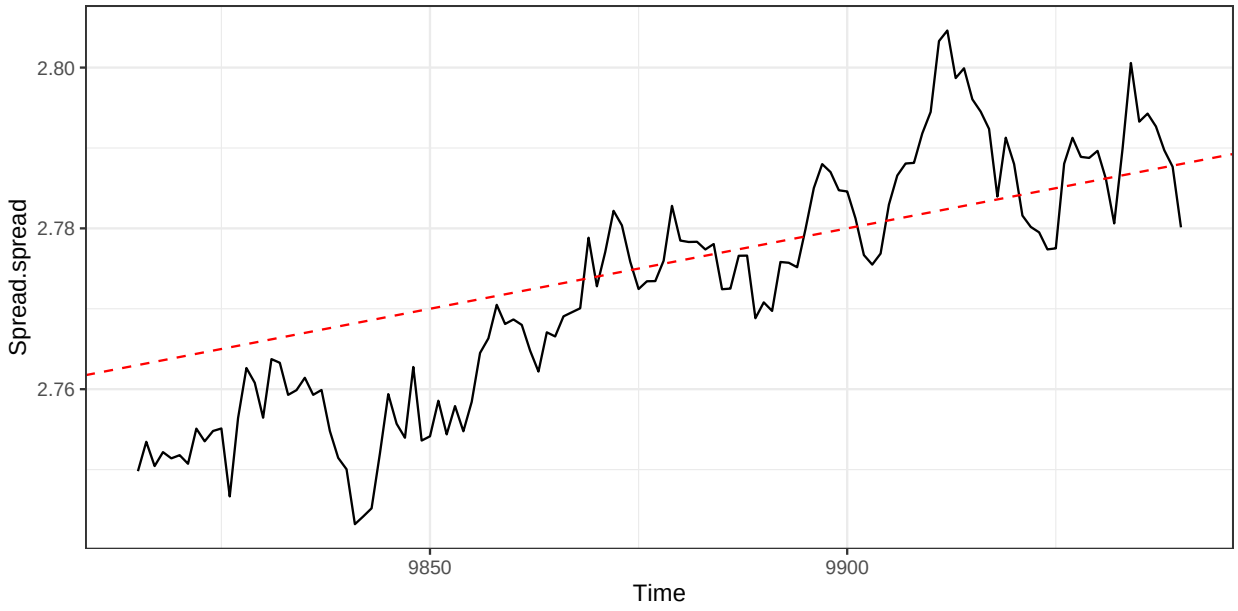

FIGURE 5.39. Last 126 steps of the trend-stationary OU process in figure 5.33

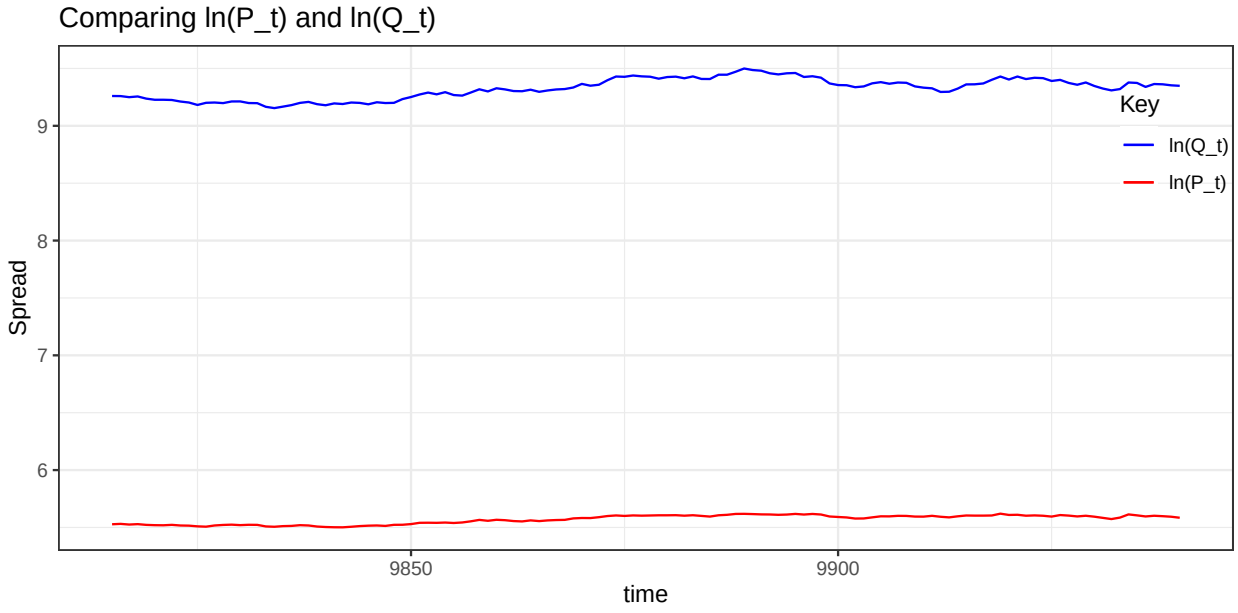

FIGURE 5.40. Last 126 steps of $ln(P_t)$ vs. $ln(Q_t)$ from figures 5.32 and 5.34

Linear model type	Lag	ADF	p-value
Type 1: no drift, no trend	0	-2.73	≤ 0.01
Type 2: with drift, no trend	0	-2.72	0.0788
Type 3: with drift and trend	0	-3.08	0.125

TABLE 5.20. Augmented Dickey-Fulller test for stationarity of the cointegration between logarithmic returns of the artificial pair P_t and Q_t.
Note 1: Alternate hypothesis: Stationary

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.29881380	-0.00118620	0.02079720	0.00043393
		LS	0.28613051	-0.01386949		
a	0.00020000	MLE	0.00020130	0.00000130	0.00000046	0.00000000
		LS	0.00030316	0.00010316		
b	0.80000000	MLE	0.70968410	-0.09031590	0.19413970	0.04584719
		LS	-0.42596944	-1.22596944		
λ	0.08000000	MLE	0.13149820	0.05149820	0.05595530	0.31575162
		LS	0.22047593	0.14047593		
σ	0.00500000	MLE	0.00443500	-0.00065600	0.00030550	0.00000965
		LS	0.00455466	-0.00044534		

TABLE 5.21. Parameter estimates by one realization, for the trend-stationary OU representation of the spread between lnP_t and lnQ_t for 126 steps, by maximum likelihood and method of least squares.

5.2.5.5. Parameter Estimation

We estimate the parameters η, a, b, λ and σ by maximum likelihood method based on equation 68 and also by the method of least squares. A comparison of the outputs is presented in table 5.21. Both methods do not perform well in estimating λ for this case, but

besides this, the least squares estimates for b is also poor, which is critical for our optimal threshold estimation as explained earlier. A visualization of this is also shown in figure 5.41

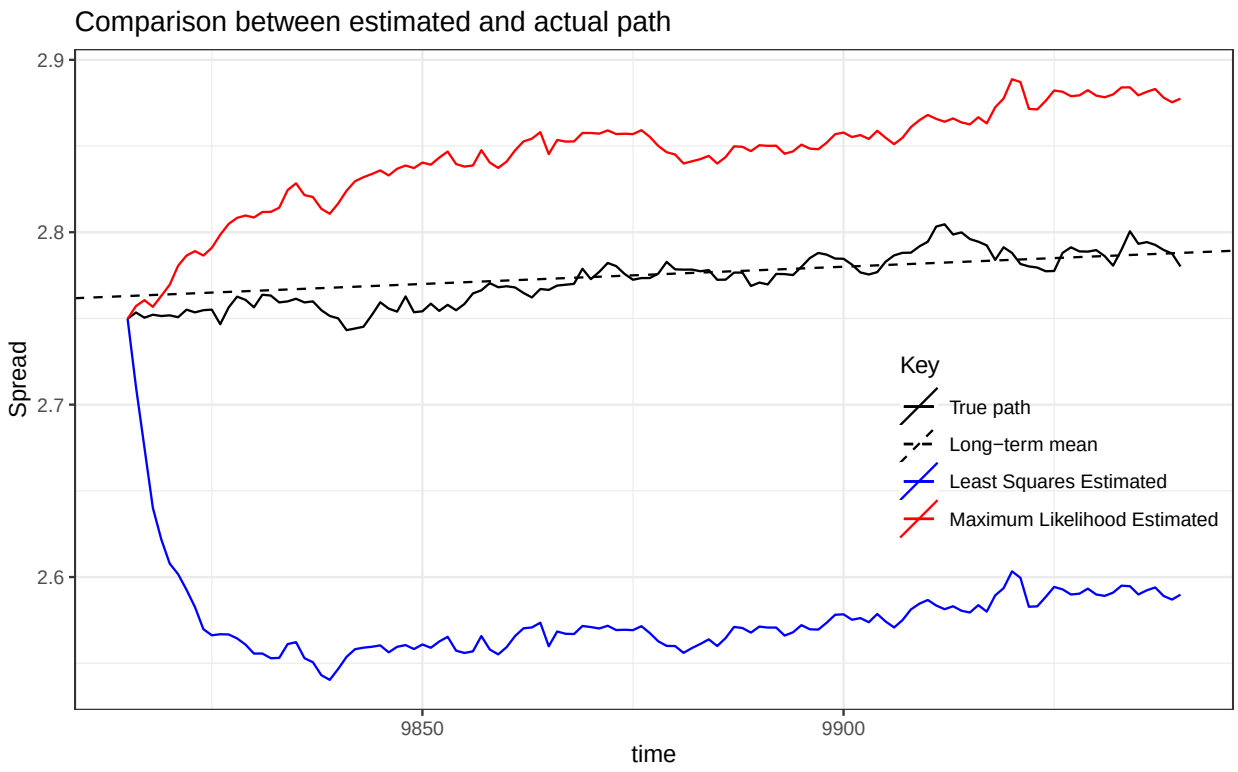

FIGURE 5.41. Trend-stationary OU process representation of spreads of length 126 each, generated from one realization of parameter estimates from maximum likelihood and least squares methods, using $ln(Q_t)$ and $ln(P_t)$

We perform Monte Carlo simulations with 10,000 sample paths and obtain the average estimates for the parameters, using both maximum likelihood and least squares methods. The results are shown in table 5.22. The paths reproduced from these estimates is shown in figure 5.42. We observe here that the estimates perform better than the case of one realization 5.41, although not as good as the case of long paths. We also note here that the path from the maximum likelihood estimates is closer to the true path than the path from the least squares estimates. We notice that the standard error in our estimate for b is also high for both methods.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
η	0.30000000	MLE	0.30049999	0.00049999	0.00000053	0.00000025
		LS	0.29972018	-0.00027982	0.02278868	0.00051940
a	0.00020000	MLE	0.00020155	0.00000155	0.00000056	0.00000000
		LS	0.00018973	-0.00001027	0.00020292	0.00000004
b	0.80000000	MLE	0.78000000	-0.02000000	0.00000001	0.00040000
		LS	0.84965090	0.04965090	1.57989860	2.49854479
λ	0.08000000	MLE	0.08600188	0.00600188	0.00000243	0.00003602
		LS	0.16250616	0.08250616	0.07230137	0.01203475
σ	0.00500000	MLE	0.00488397	-0.00011603	0.00024220	0.00000007
		LS	0.00509510	0.00009510	0.00034781	0.00000013

TABLE 5.22. Parameter estimates by Maximum Likelihood and Least Squares methods for the 126 steps OU process, by Monte Carlo simulations

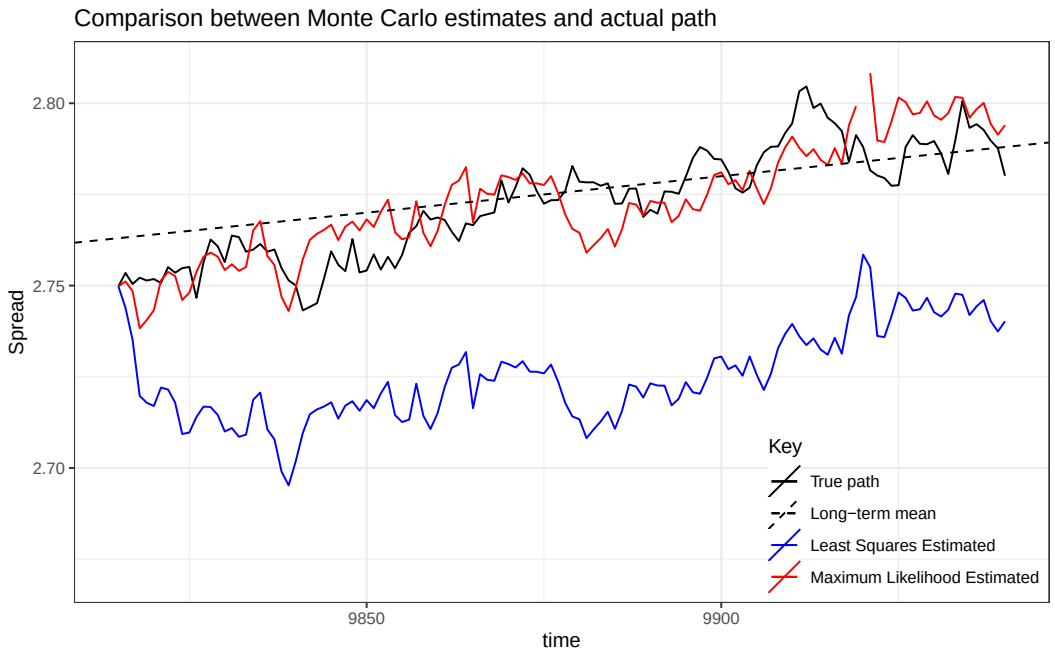

FIGURE 5.42. Spreads of length 126 each, generated from parameter estimates by maximum likelihood and least squares methods from $ln(Q_t)$ and $ln(P_t)$, based on Monte Carlo averages. See table 5.22

5.2.6. Optimal Threshold

For similar reasons as discussed in subsection 5.1.6, we will estimate η separately by simple linear regression of P_t on Q_t and then estimate the remaining four parameters using maximum likelihood estimates as shown in subsections 5.2.3 and 5.2.4.

We will scaled the time from years to days in the trend. Thus $t^{\dagger} = 252t$, since there are 252 trading days in a year, based on the US financial market.

5.2.6.1. Long Path

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00018800	-0.00001200	0.00000017	0.00000000
		LS	0.00018798	-0.00001202		
b	0.80000000	MLE	0.91314500	0.11314500	0.00159200	0.01280433
		LS	0.91315080	0.11315080		
λ	0.08000000	MLE	0.09000600	0.01000600	0.01251900	0.00025663
		LS	0.09004500	0.01004500		
σ	0.00500000	MLE	0.00508300	0.00008300	0.00010600	0.00000002
		LS	0.00508510	0.00008510		

TABLE 5.23. One realization of parameter estimates by maximum likelihood and least squares methods, for the trend-stationary OU process representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$, with actual parameters $a = 0.0002$, $b = 0.8$, $\lambda = 0.08$ and $\sigma = 0.005$

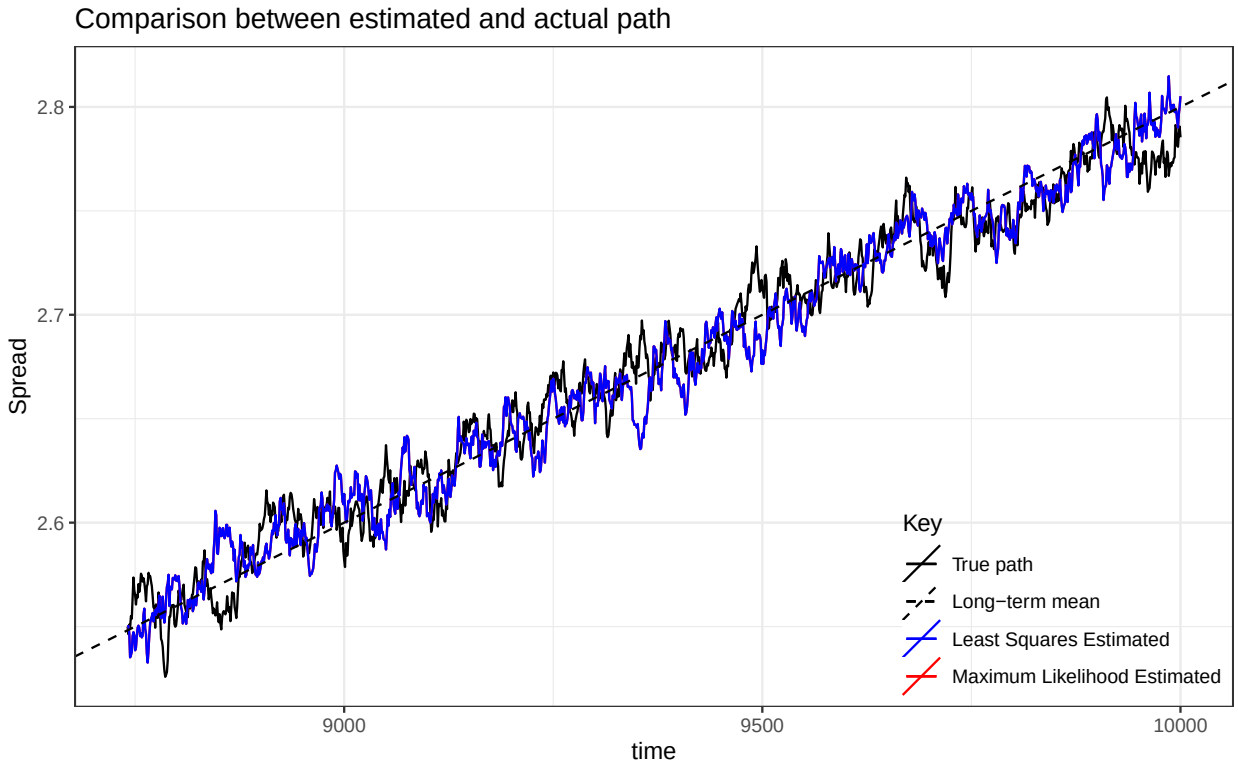

FIGURE 5.43. Paths generated from one realization of parameter estimates by maximum likelihood and least squares methods, for the trend-stationary OU process representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

5.2.6.2. Case 1: $\gamma = 0$

The thresholds for this case are given by

$$g(t) = at^{\dagger} + b \pm \frac{\sigma}{\sqrt{2\lambda}}\beta e^{-\lambda t}. \tag{74}$$

Substituting the time horizon $T = 5$ into the corresponding $h(\beta)$ in chapter 4 and solving gives the optimizer as $\tilde{\beta} \approx 0.9715$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{a}t^{\dagger} + \tilde{b} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(0.9715)e^{-\tilde{\lambda} t}.$$

From maximum likelihood estimates in table 5.23, we take the values of the parameters as $\tilde{a} = 0.0002$, $\tilde{b} = 0.9131$, $\tilde{\lambda} = 0.0900$ and $\tilde{\sigma} = 0.0051$. Hence we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &= 0.0002t^{\dagger} + 0.9131 \pm \frac{0.0051}{\sqrt{2(0.0900)}}(0.9715)e^{-0.0900t} \\ &= 0.9131 + 0.0002t^{\dagger} \pm 0.0117e^{-0.0602t} \end{aligned} \tag{75}$$

A graphical representation of this result is shown in figure 5.44.

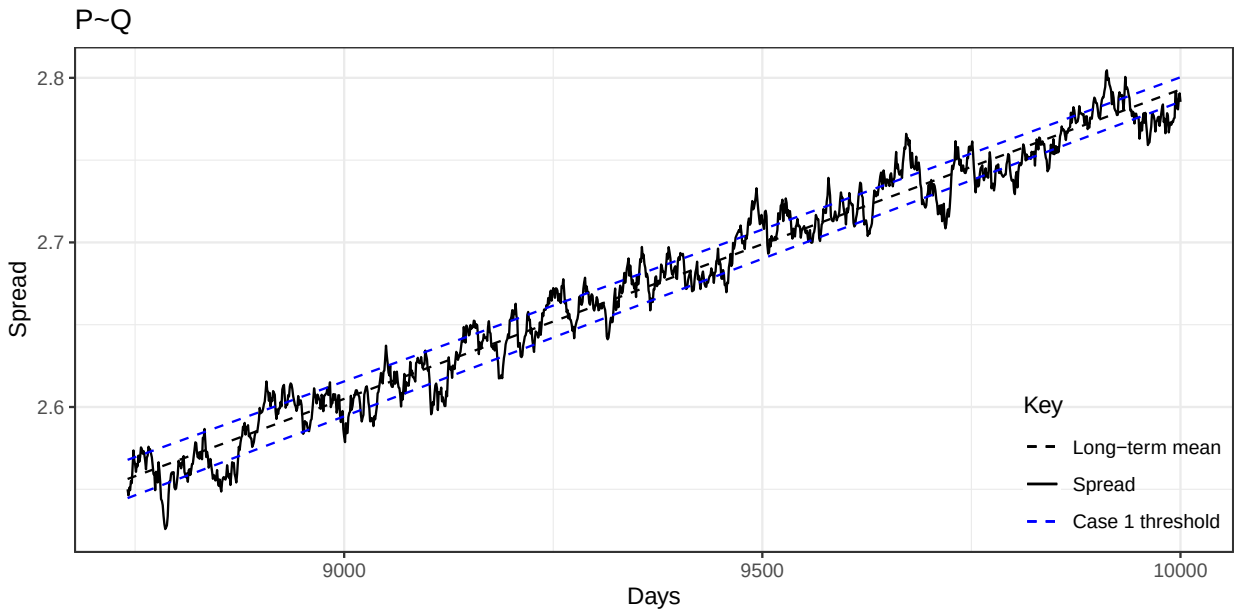

FIGURE 5.44. New threshold case 1 on the trend-stationary OU process representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

5.2.6.3. Case 2: $\gamma \geq 0$ and $\rho = 1$

For this, the threshold is given by:

$$g(t) = at^{\dagger} + b \pm \frac{\sigma}{\sqrt{2\lambda}}(\beta e^{-\lambda t} - \gamma e^{\lambda t}). \tag{76}$$

Substituting the time horizon $T = 5$ into the corresponding $h(\beta, \gamma)$ in chapter 4 and solving gives the optimizer as $(\tilde{\beta}, \tilde{\gamma}) \approx (1.3351, 0.1856)$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{a}t^{\dagger} + \tilde{b} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(1.3351e^{-\tilde{\lambda}t} - 0.1856e^{\tilde{\lambda}t}).$$

Taking the estimates of the OU process parameters to be $\tilde{a} = 0.0002$, $\tilde{b} = 0.9131$, $\tilde{\lambda} = 0.0900$ and $\tilde{\sigma} = 0.0051$, as in case 1, we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &\approx 0.0002t^{\dagger} + 0.9131 \pm \frac{0.0051}{\sqrt{2(0.0900)}}(1.3351e^{-0.0900t} - 0.1856e^{0.0900t}) \\ &= 0.9131 + 0.0002t^{\dagger} \pm 0.0117(1.3351e^{-0.0900t} - 0.1856e^{0.0900t}) \end{aligned} \tag{77}$$

A graphical representation of this result is also shown in figure 5.45.

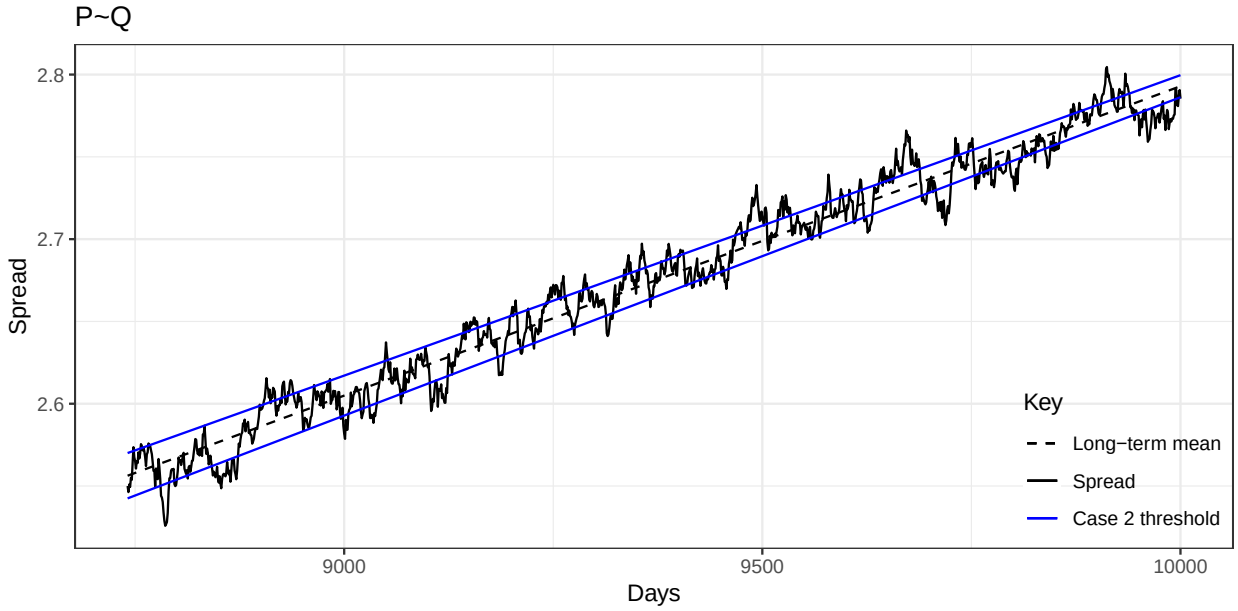

FIGURE 5.45. New threshold case 2 on the trend-stationary OU process representation of the spread of length 1260 between $ln(Q_t)$ and $ln(P_t)$

5.2.6.4. Short Path

We now look at the short-term trade time horizon.

	True Value	Method	Estimate	Bias	Standard Error	RMSE
a	0.00020000	MLE	0.00034270	0.00014270	0.00000020	0.00000002
		LS	0.00034272	0.00014272		
b	0.80000000	MLE	-0.61274830	-1.41274830	0.00187640	1.99586128
		LS	-0.61275034	-1.41274830		
λ	0.08000000	MLE	0.21816210	0.13816210	0.06937040	0.02390102
		LS	0.21808244	0.13808244		
σ	0.00500000	MLE	0.00455430	-0.00044570	0.00032350	0.00000030
		LS	0.00457163	-0.00042837		

TABLE 5.24. One realization of parameter estimates by maximum likelihood and least squares methods, for the OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$, with true parameters $a = 0.0002$, $= 0.08$, $\lambda = 0.08$ and $\sigma = 0.005$

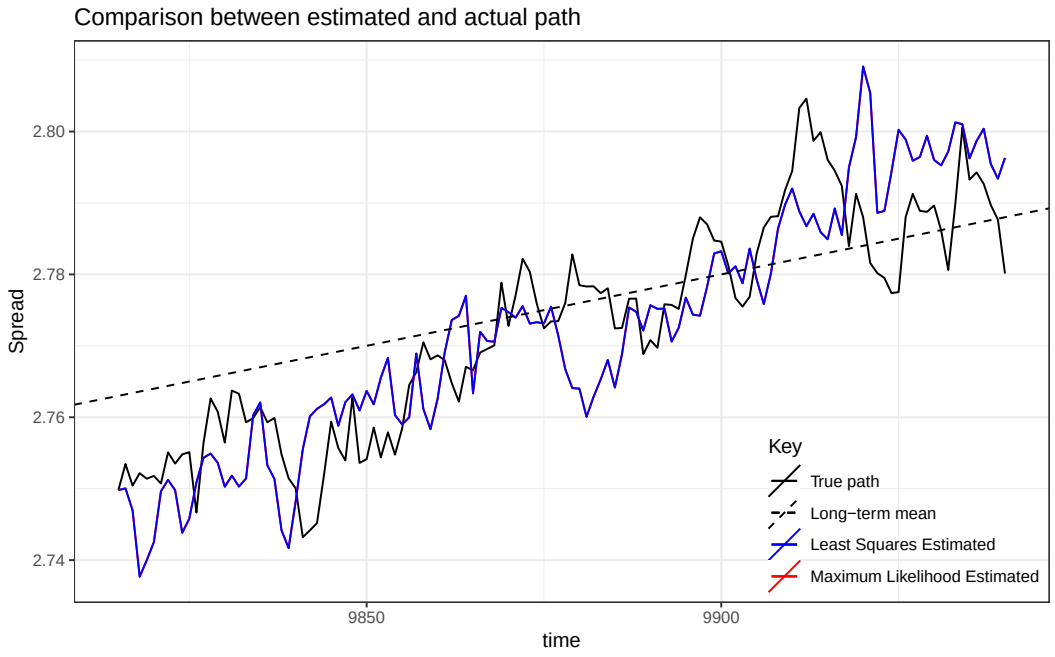

FIGURE 5.46. Paths generated from one realization of parameter estimates by maximum likelihood and least squares methods, for the trend-stationary OU process representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

5.2.6.5. Case 1: $\gamma = 0$

Substituting the time horizon $T = \frac{1}{2}$ into equation the corresponding $h(\beta)$ in chapter 4 and solving gives the optimizer of the expected return as $\tilde{\beta} \approx 0.8735$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{a}t^{\dagger} + \tilde{b} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(0.8735)e^{-\tilde{\lambda}t}.$$

From maximum likelihood estimates in table 5.23, we take the values of the parameters as $\tilde{a} = 0.0003$, $\tilde{b} = -0.6127$, $\tilde{\lambda} = 0.2181$ and $\tilde{\sigma} = 0.0046$. Hence we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &= 0.0003t^{\dagger} - 0.6127 \pm \frac{0.0046}{\sqrt{2(0.2181)}}(0.8735)e^{-0.2181t} \\ &= -0.6127 + 0.0003t^{\dagger} \pm 0.0061e^{-0.2181t} \end{aligned} \tag{78}$$

A graphical representation of this result is shown in figure 5.47.

5.2.6.6. Case 2: $\gamma \geq 0$ and $\rho = 1$

Substituting the time horizon $T = \frac{1}{2}$ into corresponding $h(\beta, \gamma)$ in chapter 4 and solving gives the optimizer of the expected return as $(\tilde{\beta}, \tilde{\gamma}) \approx (2.2621, 0.6944)$, approximated to four decimal places. Thus the optimal thresholds for this case are:

$$\tilde{g}(t) \approx \tilde{a}t^{\dagger} + \tilde{b} \pm \frac{\tilde{\sigma}}{\sqrt{2\tilde{\lambda}}}(2.2621e^{-\tilde{\lambda}t} - 0.6944e^{\tilde{\lambda}t}).$$

Taking the estimates of the OU process parameters to be $\tilde{a} = 0.0003$, $\tilde{b} = -0.6127$, $\tilde{\lambda} = 0.2181$ and $\tilde{\sigma} = 0.0046$, as in case 1, we approximate the optimal thresholds as:

$$\begin{aligned} \tilde{g}(t) &\approx 0.0003t^{\dagger} - 0.6127 \pm \frac{0.0046}{\sqrt{2(0.2181)}}(2.2621e^{-0.2181t} - 0.6944e^{0.2181t}) \\ &= -0.6127 + 0.0003t^{\dagger} \pm 0.0070(2.2621e^{-0.2181t} - 0.6944e^{0.2181t}) \end{aligned} \tag{79}$$

A graphical representation of this result is shown in figure 5.48.

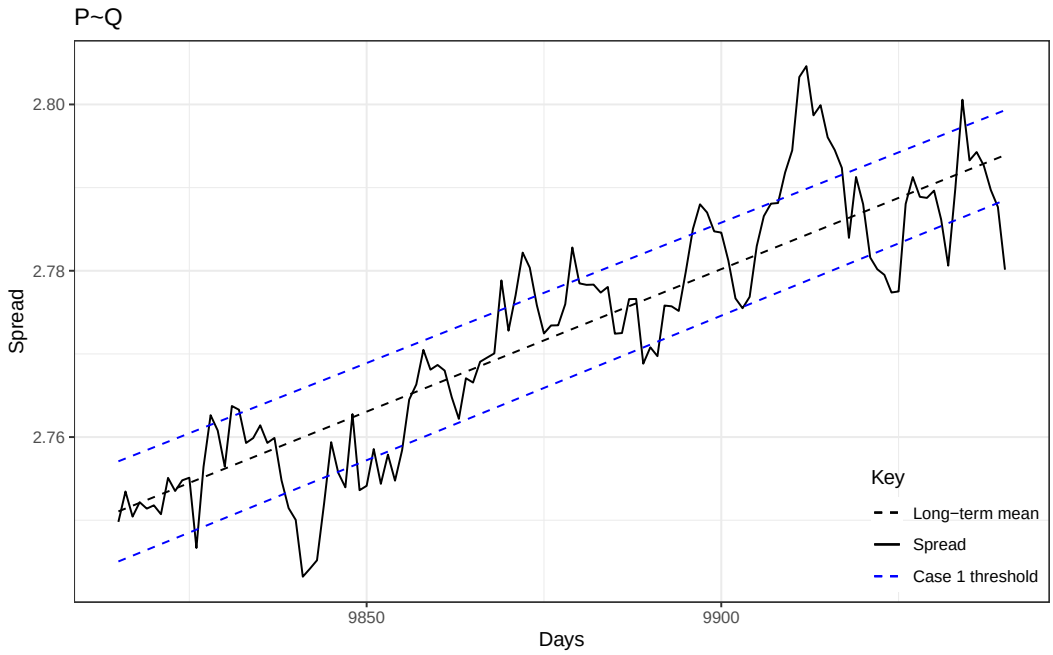

FIGURE 5.47. New threshold case 1 on the trend-stationary OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

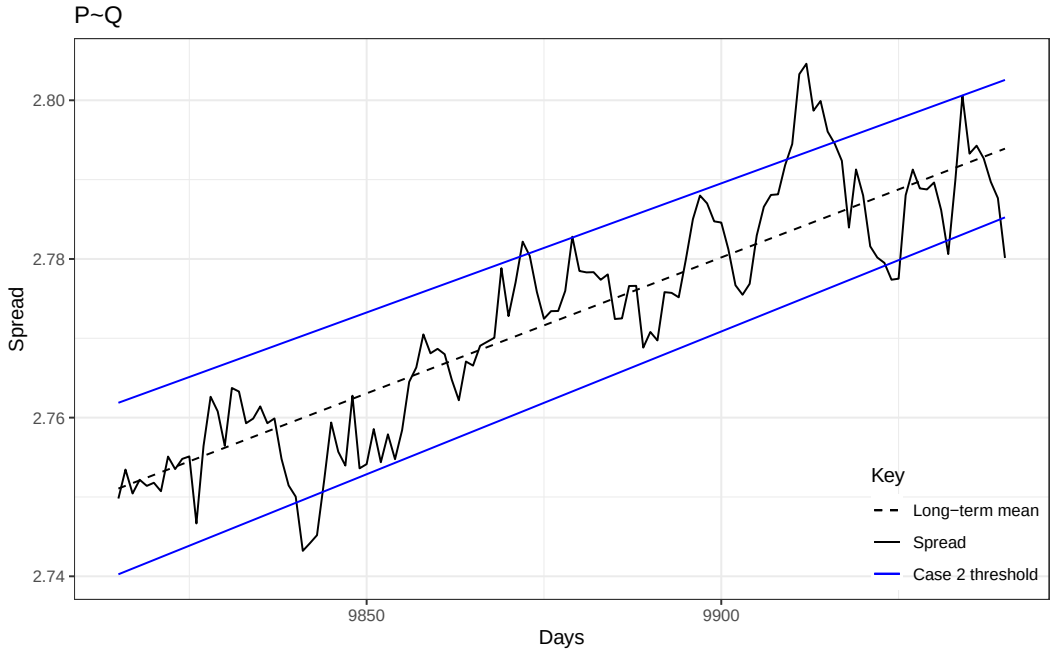

FIGURE 5.48. New threshold case 2 on the trend-stationary OU representation of the spread of length 126 between $ln(Q_t)$ and $ln(P_t)$

Similar to previous cases, the optimal thresholds provided by case 1 has a narrower

band and as such is crossed more often by the path, while the optimal threshold presented in case 2 has a broader band, and as a result may have fewer number of crossings than case 1, but may generate higher return per trade cycle.

CHAPTER 6

APPLICATIONS

In this chapter we apply our results to four pairs of stocks that are often traded together. Namely, Coca-Cola (KO)/Pepsi (PEP), Chevron (CVX)/ Exxon Mobil(XOM), Target (TGT)/Walmart (WMT), and RWE AG (RWE.DE)/E.OnSe (EOAN.DE), [40] and [17]. We will consider both short and long term trade time horizons.

We will first use simple linear regression to obtain the value of the parameter η for the cointegration model and then proceed to estimate the remaining parameters by least squares and maximum likelihood estimation. This conforms with conventional practice as shown in [17] and [40].

We then compare our results with theirs and show that our strategy could potentially yield higher returns than the constant threshold methods in their strategies.

6.1. Thresholds

6.1.1. Zeng and Lee's Threshold

The strategy of Zeng and Lee [40] seeks to find optimal thresholds for the trade, that maximize the expected return per unit time. They achieve this by utilizing the elementary renewal theorem to derive this expectation and the optimization methods resulted in an implicit equation involving infinite series:

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)=(a-\frac{c}{2})\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right), \tag{80}$$

where c is the transaction cost and a is the level of the optimal threshold. Thus given the transaction cost c, one can solve for a numerically.

6.1.2. Goncu and Akyildirim's Threshold

The method Goncu and Akyildirim [17] on the other hand is that given an investment time horizon T, they seek to find the level c such that the probability of a successful trade

is maximized. By this they mean successfully reaching the long term mean, from the level c, within the time horizon T. Their optimization method resulted in the equation,

$$c_*(T) = \sqrt{\frac{1 - e^{-2T}}{e^{-2T}}}, \quad \text{where} \quad T > 0 \tag{81}$$

6.1.3. New Thresholds

The new strategy we present, as discussed in chapter 4, differs from the other strategies, in that the thresholds are not constant, but rather vary over time, and we obtain the optimal threshold by finding parameter values that maximize the expected return of a complete trade cycle. We look into two cases of our general threshold 12, and in addition we also consider versions of these two thresholds that take into account trend information in the data.

6.1.4. Time Scaling

As explained earlier $t^\dagger$ is a scaled form of t, and the choice of scale depends on the context of application. In our case, we are considering days and years. Instead of using different notations for the two, we will convert all times to days at this point and let Δt be equivalent to one day. So t will represent days for the rest of this material and not years.

6.2. Short-Term Trades

For the short term, we will consider the last 126 days up until June 30th, 2021, which is equivalent to six months in the US financial market. RWE and EOAN are German utility companies and have more trading days in a week than the US market, so the start date of this pair is different from the others. The restriction of 126 days for short term trading cycles makes it difficult to obtain accurate estimates for our parameters for the OU model in this case, as such we will be using the least squares method to obtain initial estimates which will then be passed on as starting point for the Newton-Raphson approximation in the maximum likelihood estimation. This way we have good starting points and the maximum likelihood estimation of the maxLik package in R will converge faster and be more accurate.

6.2.1. Pepsi(PEP) and Coca-Cola(KO)

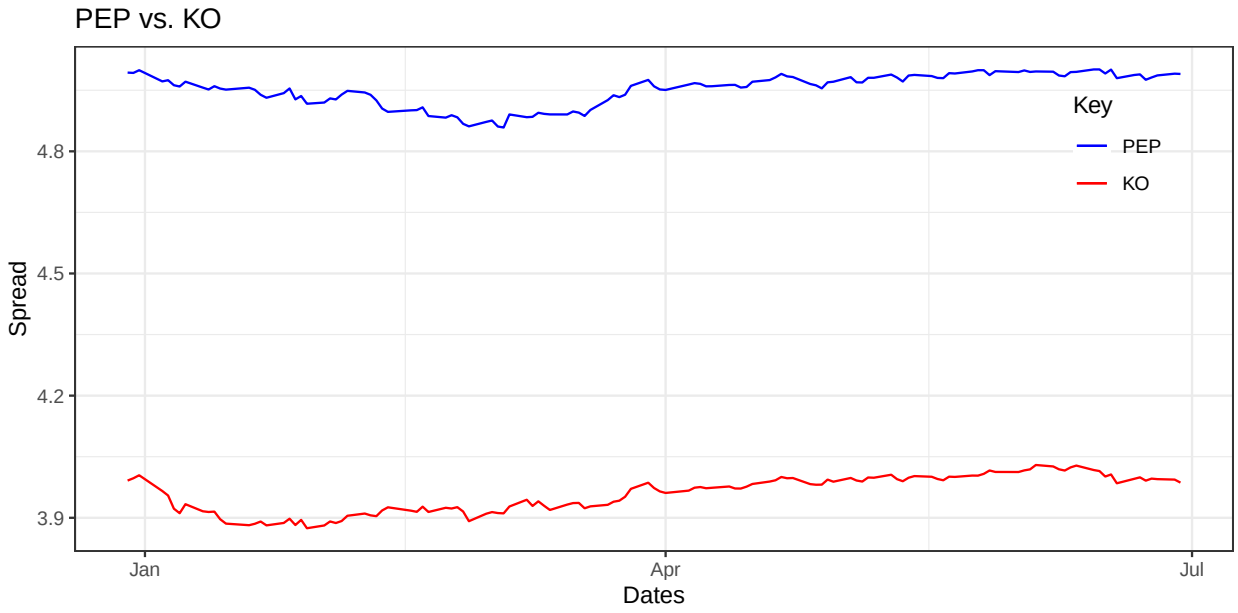

FIGURE 6.1. Logarithmic returns of PEP and KO for first half of 2021

6.2.1.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	2.2766972	0.0154464
	LS	2.276706548	
λ	MLE	0.0438524	0.0270987
	LS	0.043905319	
σ	MLE	0.0075725	0.0004896
	LS	0.007603501	

TABLE 6.1. Parameter estimates for OU process representation of spread between PEP and KO for first half of 2021

	Method	Estimate	Standard Error
a	MLE	0.0002681	0.0005025
	LS	0.0002680341	
b	MLE	2.2538407	0.0477373
	LS	2.2538261300	
λ	MLE	0.0419085	0.0273073
	LS	0.0419038131	
σ	MLE	0.0075550	0.0004887
	LS	0.0075852657	

TABLE 6.2. Parameter estimates for the trend-stationary OU process representation of spread between PEP and KO for first half of 2021

6.2.1.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	2.3060	2.2474
Goncu	2.3102	2.2432
New case 1, no trend	2.2767 + 0.0228 $e^{-0.0419(\frac{t}{252})}$	2.2767 - 0.0228 $e^{-0.0419(\frac{t}{252})}$
New case 2, no trend	2.2767 + 0.0590 $e^{-0.0419(\frac{t}{252})}$ - 0.0181 $e^{0.0419(\frac{t}{252})}$	2.2767 - 0.0590 $e^{-0.0419(\frac{t}{252})}$ + 0.0181 $e^{0.0419(\frac{t}{252})}$
New case 1, with trend	2.2538 + 0.0003 t + 0.0228 $e^{-0.0419(\frac{t}{252})}$	2.2538 + 0.0003 t - 0.0228 $e^{-0.0419(\frac{t}{252})}$
New case 2, with trend	2.2538 + 0.0003 t + 0.0590 $e^{-0.0419(\frac{t}{252})}$ - 0.0181 $e^{0.0419(\frac{t}{252})}$	2.2538 + 0.0003 t - 0.0590 $e^{-0.0419(\frac{t}{252})}$ + 0.0181 $e^{0.0419(\frac{t}{252})}$

TABLE 6.3. Thresholds for the various trading strategies for the spread between PEP and KO for first half of 2021

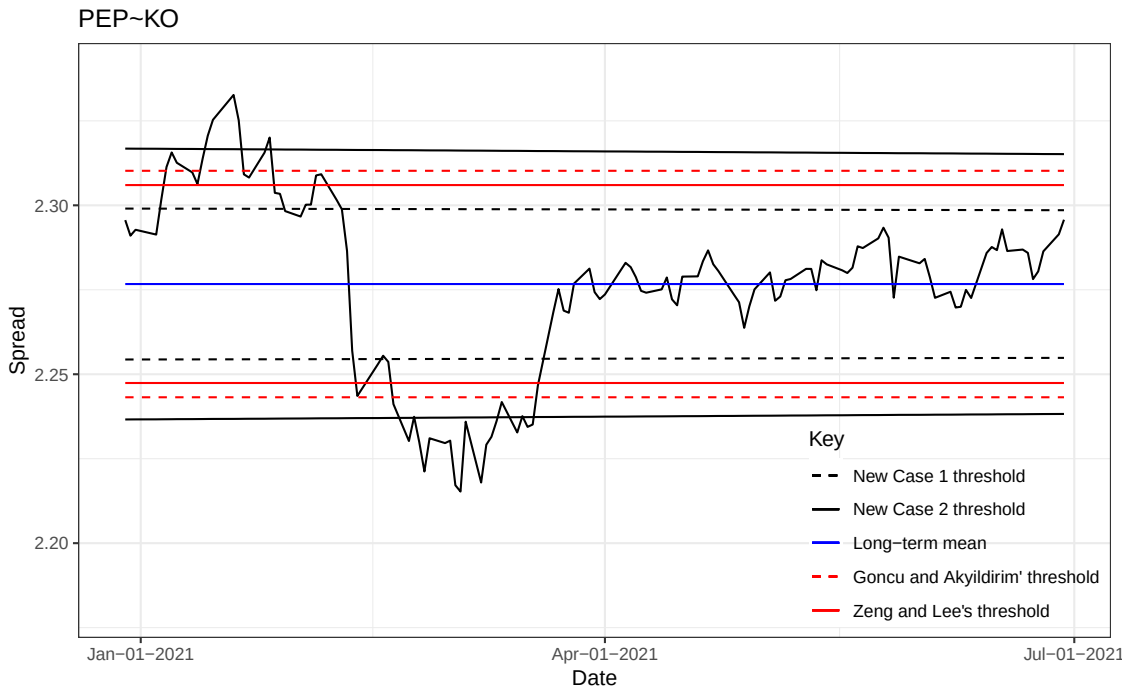

FIGURE 6.2. Zero trend new thresholds against Old thresholds for PEP and KO

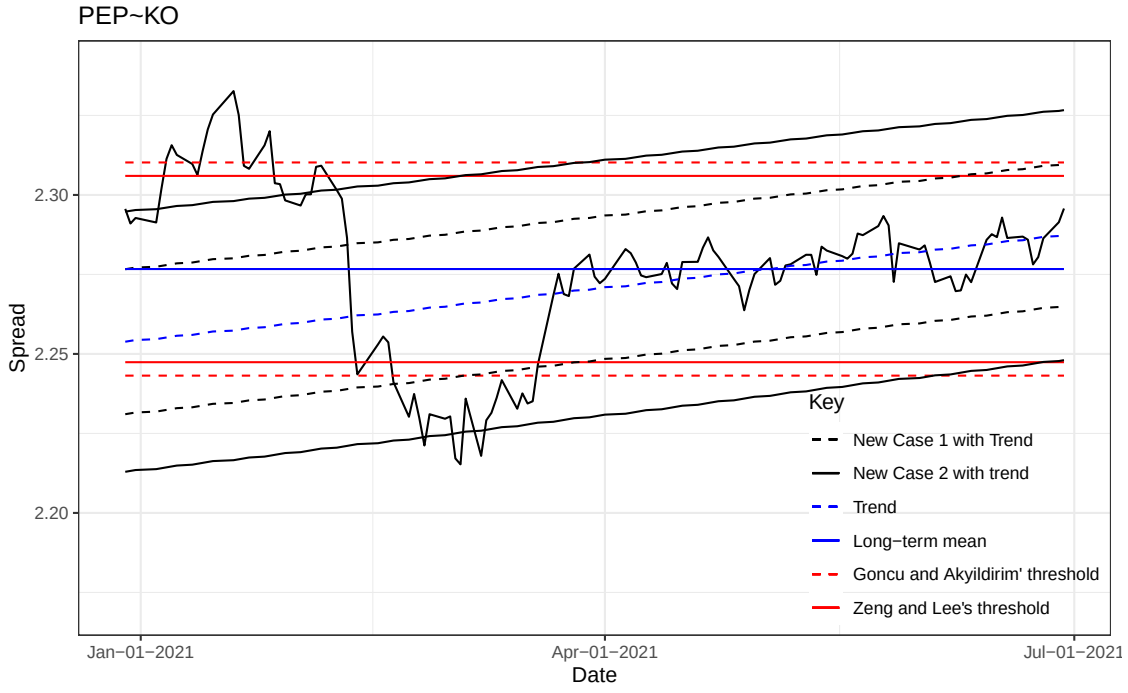

FIGURE 6.3. Nonzero trend new thresholds against Old thresholds for PEP and KO

From figures 6.2 and 6.3, we see that for the short term pairs trading between Pepsi and Coca cola, all the strategies yield exactly two trades. However case 2 of our new threshold provides a wider band, and consequently yields higher return than the others, and it is even much higher for the trending case. See table 6.13 fro the comparison.

6.2.2. E.OnSe (EOAN) and RWE AG(RWE) (German Utility Companies)

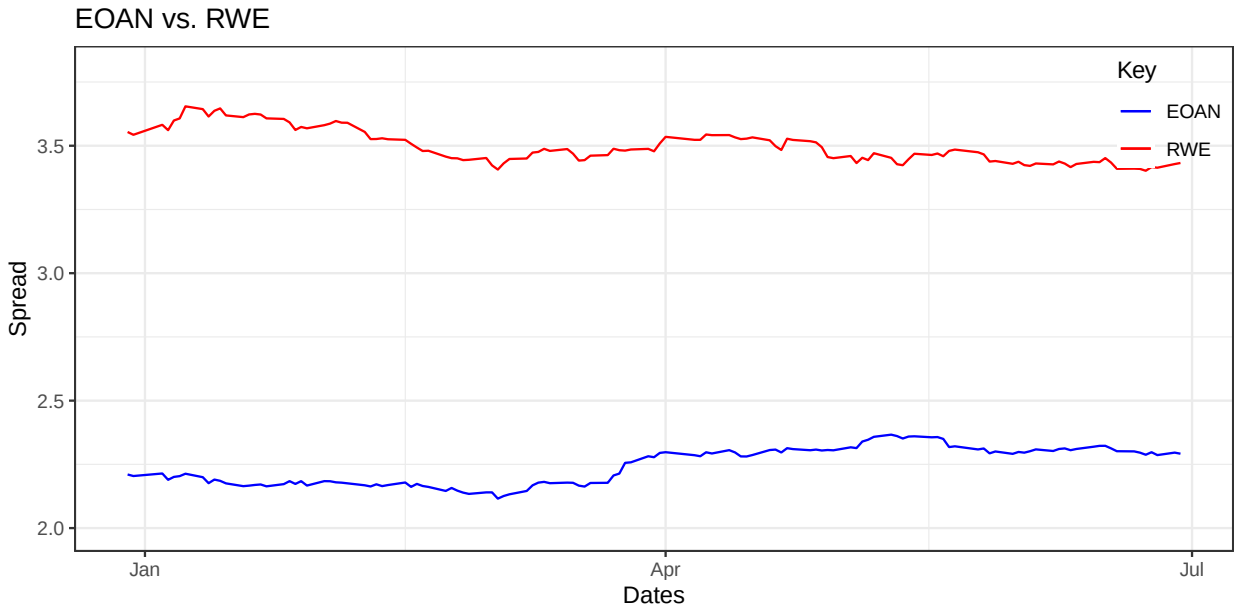

FIGURE 6.4. Logarithmic returns of EOAN and RWE for first half of 2021

6.2.2.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	3.80459000	0.04798800
	LS	3.80467018	
λ	MLE	0.03106200	0.02274100
	LS	0.03106018	
σ	MLE	0.01657000	0.00106500
	LS	0.01663610	

TABLE 6.4. Parameter estimates for OU process representation of spread between EOAN and RWE for first half of 2021

	Method	Estimate	Standard Error
a	MLE	0.0010911	0.0008929
	LS	0.001089163	
b	MLE	3.7117340	0.0822731
	LS	3.711795198	
λ	MLE	0.0462500	0.0274028
	LS	0.046237329	
σ	MLE	0.0166238	0.0010752
	LS	0.016691165	

TABLE 6.5. Parameter estimates for the trend-stationary OU process representation of spread between EOAN and RWE for first half of 2021

6.2.2.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	3.8577	3.7514
Goncu	3.8917	3.7174
New case 1, no trend	3.8045 + 0.0477 $e^{-0.0462(\frac{t}{252})}$	3.8046 - 0.0477 $e^{-0.0462(\frac{t}{252})}$
New case 2, no trend	3.8045 + 0.0477 $e^{-0.0462(\frac{t}{252})}$ - 0.1236 $e^{0.0462(\frac{t}{252})}$	3.8046 - 0.0477 $e^{-0.0462(\frac{t}{252})}$ + 0.1236 $e^{0.0462(\frac{t}{252})}$
New case 1, with trend	3.7117 + 0.0011 t + 0.0477 $e^{-0.0462(\frac{t}{252})}$	3.7117 + 0.0011 t - 0.0477 $e^{-0.0462(\frac{t}{252})}$
New case 2, with trend	3.7117 + 0.0011 t + 0.0477 $e^{-0.0462(\frac{t}{252})}$ - 0.1236 $e^{0.0462(\frac{t}{252})}$	3.7117 + 0.0011 t - 0.0477 $e^{-0.0462(\frac{t}{252})}$ + 0.1236 $e^{0.0462(\frac{t}{252})}$

TABLE 6.6. Thresholds for the various trading strategies for the spread between EOAN and RWE for first half of 2021

FIGURE 6.5. Zero trend new thresholds against Old thresholds for EOAN and RWE

FIGURE 6.6. Nonzero trend new thresholds against Old thresholds for EOAN and RWE

From figures 6.5 and 6.6, the short term pairs trading between EOAN and RWE yields a maximum of two trades without considering trend, while including the trend yields three. Apart from the lower return in Zeng's strategy upon taking out transaction cost, there are no significant differences in returns in this case.

6.2.3. Exxon Mobil(XOM) and Chevron(CVX)

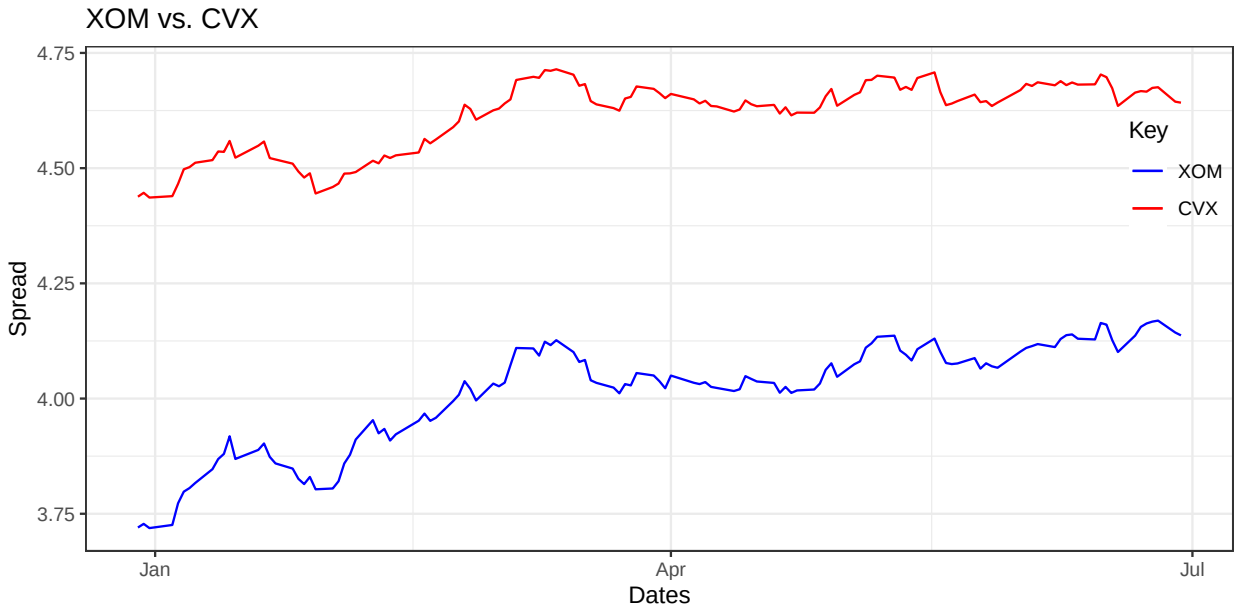

FIGURE 6.7. Logarithmic returns of XOM and CVX for first half of 2021

6.2.3.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	-2.5825843	0.0251620
	LS	-2.58257627	
λ	MLE	0.0504110	0.0354427
	LS	0.05042265	
σ	MLE	0.0118340	0.0007767
	LS	0.01187945	

TABLE 6.7. Parameter estimates for OU process representation of spread between XOM and CVX for first half of 2021

	Method	Estimate	Standard Error
a	MLE	0.0005872	0.0004374
	LS	0.0005870524	
b	MLE	-2.6326320	0.0355792
	LS	-2.6326459124	
λ	MLE	0.0706605	0.0393406
	LS	0.0706843357	
σ	MLE	0.0118702	0.0007846
	LS	0.0119189268	

TABLE 6.8. Parameter estimates for the trend-stationary OU process representation of spread between XOM and CVX for first half of 2021

6.2.3.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	-2.5457	-2.6195
Goncu	-2.5337	-2.6314
New case 1, no trend	-2.5826 + 0.0276 $e^{-0.0707(\frac{t}{252})}$	-2.5826 - 0.0276 $e^{-0.0707(\frac{t}{252})}$
New case 2, no trend	-2.5826 + 0.0714 $e^{-0.0707(\frac{t}{252})}$ - 0.0219 $e^{0.0707(\frac{t}{252})}$	-2.5826 - 0.0714 $e^{-0.0707(\frac{t}{252})}$ + 0.0219 $e^{0.0707(\frac{t}{252})}$
New case 1, with trend	-2.6326 + 0.0006 t + 0.0276 $e^{-0.0707(\frac{t}{252})}$	-2.6326 + 0.0006 t - 0.0276 $e^{-0.0707(\frac{t}{252})}$
New case 2, with trend	-2.6326 + 0.0006 t + 0.0714 $e^{-0.0707(\frac{t}{252})}$ - 0.0219 $e^{0.0707(\frac{t}{252})}$	-2.6326 + 0.0006 t - 0.0714 $e^{-0.0707(\frac{t}{252})}$ + 0.0219 $e^{0.0707(\frac{t}{252})}$

TABLE 6.9. Thresholds for the various trading strategies for the spread between XOM and CVX for first half of 2021

FIGURE 6.8. Zero trend new thresholds against Old thresholds for XOM and CVX

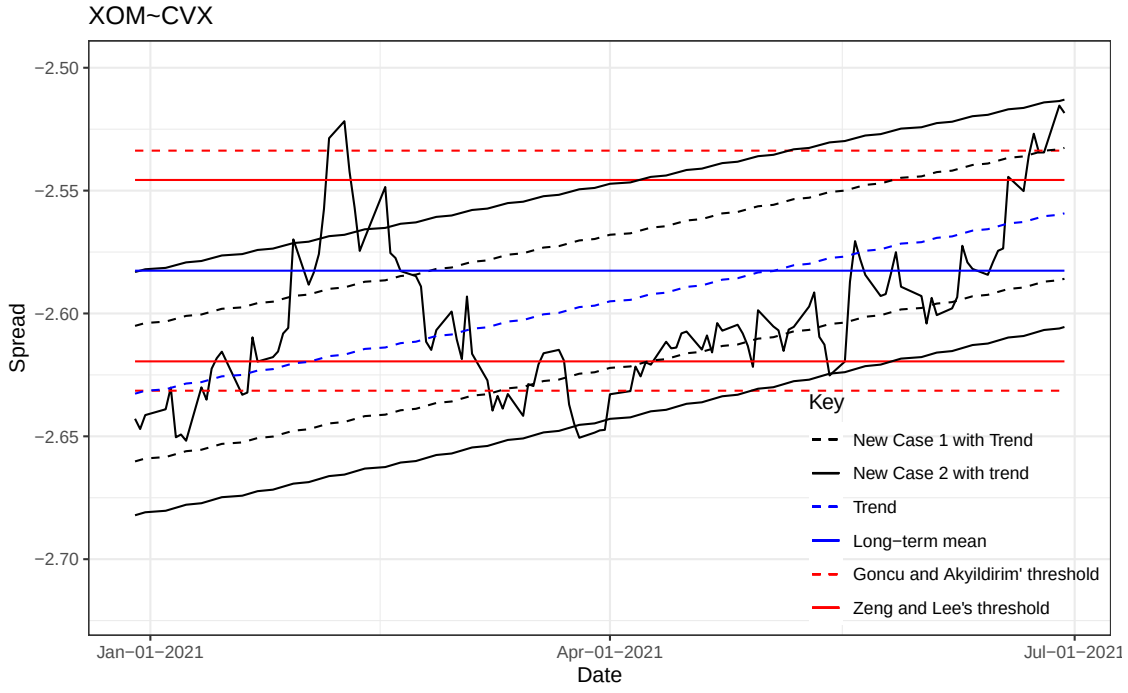

FIGURE 6.9. Nonzero trend new thresholds against Old thresholds for XOM and CVX

From figures 6.8 and 6.9, the short term pairs trading between Exxon mobile and Chevron yields four trades for all cases, except for case 2 of our new strategy, where we get only two trades. But the width of the band ensures that the strategy is still profitable and there is no significant difference in returns as shown in table 6.13.

6.2.4. Walmart(WMT) and Target(TGT)

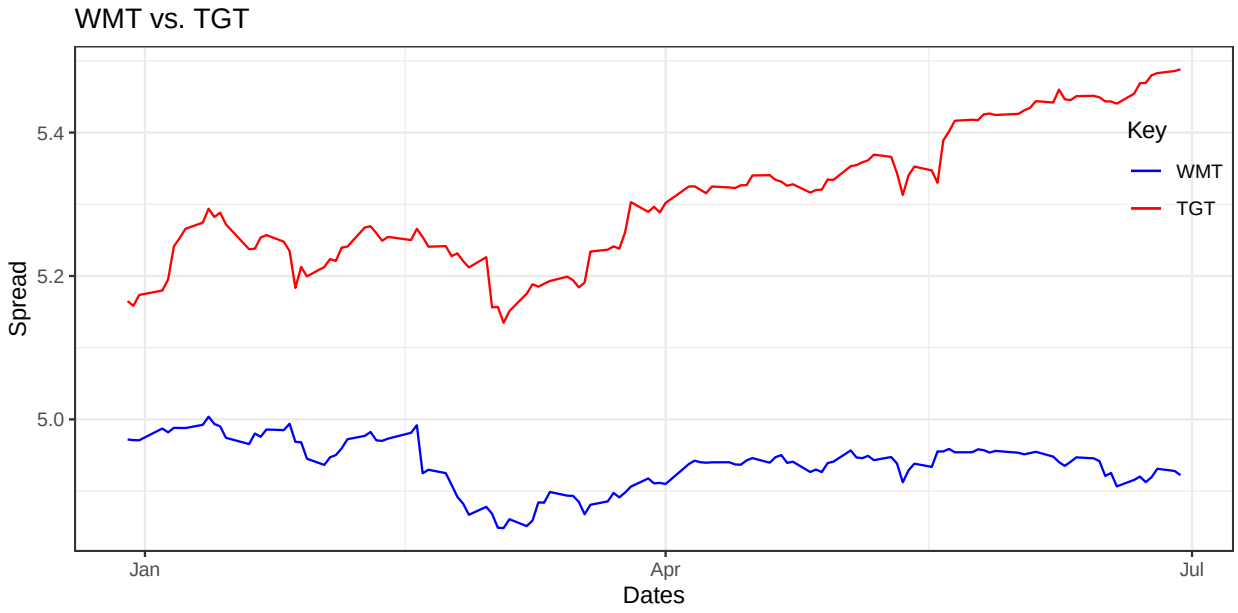

FIGURE 6.10. Logarithmic returns of WMT and TGT for first half of 2021

6.2.4.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	4.5494307	0.0184851
	LS	4.54951867	
λ	MLE	0.0569672	0.0299087
	LS	0.05709409	
σ	MLE	0.0112530	0.0007306
	LS	0.01129947	

TABLE 6.10. Parameter estimates for OU process representation of spread between WMT and TGT for first half of 2021

	Method	Estimate	Standard Error
a	MLE	-0.0002891	0.0004313
	LS	-0.0002889013	
b	MLE	4.5730214	0.0368662
	LS	4.5729951868	
λ	MLE	0.0653352	0.0329665
	LS	0.0653468594	
σ	MLE	0.0112824	0.0007365
	LS	0.0113280445	

TABLE 6.11. Parameter estimates for the trend-stationary OU process representation of spread between WMT and TGT for first half of 2021

6.2.4.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	4.5839	4.5150
Goncu	4.5931	4.5057
New case 1, no trend	4.5494 + 0.0273 $e^{-0.0653(\frac{t}{252})}$	4.5494 - 0.0273 $e^{-0.0653(\frac{t}{252})}$
New case 2, no trend	4.5494 + 0.0706 $e^{-0.0653(\frac{t}{252})}$ - 0.0217 $e^{0.0653(\frac{t}{252})}$	4.5494 - 0.0706 $e^{-0.0653(\frac{t}{252})}$ + 0.0217 $e^{0.0653(\frac{t}{252})}$
New case 1, with trend	4.5730 - 0.0003 t + 0.0273 $e^{-0.0653(\frac{t}{252})}$	4.5730 - 0.0003 t - 0.0273 $e^{-0.0653(\frac{t}{252})}$
New case 2, with trend	4.5730 - 0.0003 t + 0.0706 $e^{-0.0653(\frac{t}{252})}$ - 0.0217 $e^{0.0653(\frac{t}{252})}$	4.5730 - 0.0003 t - 0.0706 $e^{-0.0653(\frac{t}{252})}$ + 0.0217 $e^{0.0653(\frac{t}{252})}$

TABLE 6.12. Thresholds for the various trading strategies for the spread between WMT and TGT for first half of 2021

FIGURE 6.11. Zero trend new thresholds against Old thresholds for WMT and TGT

FIGURE 6.12. Nonzero trend new thresholds against Old thresholds for WMT and TGT

From figures 6.11 and 6.12, the short term pairs trading between Target and Walmart yields only two trades for all the old strategies, while our new strategy yields up to five trades for case 1 with trend, and we see significantly higher returns when the trend is considered. See table 6.13.

6.2.5. Performance Comparison

	Goncu	Zeng	Zeng-cost	New1	New2	New_Trended_1	New_Trended_2
PEP-KO	0.08	0.07	0.03	0.06	0.09	0.06	0.08
EOAN-RWE	0.17	0.10	0.06	0.10	0.10	0.12	0.21
XOM-CVX	0.10	0.07	-0.01	0.06	0.11	0.08	0.08
WMT-TGT	0.11	0.08	0.04	0.10	0.11	0.19	0.16

TABLE 6.13. Performance comparison of the strategies, by profits, for first half of 2021

Table 6.13 should be interpreted as returns made from shorting $1.00 worth of one of the stocks and using the amount to long the equivalent worth of the second stock in the pair.

6.3. Long-Term Trades

For the long term time horizon, we will use the last 1260 trading days data up until June 30th, 2021, which corresponds to five years in the US financial market.

We do not consider the strategy of [17] for long time horizons due to the fact that the thresholds become too far from the spreads and yields no trades.

6.3.1. Pepsi(PEP) and Coca-Cola(KO)

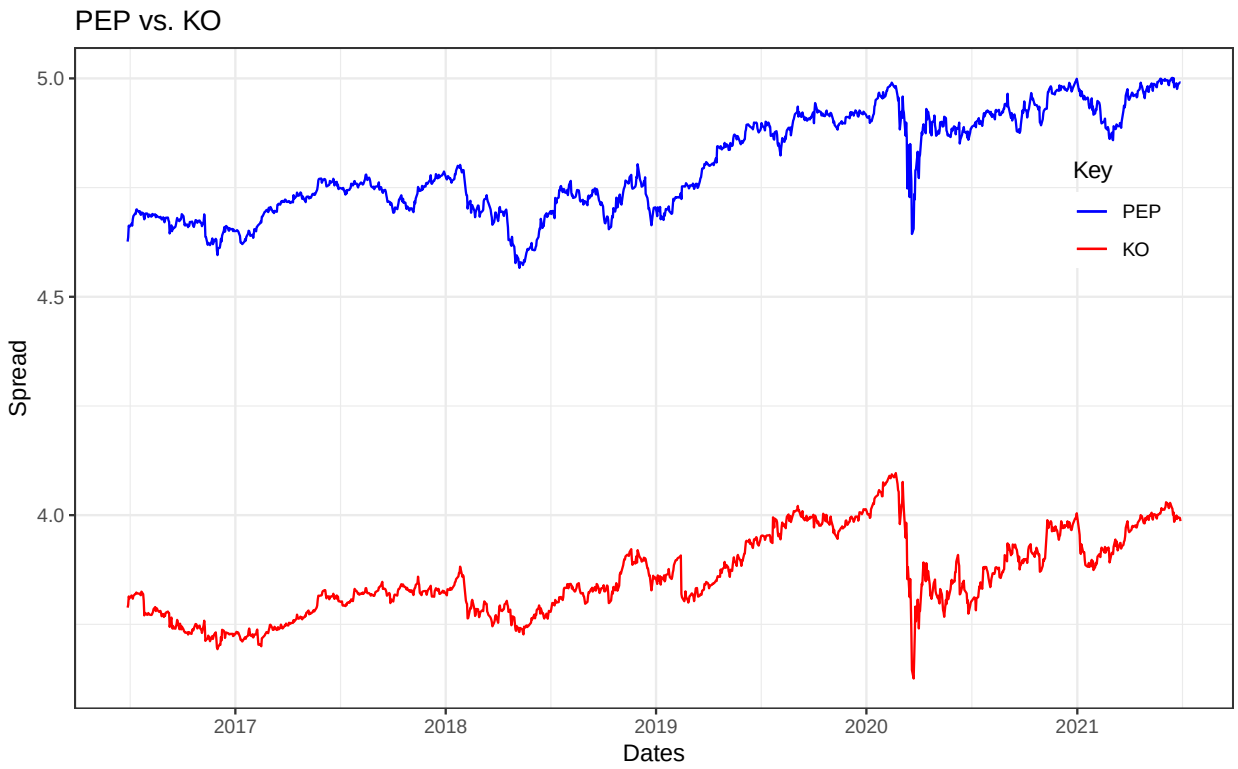

FIGURE 6.13. Logarithmic returns of PEP and KO from July, 2016 to June, 2021

6.3.1.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	0.7042167	0.0183045
	LS	0.704223932	
λ	MLE	0.0154326	0.0048586
	LS	0.015442115	
σ	MLE	0.0099308	0.0001994
	LS	0.009933762	

TABLE 6.14. Parameter estimates for OU process representation of spread between PEP and KO from July, 2016 to June, 2021

	Method	Estimate	Standard Error
a	MLE	0.00006230	0.00004115
	LS	0.00006245	
b	MLE	0.66040000	0.03183000
	LS	0.66025370	
λ	MLE	0.01873000	0.00543900
	LS	0.01874875	
σ	MLE	0.00993800	0.00019990
	LS	0.00994282	

TABLE 6.15. Parameter estimates for the trend-stationary OU process representation of spread between PEP and KO from July, 2016 to June, 2021

6.3.1.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	0.7521608	0.6562727
New case 1, no trend	0.7042 + 0.0499 $e^{-0.0187(\frac{t}{252})}$	0.7042 - 0.0499 $e^{-0.0187(\frac{t}{252})}$
New case 2, no trend	0.7042 + 0.0685 $e^{-0.0187(\frac{t}{252})}$ - 0.0095 $e^{0.0187(\frac{t}{252})}$	0.7042 - 0.0685 $e^{-0.0187(\frac{t}{252})}$ + 0.0095 $e^{0.0187(\frac{t}{252})}$
New case 1, with trend	0.6604 + 0.000062 t + 0.0499 $e^{-0.0187(\frac{t}{252})}$	0.6604 + 0.000062 t - 0.0499 $e^{-0.0187(\frac{t}{252})}$
New case 2, with trend	0.6604 + 0.000062 t + 0.0685 $e^{-0.0187(\frac{t}{252})}$ - 0.0095 $e^{0.0187(\frac{t}{252})}$	0.6604 + 0.000062 t - 0.0685 $e^{-0.0187(\frac{t}{252})}$ + 0.0095 $e^{0.0187(\frac{t}{252})}$

TABLE 6.16. Thresholds for the various trading strategies for the spread between PEP and KO from July, 2016 to June, 2021

FIGURE 6.14. Zero trend new thresholds against Old thresholds for PEP and KO

FIGURE 6.15. Nonzero trend new thresholds against Old thresholds for PEP and KO

From figures 6.14 and 6.15, we see that with considering trend, the long time horizon pairs trading between Pepsi and Coca cola yields only five trades over the five years, while with trend we obtain nine trades for case 1 threshold and eight for case 2 threshold. All our thresholds outperform Zeng's threshold, especially with trend.

6.3.2. E.OnSe (EOAN) and RWE AG(RWE) (German Utility Companies)

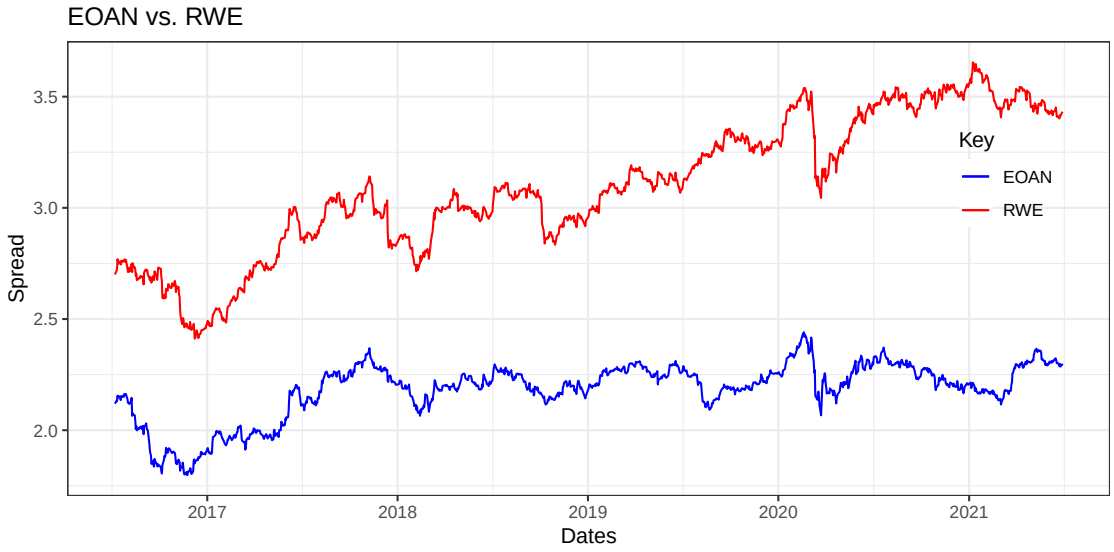

FIGURE 6.16. Logarithmic returns of EOAN and RWE from July, 2016 to June, 2021

6.3.2.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	1.2033039	0.0322562
	LS	1.20323964	
λ	MLE	0.0104885	0.0040641
	LS	0.01050636	
σ	MLE	0.0119887	0.0002402
	LS	0.01199341	

TABLE 6.17. Parameter estimates for OU process representation of spread between EOAN and RWE from July, 2016 to June, 2021

	Method	Estimate	Standard Error
a	MLE	0.0000044	0.00008919
	LS	0.0000047	
b	MLE	1.2000000	0.07236000
	LS	1.199820	
λ	MLE	0.01050000	0.00408700
	LS	0.01048860	
σ	MLE	0.01199000	0.00024030
	LS	0.01199329	

TABLE 6.18. Parameter estimates for the trend-stationary OU process representation of spread between EOAN and RWE from July, 2016 to June, 2021

6.3.2.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	1.264456	1.142152
New case 1, no trend	1.2033 + 0.0804 $e^{-0.0105(\frac{t}{252})}$	1.2033 - 0.0804 $e^{-0.0105(\frac{t}{252})}$
New case 2, no trend	1.2033 + 0.1104 $e^{-0.0105(\frac{t}{252})}$ - 0.0154 $e^{0.0105(\frac{t}{252})}$	1.2033 - 0.1104 $e^{-0.0105(\frac{t}{252})}$ + 0.0154 $e^{0.0105(\frac{t}{252})}$
New case 1, with trend	1.2000 + 0.000004 t + 0.0804 $e^{-0.0105(\frac{t}{252})}$	1.2000 + 0.000004 t - 0.0804 $e^{-0.0105(\frac{t}{252})}$
New case 2, with trend	1.2000 + 0.000004 t + 0.1104 $e^{-0.0105(\frac{t}{252})}$ - 0.0154 $e^{0.0105(\frac{t}{252})}$	1.2000 + 0.000004 t - 0.1104 $e^{-0.0105(\frac{t}{252})}$ + 0.0154 $e^{0.0105(\frac{t}{252})}$

TABLE 6.19. Thresholds for the various trading strategies for the spread between EOAN and RWE from July, 2016 to June, 2021

FIGURE 6.17. Zero trend new thresholds against Old thresholds for EOAN and RWE

FIGURE 6.18. Nonzero trend new thresholds against Old thresholds for EOAN and RWE

From figures 6.17 and 6.18, we notice there is no significant trend in the spread. However, table 6.26 shows that all our thresholds still outperform Zeng's threshold.

6.3.3. Exxon Mobil(XOM) and Chevron(CVX)

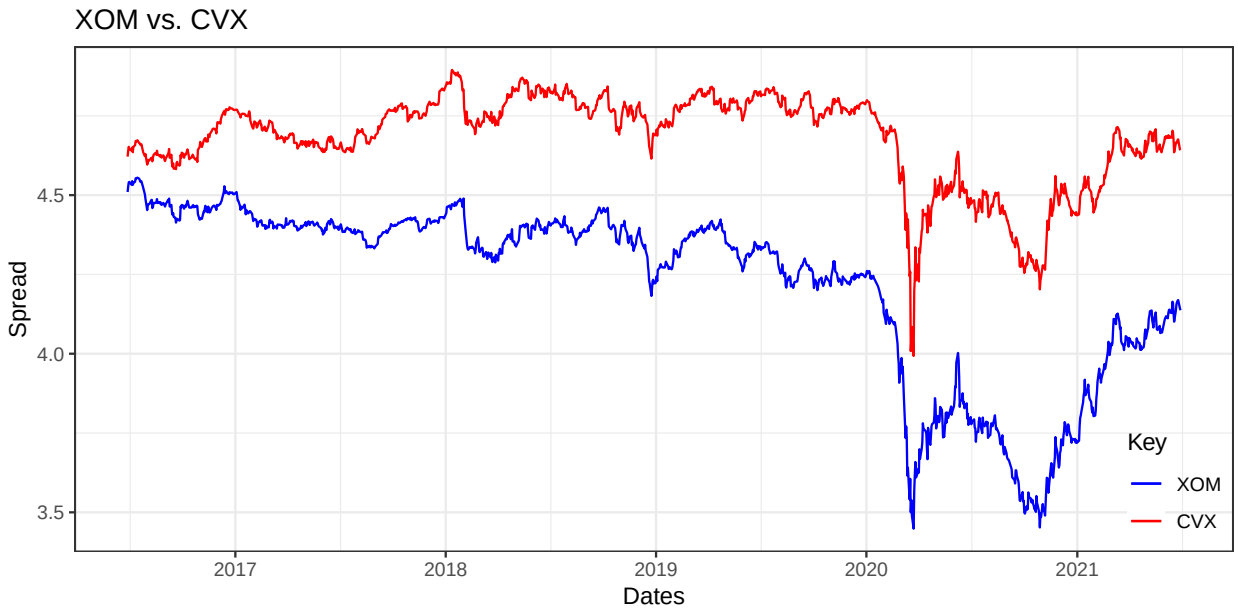

FIGURE 6.19. Logarithmic returns of XOM and CVX from July, 2016 to June, 2021

6.3.3.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	-3.2942037	0.0468415
	LS	-3.29412264	
λ	MLE	0.0128430	0.0040526
	LS	0.01284758	
σ	MLE	0.0210520	0.0004217
	LS	0.02106278	

TABLE 6.20. Parameter estimates for OU process representation of spread between XOM and CVX from July, 2016 to June, 2021

	Method	Estimate	Standard Error
a	MLE	-0.00029550	0.00005101
	LS	-0.0002953393	
b	MLE	-3.08400000	0.03824000
	LS	-3.0842283357	
λ	MLE	0.03283000	0.00740600
	LS	0.0328136899	
σ	MLE	0.02118000	0.00042920
	LS	0.0211814889	

TABLE 6.21. Parameter estimates for the trend-stationary OU process representation of spread between XOM and CVX from July, 2016 to June, 2021

6.3.3.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	-3.211814	-3.376593
New case 1, no trend	-3.2942 + 0.0803 $e^{-0.0462(\frac{t}{252})}$	-3.2942 - 0.0803 $e^{-0.0462(\frac{t}{252})}$
New case 2, no trend	-3.2942 + 0.1103 $e^{-0.0462(\frac{t}{252})}$ - 0.0153 $e^{0.0462(\frac{t}{252})}$	-3.2942 - 0.1103 $e^{-0.0462(\frac{t}{252})}$ + 0.0153 $e^{0.0462(\frac{t}{252})}$
New case 1, with trend	3.7117 - 0.0003 t + 0.0803 $e^{-0.0462(\frac{t}{252})}$	3.7117 - 0.0003 t - 0.0803 $e^{-0.0462(\frac{t}{252})}$
New case 2, with trend	3.7117 - 0.0003 t + 0.1103 $e^{-0.0462(\frac{t}{252})}$ - 0.0153 $e^{0.0462(\frac{t}{252})}$	3.7117 - 0.0003 t - 0.1103 $e^{-0.0462(\frac{t}{252})}$ + 0.0153 $e^{0.0462(\frac{t}{252})}$

TABLE 6.22. Thresholds for the various trading strategies for the spread between XOM and CVX from July, 2016 to June, 2021

FIGURE 6.20. Zero trend new thresholds against Old thresholds for XOM and CVX

FIGURE 6.21. Nonzero trend new thresholds against Old thresholds for XOM and CVX

We see here that the spread between Exxon mobile and Chevron has a significant trend. Thus we get significantly more trades when the trend is considered. See figure 6.20 and 6.21. All our thresholds still perform better than Zeng's threshold, and the trending thresholds perform even better, generating almost twice as much profit as Zeng's with transaction cost considered.

6.3.4. Walmart(WMT) and Target(TGT)

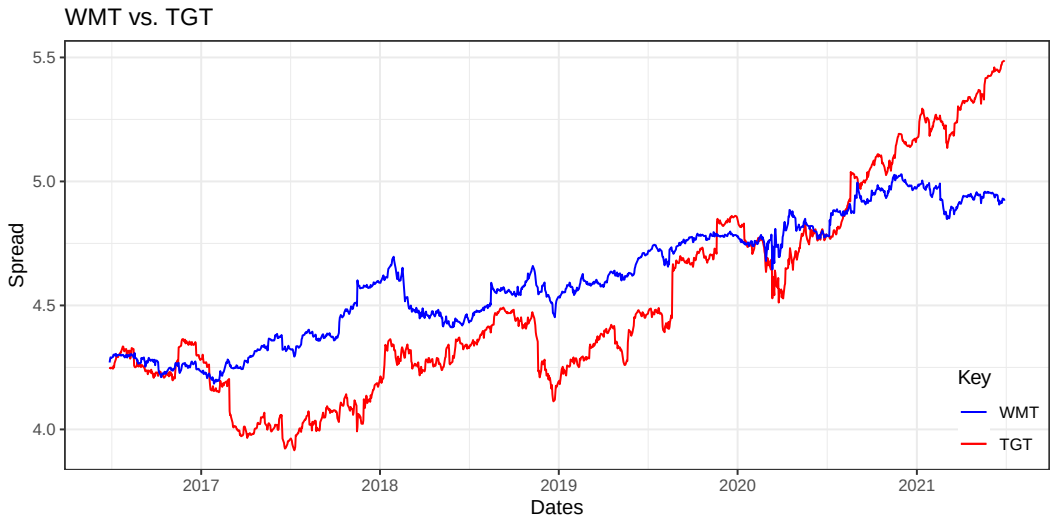

FIGURE 6.22. Logarithmic returns of WMT and TGT from July, 2016 to June, 2021

6.3.4.1. Parameter Estimate

	Method	Estimate	Standard Error
μ	MLE	2.2301072	0.0551505
	LS	2.229735861	
λ	MLE	0.0065499	0.0032483
	LS	0.006567753	
σ	MLE	0.0128164	0.0002563
	LS	0.012821651	

TABLE 6.23. Parameter estimates for OU process representation of spread between WMT and TGT from July, 2016 to June, 2021

	Method	Estimate	Standard Error
a	MLE	-0.0001508	0.0002448
	LS	-0.0001514578	
b	MLE	2.3518474	0.2201089
	LS	2.3523800692	
λ	MLE	0.0055067	0.0034584
	LS	0.0055053940	
σ	MLE	0.0128072	0.0002563
	LS	0.0128118140	

TABLE 6.24. Parameter estimates for the trend-stationary OU process representation of spread between WMT and TGT from July, 2016 to June, 2021

6.3.4.2. Thresholds

	Upper Threshold	Lower Threshold
Zeng	2.304404	2.15581
New case 1, no trend	2.2301 + 0.1186 $e^{-0.0055(\frac{t}{252})}$	2.2301 - 0.1186 $e^{-0.0055(\frac{t}{252})}$
New case 2, no trend	2.2301 + 0.1629 $e^{-0.0055(\frac{t}{252})}$ - 0.0227 $e^{0.0055(\frac{t}{252})}$	2.2301 - 0.1629 $e^{-0.0055(\frac{t}{252})}$ + 0.0227 $e^{0.0055(\frac{t}{252})}$
New case 1, with trend	2.3518 - 0.00015 t + 0.1186 $e^{-0.0055(\frac{t}{252})}$	2.3518 - 0.00015 t - 0.1186 $e^{-0.0055(\frac{t}{252})}$
New case 2, with trend	2.3518 - 0.00015 t + 0.1629 $e^{-0.0055(\frac{t}{252})}$ - 0.0227 $e^{0.0055(\frac{t}{252})}$	2.3518 - 0.00015 t - 0.1629 $e^{-0.0055(\frac{t}{252})}$ + 0.0227 $e^{0.0055(\frac{t}{252})}$

TABLE 6.25. Thresholds for the various trading strategies for the spread between WMT and TGT from July, 2016 to June, 2021

FIGURE 6.23. Zero trend new thresholds against Old thresholds for WMT and TGT

FIGURE 6.24. Nonzero trend new thresholds against Old thresholds for WMT and TGT

For the Target/Walmart pair, we notice that the trend in the spread switches from an increasing one to a decreasing trend between the first and second halves of of the time horizon, as shown in figures 6.23 and 6.24. Nonetheless, we notice from table 6.26 that our method is still superior to that of [40], with almost twice as much profit for without trend and almost trice as much for case 2 with trend, in comparison with [40] after deducting transaction cost.

6.3.5. Performance Comparison

	Zeng	Zeng-cost	New1	New2	New_Trended_1	New_Trended_2
PEP-KO	0.22	0.12	0.24	0.31	0.45	0.50
EOAN-RWE	0.71	0.45	0.78	0.75	0.78	0.74
XOM-CVX	0.79	0.59	0.82	0.94	1.06	1.01
WMT-TGT	0.48	0.34	0.59	0.62	0.53	0.94

TABLE 6.26. Performance comparison of the strategies, by profits, from July, 2016 to June, 2021

CHAPTER 7

CONCLUSION

We have looked at pairs trading as a form of statistical arbitrage, and have considered some of the resent trading strategies published in major quantitative finance journals. In particular we reviewed the strategies of Zeng and Lee [40], published as recent as 2014, and Goncu and Akyildirim [17], which was published in 2016. In our study of these two results we pointed out limitations of their strategies, such as overdependence on transaction cost, in the case of Zeng and Lee, which leads to a collapse of their strategy when transaction cost is zero, and on the other hand, the wide band of Goncu and Akyildirim's thresholds for long trading time horizons which yields zero trade under the circumstance and becomes inapplicable.

We presented a novel class of non-linear boundaries for the OU process, base on which we derived new decision rules for pairs trading. We proved the existence and uniqueness of maximizers for obtaining optimal thresholds for our new strategies. We also looked at versions of these boundaries for the trend-stationary OU process 12. Based on these boundaries, we came up with new strategies that presented four types of thresholds, two each for the mean reverting OU process and the trend-stationary OU process, and we named them case 1 and case 2, where case 1 has a tight band while case 2 is broader. We then considered both long-term and short term trade time horizons and showed with Monte Carlo simulations and real data for the pairs Pepsi(PEP)/Coca-cola(KO), Walmart(WMT)/Target(TGT), Exxon Mobil(XOM)/Chevron(CVX) and the German utility companies E.OnSe(EOAN)/RWE AG(RWE), that our new strategies are robust and outperform existing ones in terms of total profit over the trade time horizon, generating almost trice as much profit as Zeng and Lee's strategy in the case of long-term trade over a five-year period for the Walmart(WMT)/Target(TGT) pair, when transaction cost is taken into account. Our results showed that for both Long-term and short-term trade time horizons, considering trend yields higher profits.

In addition, we also derived simple formulas for some FPT moments of the standard

Brownian motion to some class of boundaries without relying on analytic expressions for the densities, since they are mostly nonexistent, and we extended this result to the Ornstein-Uhlenbeck process.

APPENDIX

SOME RELEVANT THEOREMS

A.1. Wang's Theorem

We consider the general diffusion process, X_t, satisfying the stochastic differential equation (SDE)

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dWt, \quad X_0 = x_0, \tag{82}$$

where the drift term $\mu(t, x) : [0, \infty) \times R \to R$ and the diffusion coefficient $\sigma(t, x) : [0, \infty) \times R \to R_+$ are real, deterministic functions and $\{W_t, t \geq 0\}$ is the standard Brownian motion (BM).

We define the boundary (non)crossing probability (BCP) of the process X_t to the boundaries $a(t)$ and $b(t)$ over the time interval $[0, T)$ by

$$P_X(a, b, T) = P(a(t) < W_t < b(t), \forall t \in [0, T])$$

THEOREM A.1. *If there exists a function $f(t, x) \in C^{1,2}([0, \infty) \times R)$, such that $Y_t := f(t, X_t)$ satisfies $dY_t = \tilde{\sigma}_t dW_t$, where $\tilde{\sigma}_t \in C([0, \infty))$ is a real, deterministic function satisfying $\tilde{\sigma}_t \neq 0, \forall t \in [0, \infty)$, then there exists a standard BM $\{\tilde{W}_s, s \geq 0\}$, such that for any boundaries $a(t), b(t)$,*

(1) $P_X(a, b, T) = P_{\tilde{W}}(c, d, S)$, if $\tilde{\sigma}_t > 0, \forall t \in [0, \infty)$;

(2) $P_X(a, b, T) = P_{\tilde{W}}(d, c, S)$, if $\tilde{\sigma}_t < 0, \forall t \in [0, \infty)$;

where,

$$c(s) = f(t(s), a(t(s))) - f(0, x_0), 0 \leq s \leq S,$$

$$d(s) = f(t(s), b(t(s))) - f(0, x_0), 0 \leq s \leq S,$$

$t(s)$ is the inverse function of $s(t) = \int_0^t \tilde{\sigma_u}^2 du$ and $S = s(T)$ [36].

A.2. Mill's Ratio Inequality

Given $x > 0$,

$$\frac{2}{\sqrt{x^2 + 4} + x} \leq R(x) \leq \frac{4}{\sqrt{x^2 + 8} + 3x},$$

where $R(x) = \frac{1-\Phi(x)}{\phi(x)}$ is the Mills ratio, with $\Phi(x)$ and $\phi(x)$ being the standard normal CDF and PDF respectively. [32], [39].

A.3. Ito's Lemma

Given a diffusion process X_t satisfying the stochastic differential equation:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t,$$

where B_t is the standard Brownian motion, Let $f(X_t, t) \in C^2(\mathbb{R}^2, \mathbb{R})$, then $f(X_t, t)$ satisfies

$$d(f(X_t, t)) = \left(\mu(X_t, t)\frac{\partial f(X_t, t)}{\partial t} + \frac{1}{2}\sigma^2(X_t, t)\frac{\partial^{f(X_t,t)}}{\partial X_t^2}(X_t, t)\right)dt + \sigma(X_t, t)\frac{\partial^{f(X_t,t)}}{\partial X_t^2}(X_t, t)dB_t$$

.

A.4. Girsanov's Theorem

Given the probability space $(\Omega, \mathbb{F}, P)$, let B_t be Brownian motion on the filtration $\{\mathbb{F}_t\}_{t\geq 0}$. Let μ_t be a process adapted to the filtration $\mathbb{F}_t$, and define a probability measure Q by

$$\left.\frac{dQ}{dP}\right|_{\mathbb{F}_t} = e^{-\frac{1}{2}\int_0^t \mu_s^2 ds + \int_0^t \mu_s dB_s} = Z_t,$$

in other words $Q(A) = E_p[Z_t 1_A]$, for $A \in \mathbb{F}_t$.

Then $Q \sim P$, and the process

$$W_t = B_t - \int_0^t \mu_s ds$$

is a Brownian motion under the probability measure Q [16] .

REFERENCES

[1] Milton Abramowitz and Irene Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, 1964.

[2] L. Alili and P. Patie, *Boundary-Crossing identities for diffusions having the time-inversion property*, Journal of Theoretical Probability 23 (2010), no. 1, 65–84.

[3] L. Alili, P. Patie, and J. L. Pedersen, *Representations of the first hitting time density of an ornstein-uhlenbeck process 1*, Stochastic Models 21 (2005), no. 4, 967–980.

[4] Marco Avellaneda and Jeong Hyun Lee, *Statistical arbitrage in the US equities market*, Quantitative Finance 10 (2010), no. 7, 761–782.

[5] William Karel Bertram, *Analytic Solutions for Optimal Statistical Arbitrage Trading*, SSRN Electronic Journal (2011), no. November 2009.

[6] Randall S. Billingsley, *Understanding Arbitrage An Intuitive Approach to Financial Analysis*, 2005.

[7] G E P Box and G C Tiao, *A canonical analysis of multiple time series Downloaded from*, Tech. Report 2, 1977.

[8] Leo Breiman, *First exit time from a square root boundary.*, Proc. 5th Berkley Symp. Math. Statist. Prob. 2, vol. 5, 1966, pp. 9–16.

[9] Michael J. Brennan and Eduardo S. Schwartz, *Savings bonds, retractable bonds and callable bonds*, Journal of Financial Economics 5 (1977), no. 1, 67–88.

[10] A. Buonocore, A. G. Nobile, and L. M. Ricciardi, *A new integral equation for the evaluation of first-passage-time probability densities*, Advances in Applied Probability 19 (1987), no. 4, 784–800.

[11] K C Chan, G Andrew Karolyi, Francis A Longstaff, Anthony B Sanders, Warren Bailey, Emilio Barone, Fischer Black, Tim Bollerslev, Stephen Buser, John Campbell, Jennifer Conrad, George Constantinides, Ken Dunn, Margaret Forster, Campbell Harvey, Patric Hendershott, David Mayers, Huston McCulloch, Daniel Nel-son, David Shimko, and

Stuart Turnbull, *An Empirical Comparison of Alternative Models of the Short-Term Interest Rate*, The Journal of Finance XLVII, (1992), no. 3.

[12] John C Cox, Jonathan E Ingersoll, and Stephen A Ross, *A Theory of the Term Structure of Interest Rates A THEORY OF THE TERM STRUCTURE OF INTEREST RATES1*, Tech. Report 2, 1985.

[13] D. A. Darling and A. J. F. Siegert, *The First Passage Problem for a Continuous Markov Process*, The Annals of Mathematical Statistics 24 (1953), no. 4, 624–639.

[14] Andrew N. Downes and Konstantin Borovkov, *First passage densities and boundary crossing probabilities for diffusion processes*, Methodology and Computing in Applied Probability 10 (2008), no. 4, 621–644.

[15] Jianqing Fan and Qiwei Yao, *The Elements of Financial Econometrics*, Cambridge University Press, 2017.

[16] I V Girsanov, *ON TRANSFORMING A CERTAIN CLASS OF STOCHASTIC PROCESSES BY ABSOLUTELY CONTINUOUS SUBSTITUTION OF MEASURES*, Theory of Probability and its Applications V (1960), no. 3.

[17] Ahmet Goncu and Erdinc Akyildirim, *Statistical Arbitrage with Pairs Trading*, International Review of Finance 16 (2016), no. 2, 1–11.

[18] Clive W. J. Granger, Granger, and Clive, *Some properties of time series data and their use in econometric model specification*, Journal of Econometrics 16 (1981), no. 1, 121–130.

[19] Piet Groeneboom, *Brownian motion with a parabolic drift and airy functions*, Probability Theory and Related Fields 81 (1989), no. 1, 79–109.

[20] Nicolas Huck and Komivi Afawubo, *Pairs trading and selection methods: is cointegration superior?*, Applied Economics 47 (2015), no. 6, 599–613.

[21] Peter G. Hufton, Elizabeth Buckingham-Jeffery, and Tobias Galla, *First-passage times and normal tissue complication probabilities in the limit of large populations*, Scientific Reports 10 (2020), no. 1, 1–12.

[22] Tomoyuki Ichiba and Constantinos Kardaras, *Efficient estimation of one-dimensional*

diffusion first passage time densities via monte carlo simulation, Journal of Applied Probability 48 (2011), no. 3, 699–712.

[23] Sebastian Jaimungal, Alex Kreinin, and Angelo Valov, *Integral Equations and the First Passage Time of Brownian Motions*, arXiv (2009), 1–31.

[24] Ioannis Karatzas and Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, second edi ed., Springer, 1998.

[25] Hans Rudolf Lerche, *Boundary Crossings of Brownian Motion*, vol. 40, Springer, 1986.

[26] Freddy H Marín Sánchez and Verónica M Gallego, *Parameter Estimation in Mean Reversion Processes with Deterministic Long-Term Trend*, (2016).

[27] James Nolen, *Stochastic partial differential equations and diffusion processes*, vol. 37, 1982.

[28] Alex Novikov, *On Stopping Times For A Wiener Process*, Theory of Probability and its Applications XVI (1971), no. 3, 449–456.

[29] Volf; Kordzakhia Nino Novikov, Alex; Frishling, *Approximations of Boundary Crossing Probabilities for a Brownian Motion*, Journal of Applied Probability 36 (1999), no. 4, 1019–1030.

[30] Debin Qiu, *Package 'aTSA' Type Package Title Alternative Time Series Analysis*, Tech. report, 2015.

[31] L. M. Ricciardi, L. Sacerdote, and S. Sato, *On an Integral Equation for First-Passage-Time Probability Densities.*, Journal of Applied Probability 21 (1984), no. 2, 302–314.

[32] L R Shenton, *INEQUALITIES FOR THE NORMAL INTEGRAL INCLUDING A NEW CONTINUED FRACTION*, Oxford Journals 41 (1954), no. 1/2, 177.

[33] L. A. Shepp, *A First Passage Problem for the Wiener Process*, The Annals of Mathematical Statistics 38 (1967), no. 6, 1912–1914.

[34] Christian Thierfelder, *The trending Ornstein-Uhlenbeck Process and its Applications in*, Hertford College, University of Oxford (2015), 1–93.

[35] Oldrich Vasicek, *An equilibrium characterization of the term structure*, Journal of Financial Economics 5 (1977), no. 2, 177–188.

[36] Liqun Wang and Klaus Pötzelberger, *Crossing probabilities for diffusion processes with piecewise continuous boundaries*, Methodology and Computing in Applied Probability 9 (2007), no. 1, 21–40.

[37] George H. Weiss, *First Passage Time Problems in Chemical Physics*, Advances in chemical physics, vol. 13, 2007, pp. 1–18.

[38] Inc. Wolfram Research, *Mathematica*, 2021.

[39] Zhen-Hang Yang and Yu-Ming Chu, *On approximating Mills ratio*, Journal of Inequalities and Applications 2015 (2015), 273.

[40] Zhengqin Zeng and Chi Guhn Lee, *Pairs trading: optimal thresholds and profitability*, Quantitative Finance 14 (2014), no. 11, 1881–1893.