An Optimal Strategy for Pairs Trading Under Geometric Brownian Motions

Jingzhi Tie, Hanqin Zhang & Qing Zhang

Journal of Optimization Theory and Applications

ISSN 0022-3239

J Optim Theory Appl
DOI 10.1007/s10957-017-1065-8

Springer

J Optim Theory Appl
DOI 10.1007/s10957-017-1065-8

An Optimal Strategy for Pairs Trading Under Geometric Brownian Motions

Jingzhi Tie[1] · **Hanqin Zhang**[2,3] · **Qing Zhang**[1]

Received: 7 November 2016 / Accepted: 18 January 2017

Abstract This paper is concerned with an optimal strategy for simultaneously trading of a pair of stocks. The idea of pairs trading is to monitor their price movements and compare their relative strength over time. A pairs trade is triggered by their prices divergence and consists of a pair of positions to short the strong stock and to long the weak one. Such a strategy bets on the reversal of their price strengths. From the viewpoint of technical tractability, typical pairs-trading models usually assume a difference of the stock prices satisfies a mean-reversion equation. In this paper, we consider the optimal pairs-trading problem by allowing the stock prices to follow general geometric Brownian motions. The objective is to trade the pairs over time to maximize an overall return with a fixed commission cost for each transaction. The optimal policy is characterized by threshold curves obtained by solving the associated HJB equations. Numerical examples are included to demonstrate the dependence of our trading rules on various parameters and to illustrate how to implement the results in practice.

Keywords Pairs trading · Optimal policy · Quasi-variational inequalities

✉ Qing Zhang
qz@uga.edu

Jingzhi Tie
jtie@uga.edu

Hanqin Zhang
bizzhq@nus.edu.sg

1 Department of Mathematics, University of Georgia, Athens, GA 30602, USA

2 Academy of Mathematics and Systems Sciences, Academia Sinica, Beijing 100190, China

3 School of Business, National University of Singapore, Singapore 119245, Singapore

Published online: 30 January 2017

Mathematics Subject Classification 93E20 · 91G80 · 49L20

1 Introduction

This paper is about an optimal policy for simultaneously trading of a pair of stocks. The idea of pairs trading is to track the price movements of these two securities over time and compare their relative price strengths. A pairs trade is triggered when their prices diverge, e.g., one stock moves up substantially relative to the other. A pairs trade is entered and consists of a short position in the stronger stock and a long position in the weaker one. Such a strategy bets on the reversal of their price strength. A major advantage of pairs trading is its 'market neutral' nature in the sense that it can be profitable under any market conditions.

Pairs trading was initially introduced by Bamberger and followed by Tartaglia's quantitative group at Morgan Stanley in the 1980s; see Gatev et al. [1] for related history and background details. There are many in-depth discussions in connection with the cause of the divergence and subsequent convergence; see the book by Vidyamurthy [2] and references therein. In addition to these studies, an advanced mathematical method was developed in Song and Zhang [3] to address issues in connection with pairs trading when the underlying pairs follow a mean-reversion model. It is shown in [3] that the optimal trading rule can be determined by threshold levels. These levels can be obtained by solving algebraic equations. A set of sufficient conditions are also provided to establish the desired optimality. One of the key assumptions in [3] is that the pairs value has to be a mean-reversion process. This clearly adds a severe limitation on its potential applications. In order to meet the mean-reversion requirement, tradable pairs are typically selected among stocks from the same industrial sector. From a practical viewpoint, it is highly desirable to have a broad range of stock selections for pairs trading. Mathematically speaking, this amounts to the possibility of extending the pairs-trading results in [3] under traditional stock price models such as geometric Brownian motions. To address the practical needs, in this paper, we develop a new method to treat the pairs-trading problem under general geometric Brownian motions.

By and large, the optimal timing of investments in irreversible projects can also be considered as a pairs-trading problem. Back in 1986, McDonald and Siegel [4] considered optimal timing of investment in an irreversible project. Two factors are included in their model: the growth of the investment capital and the change in project cost. Greater capital growth potential and lesser future project cost will postpone the transaction. Further studies along this line were carried out by Hu and Øksendal [5] to specify precise optimality conditions and to provide a new proof of the following variational inequalities among others. Their results can be easily interpreted in terms of pairs trading. It is simply a pairs-trading selling rule! This is, assuming an existing pairs position with a long position in one stock and a short position in another, the problem is to determine when to exit and close the position, i.e., to sell the long position and cover the short position. In this paper, we extend these results by allowing sequentially and simultaneously trading of these pairs. We focus on simple and easily implementable pairs-trading strategy and its optimality and closed-form solution. As expected, the value function incurred by the sequential decisions becomes more complicated comparing with that incurred by one-time decision (Hu and Øksendal [5]).

Mathematical portfolio selection and trading rules have been studied for many years. For example, Davis and Norman [6] studied Merton's investment/consumption problem with transaction costs and established wedge-shaped regions for the pair of bank and stock holdings. In-depth studies and a complete solution to this problem can be found in Shreve and Soner [7]. A basic assumption underlying these works is that a fraction of shares can be traded and the performance is evaluated via a hyperbolic absolute risk aversion utility function. A more realistic setting under a regime switching model was considered in Zhang [8] in connection with stock selling rule determined by two threshold levels, a target price and a stop-loss limit. In [8], such optimal threshold levels are obtained by solving a set of two-point boundary value problems. Guo and Zhang [9] studied the optimal selling rule under a model with switching Geometric Brownian motion. Using a smooth-fit technique, they obtained the optimal threshold levels by solving a set of algebraic equations. These papers are concerned with the selling side of trading in which the underlying price models are of GBM type. Recently, Dai et al. [10] developed a trend following rule based on a conditional probability indicator. They showed that the optimal trading rule can be determined by two threshold curves which can be obtained by solving the associated Hamilton–Jacobi–Bellman (HJB) equations. Similar idea was developed following a confidence interval approach by Iwarere and Barmish [11]. In addition, Merhi and Zervos [12] studied an investment capacity expansion/reduction problem following a dynamic programming approach under a geometric Brownian motion market model. Similar problem under a more general market model was treated by Løkka and Zervos [13]. In connection with mean-reversion trading, Zhang and Zhang [14] obtained a buy-low and sell-high policy by characterizing the 'low' and 'high' levels in terms of the mean-reversion parameters.

In this paper, we consider an optimal pairs-trading rule in which a pairs (long–short) position consists of a long position of one stock and a short position of the other. The objective is to initiate (buy) and close (sell) the pairs positions sequentially to maximize a discounted payoff function. Fixed proportional transaction (commission and/or slippage) costs will be imposed to each transaction. We study the problem following a dynamic programming approach and establish the associated HJB equations for the value functions. We show that the corresponding optimal stopping times can be determined by two threshold curves (lines with slopes k_1 and k_2). These key levels are given by the ratio of one-share long position to the one-share short position and can be obtained in closed form. We also examine the dependence of these threshold levels on various parameters in numerical examples. Finally, we demonstrate how to implement the results using a pair of stocks and their historical prices. To conclude, we highlight the main new features and contributions: (a) The typical mean-reversion requirement for pairs trading is dropped, and traditional geometric Brownian motion models are used to capture stock price movements. The one-time selling decision treated in [5] is generalized to a sequence of trading decisions. (b) A set of new smooth-fit conditions are provided by solving multi-variable partial differential equations, and new threshold curves are obtained rather than typical constant threshold levels. (c) A closed-form solution for the optimal pairs-trading problem is obtained.

This paper is organized as follows. In Sect. 2, we formulate the pairs-trading problem under consideration. In Sect. 3, we study basic properties of the value functions. In Sect. 4, we consider the associate HJB equations and their solutions. In Sect. 5, we present a verification theorem. Finally, a numerical example is given in Sect. 6 to illustrate the results.

2 Problem Formulation

We consider two stocks $\mathbf{S}^1$ and $\mathbf{S}^2$. Let $\{X_t^1, t \geq 0\}$ denote the prices of stock $\mathbf{S}^1$ and $\{X_t^2, t \geq 0\}$ that of stock $\mathbf{S}^2$. They satisfy the following stochastic differential equation:

$$d\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} X_t^1 & \\ & X_t^2 \end{pmatrix} \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} \right], \tag{1}$$

where μ_i, $i = 1, 2$, are the return rates, σ_{ij}, $i, j = 1, 2$, the volatility constants, and (W_t^1, W_t^2) a 2-dimensional standard Brownian motion.

In this paper, we consider a pairs-trading strategy. For simplicity, we assume the corresponding pairs position consists of one-share long position in stock $\mathbf{S}^1$ and one-share short position in stock $\mathbf{S}^2$. Also, the notation $\mathbf{S}^i$, $i = 1, 2$, are reserved for the underlying stocks and $\mathbf{Z}$ the corresponding pairs position. One share in $\mathbf{Z}$ means the combination of one-share long position in $\mathbf{S}^1$ and one-share short position in $\mathbf{S}^2$.

We consider the case that the net position at any time can be either long (with one share of $\mathbf{Z}$) or flat (no stock position of either $\mathbf{S}^1$ or $\mathbf{S}^2$). Let $i = 0, 1$ denote the initial net position, and let $\tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$ denote a sequence of stopping times. If initially the net position is long ($i = 1$), then one should sell $\mathbf{Z}$ before acquiring any future shares. That is, to first sell the pair at τ_0, then buy at τ_1, sell at τ_2, buy at τ_3, etc. The corresponding trading sequence is denoted by $\Lambda_1 = (\tau_0, \tau_1, \tau_2, \ldots)$. Likewise, if initially the net position is flat ($i = 0$), then one should start to buy a share of $\mathbf{Z}$. That is, to first buy at τ_1, sell at τ_2, then buy at τ_3, etc. The corresponding sequence of stopping times is denoted by $\Lambda_0 = (\tau_1, \tau_2, \ldots)$.

Let K denote the fixed percentage of transaction costs associated with buying or selling of stocks $\mathbf{S}^i$, $i = 1, 2$. For example, the cost to establish the pairs position $\mathbf{Z}$ at $t = t_1$ is $(1 + K)X_{t_1}^1 - (1 - K)X_{t_2}^2$ and the proceeds to close it at a later time $t = t_2$ is $(1 - K)X_{t_2}^1 - (1 + K)X_{t_2}^2$. For ease of notation, let $\beta_\mathrm{b} = 1 + K$ and $\beta_\mathrm{s} = 1 - K$.

Given the initial state (x_1, x_2), the initial net position $i = 0, 1$, and the decision sequences Λ_0 and Λ_1, the corresponding reward functions

$$\begin{aligned} J_0(x_1, x_2, \Lambda_0) = E\Big\{ & \Big[e^{-\rho\tau_2} \Big(\beta_\mathrm{s} X_{\tau_2}^1 - \beta_\mathrm{b} X_{\tau_2}^2\Big) I_{\{\tau_2 < \infty\}} - e^{-\rho\tau_1} \Big(\beta_\mathrm{b} X_{\tau_1}^1 - \beta_\mathrm{s} X_{\tau_1}^2\Big) I_{\{\tau_1 < \infty\}} \Big] \\ & + \Big[e^{-\rho\tau_4} \Big(\beta_\mathrm{s} X_{\tau_4}^1 - \beta_\mathrm{b} X_{\tau_4}^2\Big) I_{\{\tau_4 < \infty\}} - e^{-\rho\tau_3} \Big(\beta_\mathrm{b} X_{\tau_3}^1 - \beta_\mathrm{s} X_{\tau_3}^2\Big) I_{\{\tau_3 < \infty\}} \Big] + \cdots \Big\}, \\ J_1(x_1, x_2, \Lambda_1) = E\Big\{ & e^{-\rho\tau_0} \Big(\beta_\mathrm{s} X_{\tau_0}^1 - \beta_\mathrm{b} X_{\tau_0}^2\Big) I_{\{\tau_0 < \infty\}} \end{aligned}$$

$$+\left[e^{-\rho\tau_2}\left(\beta_s X^1_{\tau_2}-\beta_b X^2_{\tau_2}\right)I_{\{\tau_2<\infty\}}-e^{-\rho\tau_1}\left(\beta_b X^1_{\tau_1}-\beta_s X^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right]$$
$$+\left[e^{-\rho\tau_4}\left(\beta_s X^1_{\tau_4}-\beta_b X^2_{\tau_4}\right)I_{\{\tau_4<\infty\}}-e^{-\rho\tau_3}\left(\beta_b X^1_{\tau_3}-\beta_s X^2_{\tau_3}\right)I_{\{\tau_3<\infty\}}\right]+\cdots\Bigg\}, \quad (2)$$

where $\rho > 0$ is a given discount factor and I_A is the indicator function of an event A.

For $i = 0, 1$, let $V_i(x_1, x_2)$ denote the value functions with $(X^1_0, X^2_0) = (x_1, x_2)$ and initial net positions $i = 0, 1$. That is, $V_i(x_1, x_2) = \sup_{\Lambda_i} J_i(x_1, x_2, \Lambda_i)$, $i = 0, 1$.

Remark 2.1 Note that the 'one-share' assumption can be easily relaxed. For example, one can consider any pairs $\mathbf{Z}$ consisting of n_1 shares of long position in $\mathbf{S}^1$ and n_2 shares of short position in $\mathbf{S}^2$. This case can be treated by changing the state variables $(X^1_t, X^2_t) \to (n_1X^1_t, n_2X^2_t)$. Due to the nature of GBMs, the corresponding system equation in (1) will stay the same. The new allocations will only affect the reward function in (2) implicitly. In addition, we only focus on the 'long' side of pairs trading and note that the 'short' side of trading can also be treated by simply switching the roles of the two stocks $\mathbf{S}^1$ and $\mathbf{S}^2$.

Example 2.1 In this example, we consider stock prices of Target Corp. (TGT) and Wal-Mart Stores Inc. (WMT). In Fig. 1, daily closing prices of both stocks from 1985 to 2014 are plotted. The data are divided into two parts. The first part (1985–1999) will be used to calibrate the model and the second part (2000–2014) to backtest the performance of our results. Using the prices (1985–1999) and following the traditional

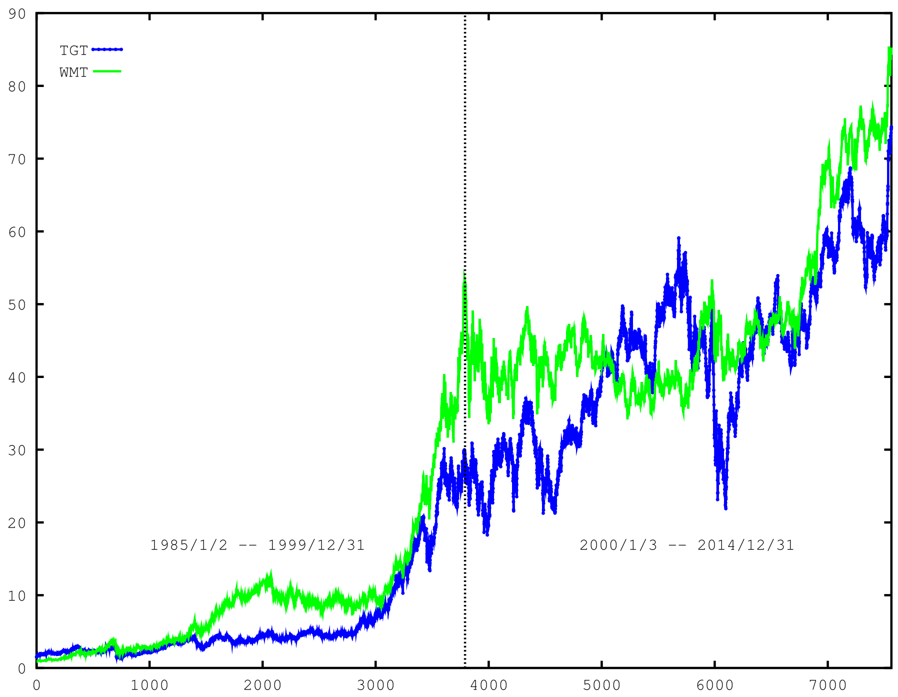

Fig. 1 Daily closing prices of TGT and WMT from 1985 to 2014

least squares method, we obtain $\mu_1 = 0.2059$, $\mu_2 = 0.2459$, $\sigma_{11} = 0.3112$, $\sigma_{12} = 0.0729$, $\sigma_{21} = 0.0729$, $\sigma_{22} = 0.2943$.

In this paper, we will assume that the discount factor ρ is greater than the return rates of each stock prices., i.e., $\rho > \mu_1$ and $\rho > \mu_2$.

Remark 2.2 Note that ρ serves as a combined discounting and risk aversion rate. The above inequalities are often imposed when treating related decision-making problems (see, e.g., Hu and Øksendal [5] and McDonald and Siegel [4]). They are also used to ensure that the corresponding value functions are finite. If some of these conditions are violated, e.g., taking $\mu_1 = 0.20$, $\mu_2 = 0.10$, and $\rho = 0.15$, it can be seen that, for all $(x_1, x_2) > 0$, $V_1(x_1, x_2) \geq Ee^{-0.15n}(\beta_s X_n^1 - \beta_b X_n^2) = \beta_s x_1 e^{0.05n} - \beta_b x_2 e^{-0.05n} \to \infty$, as $n \to \infty$.

3 Properties of the Value Functions

In this section, we establish basic properties of the value functions. First, for any given sequence of stopping times $\tau_0 \leq \tau_1 \leq \tau_2, \ldots$, let $\Lambda_1 = (\tau_0, \tau_1, \tau_2, \ldots)$ and $\Lambda_0 = (\tau_1, \tau_2, \ldots)$. Note that $J_1(x_1, x_2, \Lambda_1) = E\left[e^{-\rho\tau_0}\left(\beta_s X_{\tau_0}^1 - \beta_b X_{\tau_0}^2\right) I_{\{\tau_0<\infty\}}\right] + J_0(x_1, x_2, \Lambda_0)$. In particular, if $\tau_0 = 0$, a.s., then $J_1(x_1, x_2, \Lambda_1) = \beta_s x_1 - \beta_b x_2 + J_0(x_1, x_2, \Lambda_0)$. It follows that

$$V_1(x_1, x_2) \geq \beta_s x_1 - \beta_b x_2 + V_0(x_1, x_2). \tag{3}$$

Similarly, let $\Lambda_0 = (\tau_1, \tau_2, \ldots)$ and the subsequent $\Lambda_1 = (\tau_2, \ldots)$. Then, we have $J_0(x_1, x_2, \Lambda_0) = -E\left[e^{-\rho\tau_1}\left(\beta_b X_{\tau_1}^1 - \beta_s X_{\tau_1}^2\right) I_{\{\tau_1<\infty\}}\right] + J_1(x_1, x_2, \Lambda_1)$. Setting $\tau_1 = 0$, a.s., leads to

$$V_0(x_1, x_2) \geq -\beta_b x_1 + \beta_s x_2 + V_1(x_1, x_2). \tag{4}$$

Lemma 3.1 *For all $x_1, x_2 > 0$, we have*

$$0 \leq V_0(x_1, x_2) \leq x_2, \quad \textit{and} \quad \beta_s x_1 - \beta_b x_2 \leq V_1(x_1, x_2) \leq \beta_b x_1 + K x_2.$$

Proof Recall inequalities (3) and (4). It suffices to show the bounds for $V_0(x_1, x_2)$. It is clear that $V_0(x_1, x_2) \geq 0$ by definition and taking $\tau_1 = \infty$. To show $V_0(x_1, x_2) \leq x_2$, note that $\beta_b > 1$ and $\beta_s < 1$. It follows that

$$\begin{aligned} J_0(x_1, x_2, \Lambda_0) \leq E\Big\{&\Big[e^{-\rho\tau_2}\left(X_{\tau_2}^1 - X_{\tau_2}^2\right) I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}\left(X_{\tau_1}^1 - X_{\tau_1}^2\right) I_{\{\tau_1<\infty\}}\Big] \\ &+ \Big[e^{-\rho\tau_4}\left(X_{\tau_4}^1 - X_{\tau_4}^2\right) I_{\{\tau_4<\infty\}} - e^{-\rho\tau_3}\left(X_{\tau_3}^1 - X_{\tau_3}^2\right) I_{\{\tau_3<\infty\}}\Big] + \cdots\Big\}. \end{aligned}$$

Regroup the above terms to obtain

$$\begin{aligned}
&J_0(x_1, x_2, \Lambda_0)\\
&\le E\Big\{\Big(e^{-\rho\tau_2}X^1_{\tau_2}I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X^1_{\tau_1}I_{\{\tau_1<\infty\}}\Big) + \Big(e^{-\rho\tau_4}X^1_{\tau_4}I_{\{\tau_4<\infty\}}\\
&\quad - e^{-\rho\tau_3}X^1_{\tau_3}I_{\{\tau_3<\infty\}}\Big) + \cdots\Big\}\\
&\quad - E\Big\{\Big(e^{-\rho\tau_2}X^2_{\tau_2}I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X^2_{\tau_1}I_{\{\tau_1<\infty\}}\Big) + \Big(e^{-\rho\tau_4}X^2_{\tau_4}I_{\{\tau_4<\infty\}}\\
&\quad - e^{-\rho\tau_3}X^2_{\tau_3}I_{\{\tau_3<\infty\}}\Big) + \cdots\Big\}.
\end{aligned}$$

We first consider the term $E\left[(e^{-\rho\tau_2}X^1_{\tau_2}I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X^1_{\tau_1}I_{\{\tau_1<\infty\}}\right]$. For each $m = 1, 2, \ldots$ and $n = 1, 2, \ldots$, let $\tau^n_m = \tau_m \wedge n$. Recall that $\rho > \mu_1$. Then Dynkin's formula implies

$$E\left[e^{-\rho\tau^n_2}X^1_{\tau^n_2} - e^{-\rho\tau^n_1}X^1_{\tau^n_1}\right] = E\int_{\tau^n_1}^{\tau^n_2} e^{-\rho t}X^1_t(-\rho+\mu_1)\mathrm{d}t \le 0.$$

In addition, the uniform integrability of $\{e^{-\rho\tau^n_1}X^1_{\tau^n_1}\}$ can be proved by showing the existence of a $\gamma_0 > 1$ such that $\sup_n E\left(e^{-\rho\tau^n_1}X^1_{\tau^n_1}\right)^{\gamma_0} < \infty$. Sending $n \to \infty$, we have $Ee^{-\rho\tau^n_m}X^1_{\tau^n_m} \to Ee^{-\rho\tau_m}X^1_{\tau_m}I_{\{\tau_m<\infty\}}$, $m = 1, 2$. It follows that $E\left[e^{-\rho\tau_2}X^1_{\tau_2}I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X^1_{\tau_1}I_{\{\tau_1<\infty\}}\right] \le 0$. Repeat this on each term below to obtain

$$\begin{aligned}
&E\Big\{\Big(e^{-\rho\tau_2}X^1_{\tau_2}I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X^1_{\tau_1}I_{\{\tau_1<\infty\}}\Big) + \Big(e^{-\rho\tau_4}X^1_{\tau_4}I_{\{\tau_4<\infty\}}\\
&\quad - e^{-\rho\tau_3}X^1_{\tau_3}I_{\{\tau_3<\infty\}}\Big) + \cdots\Big\} \le 0.
\end{aligned}$$

Similarly, we can show, for each $m = 1, 2, \ldots$,

$$-E\left[e^{-\rho\tau_{m+1}}X^2_{\tau_{m+1}}I_{\{\tau_{m+1}<\infty\}} - e^{-\rho\tau_m}X^2_{\tau_m}I_{\{\tau_m<\infty\}}\right] = E\int_{\tau_m}^{\tau_{m+1}} e^{-\rho t}X^2_t(\rho-\mu_2)\mathrm{d}t,$$

by noticing the monotone convergence of $E\int_0^{\tau^n_m} e^{-\rho t}X^2_t(\rho-\mu_2)\mathrm{d}t$ to $E\int_0^{\tau_m} e^{-\rho t}X^2_t(\rho-\mu_2)\mathrm{d}t$ and, therefore, the convergence of $E\int_{\tau^n_m}^{\tau^n_{m+1}} e^{-\rho t}X^2_t(\rho-\mu_2)\mathrm{d}t$ to $E\int_{\tau_m}^{\tau_{m+1}} e^{-\rho t}X^2_t(\rho-\mu_2)\mathrm{d}t$ as $n \to \infty$. It follows that

$$
\begin{aligned}
&-E\Big\{\big(e^{-\rho\tau_2}X_{\tau_2}^2 I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}X_{\tau_1}^2 I_{\{\tau_1<\infty\}}\big) + \big(e^{-\rho\tau_4}X_{\tau_4}^2 I_{\{\tau_4<\infty\}} - e^{-\rho\tau_3}X_{\tau_3}^2 I_{\{\tau_3<\infty\}}\big) + \cdots\Big\} \\
&= -E\left\{\int_{\tau_1}^{\tau_2}(-\rho+\mu_2)e^{-\rho t}X_t^2 \mathrm{d}t + \int_{\tau_3}^{\tau_4}(-\rho+\mu_2)e^{-\rho t}X_t^2 \mathrm{d}t + \cdots\right\} \\
&= E\left\{\int_{\tau_1}^{\tau_2}(\rho-\mu_2)e^{-\rho t}X_t^2 \mathrm{d}t + \int_{\tau_3}^{\tau_4}(\rho-\mu_2)e^{-\rho t}X_t^2 \mathrm{d}t + \cdots\right\} \\
&\le (\rho-\mu_2)E\int_0^\infty e^{-\rho t}X_t^2 \mathrm{d}t = (\rho-\mu_2)\int_0^\infty e^{-\rho t}(x_2 e^{\mu_2 t})\mathrm{d}t = x_2.
\end{aligned}
$$

□

4 HJB Equations

In this section, we study the associated HJB equations. Let

$$
\mathcal{A} = \frac{1}{2}\left\{a_{11}x_1^2\frac{\partial^2}{\partial x_1^2} + 2a_{12}x_1x_2\frac{\partial^2}{\partial x_1\partial x_2} + a_{22}x_2^2\frac{\partial^2}{\partial x_2^2}\right\} + \mu_1 x_1\frac{\partial}{\partial x_1} + \mu_2 x_2\frac{\partial}{\partial x_2},
$$

where $a_{11} = \sigma_{11}^2 + \sigma_{12}^2$, $a_{12} = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}$, and $a_{22} = \sigma_{21}^2 + \sigma_{22}^2$. Formally, the associated HJB equations have the form: For $x_1, x_2 > 0$,

$$
\begin{aligned}
&\min\Big\{\rho v_0(x_1,x_2) - \mathcal{A}v_0(x_1,x_2),\ v_0(x_1,x_2) - v_1(x_1,x_2) + \beta_{\mathrm{b}}x_1 - \beta_{\mathrm{s}}x_2\Big\} = 0, \\
&\min\Big\{\rho v_1(x_1,x_2) - \mathcal{A}v_1(x_1,x_2),\ v_1(x_1,x_2) - v_0(x_1,x_2) - \beta_{\mathrm{s}}x_1 + \beta_{\mathrm{b}}x_2\Big\} = 0.
\end{aligned} \tag{5}
$$

To solve the above HJB equations, we first convert them into single-variable equations. Let $y = x_2/x_1$ and $v_i(x_1,x_2) = x_1 w_i(x_2/x_1)$, for some function $w_i(y)$ and $i = 0, 1$. Then we have by direct calculation that

$$
\begin{aligned}
&\frac{\partial v_i}{\partial x_1} = w_i(y) - y w_i'(y),\quad \frac{\partial v_i}{\partial x_2} = w_i'(y), \\
&\frac{\partial^2 v_i}{\partial x_1^2} = \frac{y^2 w_i''(y)}{x_1},\quad \frac{\partial^2 v_i}{\partial x_2^2} = \frac{w_i''(y)}{x_1},\ \text{and}\ \frac{\partial^2 v_1}{\partial x_1\partial x_2} = -\frac{y w_i''(y)}{x_1}.
\end{aligned}
$$

Write $\mathcal{A}v_i$ in terms of w_i to obtain

$$
\mathcal{A}v_i = x_1\left\{\frac{1}{2}[a_{11} - 2a_{12} + a_{22}]\, y^2 w_i''(y) + (\mu_2-\mu_1) y w_i'(y) + \mu_1 w_i(y)\right\}.
$$

Then, the HJB equations can be given in terms of y and w_i as follows:

$$
\begin{aligned}
&\min\Big\{\rho w_0(y) - \mathcal{L}w_0(y),\ w_0(y) - w_1(y) + \beta_{\mathrm{b}} - \beta_{\mathrm{s}}y\Big\} = 0, \\
&\min\Big\{\rho w_1(y) - \mathcal{L}w_1(y),\ w_1(y) - w_0(y) - \beta_{\mathrm{s}} + \beta_{\mathrm{b}}y\Big\} = 0,
\end{aligned} \tag{6}
$$

where $\mathcal{L}[w_i(y)] = \lambda y^2 w_i''(y) + (\mu_2-\mu_1) y w_i'(y) + \mu_1 w_i(y)$ and $\lambda = (a_{11} - 2a_{12} + a_{22})/2$. In this paper, we only consider the case when $\lambda \neq 0$. If $\lambda = 0$, the problem

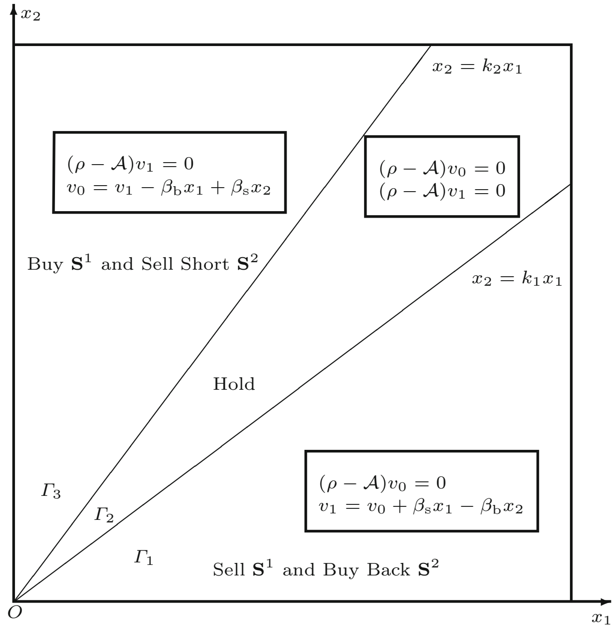

Fig. 2 Regions Γ_1, Γ_2, and Γ_3

reduces to a first-order case and can be treated in a similar way. To solve the above HJB equations, we first consider the equations $(\rho - \mathcal{L})w_i(y) = 0$, $i = 0, 1$. Clearly, these are the Euler equations and their solutions are of the form y^δ, for some δ. Substitute this into the equation $(\rho - \mathcal{L})w_i = 0$ to obtain $\delta^2 - (1 + (\mu_1 - \mu_2)/\lambda)\,\delta - (\rho - \mu_1)/\lambda = 0$. There are two real roots δ_1 and δ_2 (by direct calculation $\delta_1 > 1$ and $\delta_2 < 0$) given by

$$\begin{aligned} \delta_1 &= \frac{1}{2}\left(1 + \frac{\mu_1 - \mu_2}{\lambda} + \sqrt{\left(1 + \frac{\mu_1 - \mu_2}{\lambda}\right)^2 + \frac{4\rho - 4\mu_1}{\lambda}}\right) > 1, \\ \delta_2 &= \frac{1}{2}\left(1 + \frac{\mu_1 - \mu_2}{\lambda} - \sqrt{\left(1 + \frac{\mu_1 - \mu_2}{\lambda}\right)^2 + \frac{4\rho - 4\mu_1}{\lambda}}\right) < 0. \end{aligned} \tag{7}$$

Therefore, the general solution of $(\rho - \mathcal{L})w_i(y) = 0$ should be of the form: $w_i(y) = c_{i1} y^{\delta_1} + c_{i2} y^{\delta_2}$, for some constants c_{i1} and c_{i2}, $i = 1, 2$.

Intuitively, if X_t^1 is small and X_t^2 is large, then one should buy $\mathbf{S}^1$ and sell (short) $\mathbf{S}^2$. That is to open a pairs position $\mathbf{Z}$. If, on the other hand, X_t^1 is large and X_t^2 is small, then one should close the position $\mathbf{Z}$ by selling $\mathbf{S}^1$ and buying back $\mathbf{S}^2$. In view of this, we divide the first quadrant $P = \{(x_1, x_2) : x_1 > 0 \text{ and } x_2 > 0\}$ into three regions Γ_1, Γ_2, and Γ_3 where $\Gamma_1 = \{(x_1, x_2) \in P : x_2 \le k_1 x_1\}$, $\Gamma_2 = \{(x_1, x_2) \in P : k_1 x_1 < x_2 < k_2 x_1\}$, and $\Gamma_3 = \{(x_1, x_2) \in P : x_2 \ge k_2 x_1\}$. This is illustrated in Fig. 2.

Recall $y = x_2/x_1$. With a little bit abuse of notation, we write the corresponding $\Gamma_i, i = 1, 2, 3$, in terms of y: $\Gamma_1 = \{y : 0 < y \le k_1\}$, $\Gamma_2 = \{y : k_1 < y < k_2\}$, and $\Gamma_3 = \{y : y \ge k_2\}$. Here $0 < k_1 < k_2$ are thresholds to be determined so that on

$$\begin{array}{ll} \Gamma_1 : (\rho - \mathcal{L})w_0 = 0, & w_1 = w_0 + \beta_\mathrm{s} - \beta_\mathrm{b} y; \\ \Gamma_2 : (\rho - \mathcal{L})w_0 = 0, & (\rho - \mathcal{L})w_1 = 0; \\ \Gamma_3 : w_0 = w_1 - \beta_\mathrm{b} + \beta_\mathrm{s} y, & (\rho - \mathcal{L})w_1 = 0. \end{array} \tag{8}$$

In view of Lemma 3.1, the value functions have to be bounded near the origin. Recall that $\delta_2 < 0$. In order to have bounded value function w_0 on Γ_1, the coefficient of the term y^{δ_2} in its general form has to be zero. Therefore, $w_0 = C_0 y^{\delta_1}$ for some constant C_0 on Γ_1. Likewise, on Γ_3, the coefficient of y^{δ_1} must be zero because $\delta_1 > 1$. The solution $w_1 = C_1 y^{\delta_2}$ for some C_1 on Γ_3. Finally, these functions are extended to Γ_2 and are given by $w_0 = C_0 y^{\delta_1}$ and $w_1 = C_1 y^{\delta_2}$. Therefore, the solutions on each region should have the form:

$$\begin{array}{ll} \Gamma_1 : w_0 = C_0 y^{\delta_1}, & w_1 = C_0 y^{\delta_1} + \beta_\mathrm{s} - \beta_\mathrm{b} y; \\ \Gamma_2 : w_0 = C_0 y^{\delta_1}, & w_1 = C_1 y^{\delta_2}; \\ \Gamma_3 : w_0 = C_1 y^{\delta_2} - \beta_\mathrm{b} + \beta_\mathrm{s} y, & w_1 = C_1 y^{\delta_2}. \end{array}$$

Remark 4.1 Note that the assumptions $\rho > \mu_1$ and $\rho > \mu_2$ play a key role in the above analysis. In order to eliminate some constants in the general solutions w_0 and w_1, one needs $\delta_1 > 1$ and $\delta_2 < 0$. It can be shown by direct computation that these inequalities hold iff both $\rho > \mu_1$ and $\rho > \mu_2$ are satisfied.

Smooth-fit conditions. Next we develop smooth-fit conditions and determine the values for parameters: k_1, k_2, C_0, and C_1. In particular, we are to find C^1 solutions on the entire region $\{y > 0\}$. Necessarily, the continuity of w_1 and its first-order derivative at $y = k_1$ imply $C_1 k_1^{\delta_2} = C_0 k_1^{\delta_1} + \beta_\mathrm{s} - \beta_\mathrm{b} k_1$ and $C_1 \delta_2 k_1^{\delta_2 - 1} = C_0 \delta_1 k_1^{\delta_1 - 1} - \beta_\mathrm{b}$. We write them in matrix form:

$$\begin{pmatrix} k_1^{\delta_1} & -k_1^{\delta_2} \\ \delta_1 k_1^{\delta_1 - 1} & -\delta_2 k_1^{\delta_2 - 1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_\mathrm{b} k_1 - \beta_\mathrm{s} \\ \beta_\mathrm{b} \end{pmatrix}. \tag{9}$$

Similarly, the smooth-fit conditions for w_0 at $y = k_2$ yield the equations:

$$\begin{pmatrix} k_2^{\delta_1} & -k_2^{\delta_2} \\ \delta_1 k_2^{\delta_1 - 1} & -\delta_2 k_2^{\delta_2 - 1} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \begin{pmatrix} \beta_\mathrm{s} k_2 - \beta_\mathrm{b} \\ \beta_\mathrm{s} \end{pmatrix}. \tag{10}$$

We can solve for C_0 and C_1 and express the corresponding inverse matrices in terms of k_1 and k_2 to obtain

$$\begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = \frac{1}{\delta_1 - \delta_2} \begin{pmatrix} \beta_\mathrm{b}(1 - \delta_2) k_1^{1 - \delta_1} + \beta_\mathrm{s} \delta_2 k_1^{-\delta_1} \\ \beta_\mathrm{b}(1 - \delta_1) k_1^{1 - \delta_2} + \beta_\mathrm{s} \delta_1 k_1^{-\delta_2} \end{pmatrix}$$

$$= \frac{1}{\delta_1 - \delta_2} \begin{pmatrix} \beta_s(1-\delta_2)k_2^{1-\delta_1} + \beta_b\delta_2 k_2^{-\delta_1} \\ \beta_s(1-\delta_1)k_2^{1-\delta_2} + \beta_b\delta_1 k_2^{-\delta_2} \end{pmatrix}. \tag{11}$$

The second equality yields two equations of k_1 and k_2. Simplify them and write

$$\begin{aligned} (1-\delta_2)\left(\beta_b k_1^{1-\delta_1} - \beta_s k_2^{1-\delta_1}\right) &= \delta_2\left(\beta_b k_2^{-\delta_1} - \beta_s k_1^{-\delta_1}\right), \\ (1-\delta_1)\left(\beta_b k_1^{1-\delta_2} - \beta_s k_2^{1-\delta_2}\right) &= \delta_1\left(\beta_b k_2^{-\delta_2} - \beta_s k_1^{-\delta_2}\right). \end{aligned}$$

Let $r = k_2/k_1$. Replace k_2 by rk_1 to obtain $(1-\delta_2)(\beta_b - \beta_s r^{1-\delta_1})k_1 = \delta_2(\beta_b r^{-\delta_1} - \beta_s)$ and $(1-\delta_1)(\beta_b - \beta_s r^{1-\delta_2})k_1 = \delta_1(\beta_b r^{-\delta_2} - \beta_s)$. It follows that

$$\begin{aligned} k_1 &= \delta_2(\beta_b r^{-\delta_1} - \beta_s)/[(1-\delta_2)(\beta_b - \beta_s r^{1-\delta_1})] \\ &= \delta_1(\beta_b r^{-\delta_2} - \beta_s)/[(1-\delta_1)(\beta_b - \beta_s r^{1-\delta_2})]. \end{aligned}$$

The second equality yields an equation in terms of r: $\delta_2(\beta_b r^{-\delta_1} - \beta_s)/[(1-\delta_2)(\beta_b - \beta_s r^{1-\delta_1})] = \delta_1(\beta_b r^{-\delta_2} - \beta_s)/[(1-\delta_1)(\beta_b - \beta_s r^{1-\delta_2})]$. To show the existence of solution r_0, we let $\beta = \beta_b/\beta_s (>1)$ and

$$\begin{aligned} f(r) = \delta_1(1-\delta_2)&(\beta_b r^{-\delta_2} - \beta_s)(\beta_b - \beta_s r^{1-\delta_1}) \\ &-\delta_2(1-\delta_1)(\beta_b r^{-\delta_1} - \beta_s)(\beta_b - \beta_s r^{1-\delta_2}). \end{aligned}$$

Then we can show $f(\beta^2) = (\delta_1 - \delta_2)\beta_s^2\beta(\beta^{1-2\delta_2} - 1)(1 - \beta^{1-2\delta_1}) > 0$ and $f(r) \approx -(1-\delta_1)\delta_2\beta_s^2 r^{1-\delta_2} \to -\infty$ as $r \to \infty$ by taking the leading terms in $f(r)$. Therefore, there exists $r_0 > \beta^2$ so that $f(r_0) = 0$. Using this r_0, we can write k_1 and k_2:

$$\begin{aligned} k_1 &= \frac{\delta_2(\beta_b r_0^{-\delta_1} - \beta_s)}{(1-\delta_2)(\beta_b - \beta_s r_0^{1-\delta_1})} = \frac{\delta_1(\beta_b r_0^{-\delta_2} - \beta_s)}{(1-\delta_1)(\beta_b - \beta_s r_0^{1-\delta_2})}, \\ k_2 &= \frac{\delta_2(\beta_b r_0^{1-\delta_1} - \beta_s r_0)}{(1-\delta_2)(\beta_b - \beta_s r_0^{1-\delta_1})} = \frac{\delta_1(\beta_b r_0^{1-\delta_2} - \beta_s r_0)}{(1-\delta_1)(\beta_b - \beta_s r_0^{1-\delta_2})}. \end{aligned} \tag{12}$$

Finally, we can use these k_1 and k_2 to express C_0 and C_1 given in (11).

Theorem 4.1 *Let δ_i be given by (7) and k_i be given by (12). Then, the following functions w_0 and w_1 satisfy the HJB Eq. (6):*

$$w_0(y) = \begin{cases} \left(\dfrac{\beta_b(1-\delta_2)k_1^{1-\delta_1} + \beta_s\delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}\right) y^{\delta_1}, & \textit{if } 0 < y < k_2, \\ \left(\dfrac{\beta_b(1-\delta_1)k_1^{1-\delta_2} + \beta_s\delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}\right) y^{\delta_2} + \beta_s y - \beta_b, & \textit{if } y \ge k_2, \end{cases}$$

$$w_1(y) = \begin{cases} \left(\dfrac{\beta_b(1-\delta_2)k_1^{1-\delta_1} + \beta_s\delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}\right) y^{\delta_1} + \beta_s - \beta_b y, & \textit{if } 0 < y \le k_1, \\ \left(\dfrac{\beta_b(1-\delta_1)k_1^{1-\delta_2} + \beta_s\delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}\right) y^{\delta_2}, & \textit{if } y > k_1. \end{cases}$$

In view of this theorem, $v_0(x_1, x_2) = x_1 w_0(x_2/x_1)$ and $v_1(x_1, x_2) = x_1 w_1(x_2/x_1)$ satisfy the original HJB Eq. (5). Next, we first show a lemma needed in the proof.

Lemma 4.1 *Let k_1 and k_2 be given as in (12) with $r = r_0$. Then*

$$k_1 < \frac{\beta_s(\rho-\mu_1)}{\beta_b(\rho-\mu_2)} \quad \textit{and} \quad k_2 > \frac{\beta_b(\rho-\mu_1)}{\beta_s(\rho-\mu_2)}. \tag{13}$$

Proof First, we note that $\delta_1 + \delta_2 = 1 + (\mu_1 - \mu_2)/\lambda$ and $\delta_1\delta_2 = -(\rho - \mu_1)/\lambda$. It follows that

$$\frac{\rho-\mu_1}{\rho-\mu_2} = \frac{\delta_1(-\delta_2)}{(\delta_1-1)(1-\delta_2)}. \tag{14}$$

Use this equality and recall that $r_0 > \beta^2$ to obtain

$$\begin{aligned}
k_1 < \frac{\beta_s(\rho-\mu_1)}{\beta_b(\rho-\mu_2)} &\Longleftrightarrow \frac{\delta_2\left(\beta_{\mathrm{b}} r_0^{-\delta_1} - \beta_{\mathrm{s}}\right)}{(1-\delta_2)\left(\beta_{\mathrm{b}} - \beta_{\mathrm{s}} r_0^{1-\delta_1}\right)} < \frac{\delta_1(-\delta_2)}{(\delta_1-1)(1-\delta_2)} \cdot \frac{\beta_{\mathrm{s}}}{\beta_{\mathrm{b}}} \\
&\Longleftrightarrow \frac{\beta_{\mathrm{s}} - \beta_{\mathrm{b}} r_0^{-\delta_1}}{\beta_{\mathrm{b}} - \beta_{\mathrm{s}} r_0^{1-\delta_1}} < \frac{\delta_1}{\delta_1 - 1} \cdot \frac{\beta_{\mathrm{s}}}{\beta_{\mathrm{b}}} \quad (\because \delta_1 > 1,\ \delta_2 < 0) \\
&\Longleftrightarrow \frac{1 - \beta r_0^{-\delta_1}}{\beta - r_0^{1-\delta_1}} < \frac{\delta_1}{\delta_1 - 1} \cdot \frac{1}{\beta} \quad (\because \beta = \beta_{\mathrm{b}}/\beta_{\mathrm{s}} > 1) \\
&\Longleftrightarrow (\delta_1 - 1)\beta(1 - \beta r_0^{-\delta_1}) < \delta_1(\beta - r_0^{1-\delta_1}) \quad (\because \beta - r_0^{1-\delta_1} > 0) \\
&\Longleftrightarrow \frac{1}{\delta_1}\beta r_0^{\delta_1 - 1} + \left(1 - \frac{1}{\delta_1}\right)\beta^2 r_0^{-1} > 1 \quad \text{(simple algebra).}
\end{aligned}$$

Apply the arithmetic–geometric mean inequality ($\theta A + (1-\theta)B \geq A^\theta B^{1-\theta}$ for any nonnegative A and B and $0 \leq \theta \leq 1$) to obtain

$$\frac{1}{\delta_1}\beta r_0^{\delta_1-1} + \left(1 - \frac{1}{\delta_1}\right)\beta^2 r_0^{-1} \geq \left(\beta r_0^{\delta_1-1}\right)^{\frac{1}{\delta_1}} \cdot \left(\beta^2 r_0^{-1}\right)^{1-\frac{1}{\delta_1}} = \beta^{2-\frac{1}{\delta_1}} > 1,$$

because $\beta > 1$ and $2 - 1/\delta_1 > 1$. So the first inequality in (13) holds.

Similarly, we have

$$\begin{aligned}
k_2 > \frac{\beta_b(\rho-\mu_1)}{\beta_s(\rho-\mu_2)} &\Longleftrightarrow \frac{\delta_1\left(\beta_{\mathrm{b}} r_0^{1-\delta_2} - \beta_{\mathrm{s}} r_0\right)}{(1-\delta_1)(\beta_{\mathrm{b}} - \beta_{\mathrm{s}} r_0^{1-\delta_2})} > \frac{\delta_1(-\delta_2)}{(\delta_1-1)(1-\delta_2)} \cdot \frac{\beta_{\mathrm{b}}}{\beta_{\mathrm{s}}} \\
&\Longleftrightarrow \frac{\beta_{\mathrm{b}} r_0^{1-\delta_2} - \beta_{\mathrm{s}} r_0}{\beta_{\mathrm{s}} r_0^{1-\delta_2} - \beta_b} > \frac{(-\delta_2)}{1-\delta_2} \cdot \frac{\beta_{\mathrm{b}}}{\beta_{\mathrm{s}}} \\
&\Longleftrightarrow \frac{\beta r_0^{1-\delta_2} - r_0}{r_0^{1-\delta_2} - \beta} > \frac{-\delta_2}{1-\delta_2} \cdot \beta \\
&\Longleftrightarrow (1-\delta_2)(\beta r_0^{1-\delta_2} - r_0) > (-\delta_2)\beta(r_0^{1-\delta_2} - \beta) \quad (\because r_0^{1-\delta_2} > \beta) \\
&\Longleftrightarrow \frac{1}{1-\delta_2}\beta r_0^{-\delta_2} + \frac{-\delta_2}{1-\delta_2}\beta^2 r_0^{-1} > 1 \text{(simple algebra).}
\end{aligned}$$

Again, the AG mean inequality yields:

$$\frac{1}{1-\delta_2}\beta r_0^{-\delta_2}+\frac{-\delta_2}{1-\delta_2}\beta^2 r_0^{-1}>(\beta r_0^{-\delta_2})^{\frac{1}{1-\delta_2}}\cdot(\beta^2 r_0^{-1})^{\frac{-\delta_2}{1-\delta_2}}=\beta^{2-\frac{1}{1-\delta_2}}>1,$$

because $\beta>1$ and $(2-1/(1-\delta_2))>1$. Hence, the second inequality in (13) follows. □

Proof of Theorem 4.1 Note that the functions w_0 and w_1 have to satisfy all inequalities in (6). That is, for all $y>0$, we need

$$(\rho-\mathcal{L})w_0(y)\geq 0,\ (\rho-\mathcal{L})w_1(y)\geq 0,\ -\beta_\mathrm{b}+\beta_\mathrm{s}y\leq w_0(y)-w_1(y)\leq -\beta_\mathrm{s}+\beta_\mathrm{b}y. \tag{15}$$

Recall the equalities in (8). We have

$$\begin{array}{ll}(\rho-\mathcal{L})w_0=0 \text{ on } \Gamma_1\cup\Gamma_2, & (\rho-\mathcal{L})w_1=0 \text{ on } \Gamma_2\cup\Gamma_3,\\ w_0-w_1+\beta_\mathrm{b}-\beta_\mathrm{s}y=0 \text{ on } \Gamma_3, & w_1-w_0+\beta_\mathrm{b}y-\beta_\mathrm{s}=0 \text{ on } \Gamma_1.\end{array}$$

It is sufficient to show the following inequalities

$$\begin{array}{ll}\text{On } \Gamma_1: (\rho-\mathcal{L})w_1\geq 0, & w_0-w_1+\beta_\mathrm{b}-\beta_\mathrm{s}y\geq 0,\\ \text{On } \Gamma_2: w_0-w_1+\beta_\mathrm{b}-\beta_\mathrm{s}y\geq 0, & w_1-w_0+\beta_\mathrm{b}y-\beta_\mathrm{s}\geq 0,\\ \text{On } \Gamma_3: (\rho-L)w_0\geq 0, & w_1-w_0+\beta_\mathrm{b}y-\beta_\mathrm{s}\geq 0.\end{array}$$

Variational inequalities on Γ_1 and Γ_3. We first consider these inequalities on Γ_1. Using $(\rho-\mathcal{L})w_0=0$ and $w_0-w_1=-\beta_\mathrm{s}+\beta_\mathrm{b}y$, we have

$$w_0-w_1+\beta_\mathrm{b}-\beta_\mathrm{s}y=-\beta_\mathrm{s}+\beta_\mathrm{b}y+\beta_\mathrm{b}-\beta_\mathrm{s}y=(\beta_\mathrm{b}-\beta_\mathrm{s})(y+1)>0.$$

Note also that

$$\begin{aligned}(\rho-\mathcal{L})w_1&=(\rho-\mathcal{L})(w_0-\beta_\mathrm{b}y+\beta_\mathrm{s})=(\rho-\mathcal{L})w_0+(\rho-\mathcal{L})(\beta_\mathrm{s}-\beta_\mathrm{b}y)\\ &=(\rho-\mu_1)\beta_\mathrm{s}+\beta_\mathrm{b}(\mu_2-\rho)y.\end{aligned}$$

So, $(\rho-\mathcal{L})w_1\geq 0$ on Γ_1 iff $y\leq\beta_\mathrm{s}(\rho-\mu_1)/(\beta_\mathrm{b}(\rho-\mu_2))$, $\forall y\leq k_1$, iff $k_1\leq\beta_\mathrm{s}(\rho-\mu_1)/(\beta_\mathrm{b}(\rho-\mu_2))$. The last inequality follows from the first inequality in Lemma 4.1,

Similarly, on Γ_3, we have $w_1-w_0+\beta_\mathrm{b}y-\beta_\mathrm{s}\geq 0$. Also, $(\rho-\mathcal{L})w_0\geq 0$ is equivalent to $y\geq\beta_\mathrm{b}(\rho-\mu_1)/(\beta_\mathrm{s}(\rho-\mu_2))$, $\forall y\geq k_2$, which follows from the second inequality in Lemma 4.1.

Variational inequalities on Γ_2. Finally, we show these inequalities on Γ_2. Let

$$\phi_a(y)=w_1-w_0+\beta_\mathrm{b}y-\beta_\mathrm{s},\quad \phi_b(y)=w_0-w_1+\beta_\mathrm{b}-\beta_\mathrm{s}y.$$

Recall that $w_0=C_0y^{\delta_1}$ and $w_1=C_1y^{\delta_2}$ on Γ_2. We have

$$\phi_a(y)=C_1y^{\delta_2}-C_0y^{\delta_1}+\beta_\mathrm{b}y-\beta_\mathrm{s},\quad \phi_b(y)=C_0y^{\delta_1}-C_1y^{\delta_2}+\beta_\mathrm{s}y-\beta_\mathrm{b}.$$

In view of our smooth-fit selections of C_0 and C_1, it follows that

$$\phi_a(k_1) = \phi_a'(k_1) = 0 \quad \text{and} \quad \phi_b(k_2) = \phi_b'(k_2) = 0. \tag{16}$$

In addition, the variational inequalities on Γ_1 and Γ_3 and the continuity of w_0 and w_1 imply

$$\phi_a(k_2) \geq 0 \quad \text{and} \quad \phi_b(k_1) \geq 0. \tag{17}$$

Using the expressions of C_0 and C_1 in terms of k_1 in (11), we have

$$\begin{aligned}
\phi_a''(y) &= C_1\delta_2(\delta_2 - 1)y^{\delta_2-2} - C_0\delta_1(\delta_1 - 1)y^{\delta_1-2} \\
&= \frac{\beta_\mathrm{b}(1-\delta_1)k_1^{1-\delta_2} + \beta_\mathrm{s}\delta_1 k_1^{-\delta_2}}{\delta_1 - \delta_2}\delta_2(\delta_2 - 1)y^{\delta_2-2} \\
&\quad - \frac{\beta_\mathrm{b}(1-\delta_2)k_1^{1-\delta_1} + \beta_\mathrm{s}\delta_2 k_1^{-\delta_1}}{\delta_1 - \delta_2}\delta_1(\delta_1 - 1)y^{\delta_1-2} \\
&= \frac{\delta_1(-\delta_2)\beta_\mathrm{s}}{(\delta_1 - \delta_2)k_1^2}\left[(1-\delta_2)\left(\frac{y}{k_1}\right)^{\delta_2-2} + (\delta_1 - 1)\left(\frac{y}{k_1}\right)^{\delta_1-2}\right] \\
&\quad - \frac{(\delta_1 - 1)(1-\delta_2)\beta_\mathrm{b}}{(\delta_1 - \delta_2)k_1}\left[(-\delta_2)\left(\frac{y}{k_1}\right)^{\delta_2-2} + \delta_1\left(\frac{y}{k_1}\right)^{\delta_1-2}\right]
\end{aligned}$$

Take $y = k_1$ and use (14) to obtain

$$\phi_a''(k_1) = \frac{(\delta_1 - 1)(1-\delta_2)\beta_\mathrm{b}}{k_1^2}\left[\frac{\rho - \mu_1}{\rho - \mu_2} \cdot \frac{\beta_\mathrm{s}}{\beta_\mathrm{b}} - k_1\right] > 0.$$

Next we write $\phi_a''(y)$ using the expression of C_0 and C_1 in terms of k_2:

$$\begin{aligned}
\phi_a''(y) &= \frac{\delta_1(-\delta_2)\beta_\mathrm{b}}{(\delta_1 - \delta_2)k_2^2}\left[(1-\delta_2)\left(\frac{y}{k_2}\right)^{\delta_2-2} + (\delta_1 - 1)\left(\frac{y}{k_2}\right)^{\delta_1-2}\right] \\
&\quad - \frac{(\delta_1 - 1)(1-\delta_2)\beta_\mathrm{s}}{(\delta_1 - \delta_2)k_2}\left[(-\delta_2)\left(\frac{y}{k_2}\right)^{\delta_2-2} + \delta_1\left(\frac{y}{k_2}\right)^{\delta_1-2}\right].
\end{aligned}$$

Set $y = k_2$ to obtain

$$\phi_a''(k_2) = \frac{(\delta_1 - 1)(1-\delta_2)\beta_\mathrm{s}}{k_2^2}\left[\frac{\rho - \mu_1}{\rho - \mu_2} \cdot \frac{\beta_\mathrm{b}}{\beta_\mathrm{s}} - k_2\right] < 0.$$

Note also that ϕ_a'' has a unique zero in $[k_1, k_2]$. This together with (16) and (17) implies $\phi_a \geq 0$ on Γ_2. In addition, note that $\phi_b''(y) = -\phi_a''(y)$. It follows that $\phi_b''(k_1) < 0$ and $\phi_b''(k_2) > 0$. Combining with the boundary conditions for ϕ_b in (16) and (17), we can show $\phi_b \geq 0$ on Γ_2. This completes the proof. □

5 A Verification Theorem

Theorem 5.1 *We have* $v_i(x_1, x_2) = x_1 w_i(x_2/x_1) = V_i(x_1, x_2)$, $i = 0, 1$. *Moreover, if initially* $i = 0$, *let* $\Lambda_0^* = (\tau_1^*, \tau_2^*, \tau_3^*, \ldots)$ *such that* $\tau_1^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_3^* = \inf\{t \geq \tau_2^* : (X_t^1, X_t^2) \in \Gamma_3\}$, *and so on. Similarly, if initially* $i = 1$, *let* $\Lambda_1^* = (\tau_0^*, \tau_1^*, \tau_2^*, \ldots)$ *such that* $\tau_0^* = \inf\{t \geq 0 : (X_t^1, X_t^2) \in \Gamma_1\}$, $\tau_1^* = \inf\{t \geq \tau_0^* : (X_t^1, X_t^2) \in \Gamma_3\}$, $\tau_2^* = \inf\{t \geq \tau_1^* : (X_t^1, X_t^2) \in \Gamma_1\}$, *and so on. Then* Λ_0^* *and* Λ_1^* *are optimal.*

Proof The proof is divided into five steps.
Step 1. $C_0 > 0$, $C_1 > 0$, **and** $v_0(x_1, x_2) \geq 0$.

In view of the definition of C_0 and C_1 in (11), we have $C_0 > 0 \iff \beta_s(1-\delta_2)k_2 + \beta_b\delta_2 > 0 \iff k_2 > \beta_b(-\delta_2)/(\beta_s(1-\delta_2))$. On the other hand, in view of (13) and (14), we have $k_2 > \beta_b(\rho - \mu_1)/(\beta_s(\rho - \mu_2)) > \beta_b\delta_1(-\delta_2)/(\beta_s(\delta_1 - 1)(1 - \delta_2)) > \beta_b(-\delta_2)/(\beta_s(1 - \delta_2))$. So $C_0 > 0$. Similarly, we can show $C_1 > 0$. To see $v_0 \geq 0$, it suffices to show $w_0 \geq 0$ on Γ_3. It can be seen that $w_0(k_2) > 0$, $w_0'(k_2) > 0$, and $w_0''(k_2) = C_1\delta_2(\delta_2 - 1)y^{\delta_2 - 2} > 0$. It follows that w_0' is increasing on Γ_3 and therefore $w_0' > 0$. This implies in turn that w_0 is increasing on Γ_3. Hence, $w_0 > 0$ on Γ_3.
Step 2. $-Ax_1 - Bx_2 \leq v_i(x_1, x_2) \leq Ax_1 + Bx_2$, $i = 0, 1$, **for some** A **and** B.

We only show the case when $i = 0$. The proof for the other case is similar. First, on $\Gamma_1 \cup \Gamma_2$, we have $0 \leq v_0(x_1, x_2) = C_0 x_1^{1-\delta_1} x_2^{\delta_1} = C_0 x_1 (x_2/x_1)^{\delta_1} \leq C_0 x_1 (k_2)^{\delta_1}$. Next, on Γ_3, $-\beta_b x_1 + \beta_s x_2 \leq v_0(x_1, x_2) = C_1 x_1 (x_2/x_1)^{\delta_2} - \beta_b x_1 + \beta_s x_2 \leq C_1 x_1 (k_2)^{\delta_2} - \beta_b x_1 + \beta_s x_2$. Hence, we can choose suitable A and B so that the inequalities hold.
Step 3. Let $dY_t = \mu_0 dt + \sigma_0 dW_t$, $Y_0 = y$, **with constants** μ_0, σ_0 **and a standard Brownian motion** W_t. **For any given** $l_1 < l_2$, **define a sequence of stopping times**

$$\tau_1^0 = \inf\{t \geq 0 : Y_t = l_1\},\ \tau_2^0 = \inf\{t \geq \tau_1^0 : Y_t = l_2\},\ \tau_3^0 = \inf\{t \geq \tau_2^0 : Y_t = l_1\},\ \ldots$$

Then, for any given $\rho_0 > 0$, $Ee^{-\rho_0\tau_n^0} \to 0$, **as** $n \to \infty$.

Note that

$$\begin{aligned} Ee^{-\rho_0\tau_2^0} &= Ee^{-\rho_0\tau_2^0} I_{\{\tau_1^0 < \infty\}} + Ee^{-\rho_0\tau_2^0} I_{\{\tau_1^0 = \infty\}} \\ &= E[e^{-\rho_0\tau_2^0} | \tau_1^0 < \infty] P(\tau_1^0 < \infty) + 0 \\ &= E[e^{-\rho_0(\tau_2^0 - \tau_1^0)} \cdot e^{-\rho_0\tau_1^0} | \tau_1^0 < \infty] P(\tau_1^0 < \infty) \\ &= E[e^{-\rho_0(\tau_2^0 - \tau_1^0)} | \tau_1^0 < \infty] E[e^{-\rho_0\tau_1^0} | \tau_1^0 < \infty] P(\tau_1^0 < \infty) \\ &= E[e^{-\rho_0(\tau_2^0 - \tau_1^0)} | \tau_1^0 < \infty] E[e^{-\rho_0\tau_1^0}], \end{aligned}$$

where we used independence of $(\tau_2^0 - \tau_1^0)$ and τ_1^0 on $\{\tau_1^0 < \infty\}$ and $E[e^{-\rho_0(\tau_2^0 - \tau_1^0)} | \tau_1^0 < \infty] = e^{(l_2 - l_1)\left(\mu_0 - \sqrt{2\rho_0 + \mu_0^2}\right)}$. Let $\eta_0 = \max\left\{ e^{(l_2 - l_1)\left(\mu_0 - \sqrt{2\rho_0 + \mu_0^2}\right)}, \right.$

$e^{(l_2-l_1)\left(-\mu_0-\sqrt{2\rho_0+\mu_0^2}\right)}\Bigg\}$. Then, $0<\eta_0<1$ and $Ee^{-\rho_0\tau_2^0}\le\eta_0 Ee^{-\rho_0\tau_1^0}$. Repeat this procedure to obtain $Ee^{-\rho_0\tau_{n+1}^0}\le\eta_0^n Ee^{-\rho_0\tau_1^0}$. It follows that $Ee^{-\rho_0\tau_n^0}\to 0$, as $n\to\infty$.

Step 4. $v_i(x_1,x_2)\ge J_i(x_1,x_2,\Lambda_i)$.

Note that the functions v_0 and v_1 are continuously differentiable on the entire region $\{x_1>0,x_2>0\}$ and twice continuously differentiable in the interior of $\Gamma_i, i=1,2,3$. In addition, they satisfy the quasi-variational inequalities in (15). In particular, $\rho v_i(x)-\mathcal{A}v_i(x)\ge 0$, $i=0,1$, whenever they are twice continuously differentiable. Using these inequalities, Dynkin's formula, and Fatou's lemma as in Øksendal [15, p. 226], we have, $E\left(e^{-\rho(\theta_1\wedge N)}v_i\left(X^1_{\theta_1\wedge N},X^2_{\theta_1\wedge N}\right)\right)\ge E\left(e^{-\rho(\theta_2\wedge N)}v_i\left(X^1_{\theta_2\wedge N},X^2_{\theta_2\wedge N}\right)\right)$, for any stopping times $0\le\theta_1\le\theta_2$, a.s., and any N. Note that, for each $j=1,2$,

$$\begin{aligned}E\left(e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)\right)&=E\left(e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)\right)I_{\{\theta_j<\infty\}}\\&\quad+E\left(e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)\right)I_{\{\theta_j=\infty\}}\\&=E\left(e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)\right)I_{\{\theta_j<\infty\}}\\&\quad+E\left(e^{-\rho N}v_i(X^1_N,X^2_N)\right)I_{\{\theta_j=\infty\}}.\end{aligned}$$

In view of Step 2, the second term on the right-hand side converges to zero because both $Ee^{-\rho N}X^1_N$ and $Ee^{-\rho N}X^2_N$ go to zero as $N\to\infty$. Note also that

$$e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)I_{\{\theta_j<\infty\}}\to e^{-\rho\theta_j}v_i\left(X^1_{\theta_j},X^2_{\theta_j}\right)I_{\{\theta_j<\infty\}},\ \text{a.s.}\tag{18}$$

as $N\to\infty$. Moreover, we can show as in the proof of Lemma 3.1 that both $\{e^{-\rho(\theta_j\wedge N)}X^1_{\theta_j\wedge N}\}$ and $\{e^{-\rho(\theta_j\wedge N)}X^2_{\theta_j\wedge N}\}$ are uniformly integrable. This together with Step 2 implies the uniform integrability of $\left\{e^{-\rho(\theta_j\wedge N)}v_i\left(X^1_{\theta_j\wedge N},X^2_{\theta_j\wedge N}\right)\right\}$. Sending $N\to\infty$ in (18), we have

$$E\left(e^{-\rho\theta_1}v_i\left(X^1_{\theta_1},X^2_{\theta_1}\right)I_{\{\theta_1<\infty\}}\right)\ge E\left(e^{-\rho\theta_2}v_i\left(X^1_{\theta_2},X^2_{\theta_2}\right)I_{\{\theta_2<\infty\}}\right),\quad\text{for } i=0,1.\tag{19}$$

Given $\Lambda_0=(\tau_1,\tau_2,\ldots)$, using the third inequalities in (15) and (19), we have

$$\begin{aligned}v_0(x_1,x_2)&\ge E\left(e^{-\rho\tau_1}v_0\left(X^1_{\tau_1},X^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right)\\&\ge E\left(e^{-\rho\tau_1}\left(v_1\left(X^1_{\tau_1},X^2_{\tau_1}\right)-\beta_bX^1_{\tau_1}+\beta_sX^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right)\\&=E\left(e^{-\rho\tau_1}v_1\left(X^1_{\tau_1},X^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}-e^{-\rho\tau_1}\left(\beta_bX^1_{\tau_1}-\beta_sX^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right)\\&\ge E\left(e^{-\rho\tau_2}v_1\left(X^1_{\tau_2},X^2_{\tau_2}\right)I_{\{\tau_2<\infty\}}-e^{-\rho\tau_1}\left(\beta_bX^1_{\tau_1}-\beta_sX^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right)\\&\ge E\left(e^{-\rho\tau_2}\left(v_0\left(X^1_{\tau_2},X^2_{\tau_2}\right)+\beta_sX^1_{\tau_2}-\beta_bX^2_{\tau_2}\right)I_{\{\tau_2<\infty\}}-e^{-\rho\tau_1}\left(\beta_bX^1_{\tau_1}-\beta_sX^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right)\\&=E\left(e^{-\rho\tau_2}v_0\left(X^1_{\tau_2},X^2_{\tau_2}\right)I_{\{\tau_2<\infty\}}\right)\\&\quad+E\left(e^{-\rho\tau_2}\left(\beta_sX^1_{\tau_2}-\beta_bX^2_{\tau_2}\right)I_{\{\tau_2<\infty\}}-e^{-\rho\tau_1}\left(\beta_bX^1_{\tau_1}-\beta_sX^2_{\tau_1}\right)I_{\{\tau_1<\infty\}}\right).\end{aligned}$$

Repeat this process and recall that $v_0(x_1, x_2) \geq 0$ to obtain

$$\begin{aligned} v_0(x_1, x_2) \geq E & \left(e^{-\rho\tau_2}(\beta_s X^1_{\tau_2} - \beta_b X^2_{\tau_2}))I_{\{\tau_2<\infty\}} - e^{-\rho\tau_1}(\beta_b X^1_{\tau_1} - \beta_s X^2_{\tau_1})I_{\{\tau_1<\infty\}}\right) \\ & + \cdots + E\Big(e^{-\rho\tau_{2n}}(\beta_s X^1_{\tau_{2n}} - \beta_b X^2_{\tau_{2n}}))I_{\{\tau_{2n}<\infty\}} \\ & - e^{-\rho\tau_{2n-1}}(\beta_b X^1_{\tau_{2n-1}} - \beta_s X^2_{\tau_{2n-1}})I_{\{\tau_{2n-1}<\infty\}}\Big). \end{aligned}$$

Sending $n \to \infty$ to obtain $v_0(x_1, x_2) \geq J_0(x_1, x_2, \Lambda_0)$ for all Λ_0. So, $v_0(x_1, x_2) \geq V_0(x_1, x_2)$. Similarly, we can show that $v_1(x_1, x_2) \geq V_1(x_1, x_2)$.

Step 5. $v_i(x_1, x_2) = J_i(x_1, x_2, \Lambda^*_i)$. Define $\tau^*_1 = \inf\{t \geq 0 : (X^1_t, X^2_t) \in \Gamma_3\}$. Using again Dynkin's formula and noticing that, for each n, $v_0(x_1, x_2) = E[e^{-\rho(\tau^*_1\wedge n)}v_0(X^1_{\tau^*_1\wedge n}, X^2_{\tau^*_1\wedge n})]$. Note also that $\lim_{n\to\infty} E[e^{-\rho(\tau^*_1\wedge n)}v_0 (X^1_{\tau^*_1\wedge n}, X^2_{\tau^*_1\wedge n})] = E[e^{-\rho\tau^*_1}v_0(X^1_{\tau^*_1}, X^2_{\tau^*_1})I_{\{\tau^*_1<\infty\}}]$. It follows that $v_0(x_1, x_2) = E[e^{-\rho\tau^*_1}v_0(X^1_{\tau^*_1}, X^1_{\tau^*_1})I_{\{\tau^*_1<\infty\}}] = E[e^{-\rho\tau^*_1}(v_1(X^1_{\tau^*_1}, X^1_{\tau^*_1}) - \beta_b X^1_{\tau^*_1} + \beta_s X^2_{\tau^*_2})I_{\{\tau^*_1<\infty\}}]$. Let $\tau^*_2 = \inf\{t \geq \tau^*_1 : (X^1_t, X^2_t) \in \Gamma_1\}$. We have also

$$\begin{aligned} & E\left(e^{-\rho\tau^*_1}v_1(X^1_{\tau^*_1}, X^2_{\tau^*_1})I_{\{\tau^*_1<\infty\}}\right) = E\left(e^{-\rho\tau^*_2}v_1(X^1_{\tau^*_2}, X^2_{\tau^*_2})I_{\{\tau^*_2<\infty\}}\right) \\ & \quad = E\left(e^{-\rho\tau^*_2}\left(v_0(X^1_{\tau^*_2}, X^2_{\tau^*_2}) + \beta_s X^1_{\tau^*_2} - \beta_b X^2_{\tau^*_2}\right)I_{\{\tau^*_2<\infty\}}\right). \end{aligned}$$

Combine these to obtain

$$\begin{aligned} v_0(x_1, x_2) = E\Big[& e^{-\rho\tau^*_2}v_0\left(X^1_{\tau^*_2}, X^2_{\tau^*_2}\right)I_{\{\tau^*_2<\infty\}} \\ & + e^{-\rho\tau^*_2}\left(\beta_s X^1_{\tau^*_2} - \beta_b X^2_{\tau^*_2}\right)I_{\{\tau^*_2<\infty\}} - e^{-\rho\tau^*_1}\left(\beta_b X^1_{\tau^*_1} - \beta_s X^2_{\tau^*_1}\right)I_{\{\tau^*_1<\infty\}}\Big]. \end{aligned}$$

Continue this way to obtain

$$\begin{aligned} v_0(x_1, x_2) = & E\left(e^{-\rho\tau^*_{2n}}v_0\left(X^1_{\tau^*_{2n}}, X^2_{\tau^*_{2n}}\right)I_{\{\tau^*_{2n}<\infty\}}\right) \\ & + E\left(e^{-\rho\tau^*_2}(\beta_s X^1_{\tau^*_2} - \beta_b X^2_{\tau^*_2}))I_{\{\tau^*_2<\infty\}} - e^{-\rho\tau^*_1}(\beta_b X^1_{\tau^*_1} - \beta_s X^2_{\tau^*_1})I_{\{\tau^*_1<\infty\}}\right) \\ & + \cdots + E\Big(e^{-\rho\tau^*_{2n}}(\beta_s X^1_{\tau^*_{2n}} - \beta_b X^2_{\tau^*_{2n}}))I_{\{\tau^*_{2n}<\infty\}} \\ & - e^{-\rho\tau^*_{2n-1}}\left(\beta_b X^1_{\tau^*_{2n-1}} - \beta_s X^2_{\tau^*_{2n-1}}\right)I_{\{\tau^*_{2n-1}<\infty\}}\Big). \end{aligned}$$

It remains to show that $Ee^{-\rho\tau^*_{2n}}v_0\left(X^1_{\tau^*_{2n}}, X^2_{\tau^*_{2n}}\right)I_{\{\tau^*_{2n}<\infty\}} \to 0$. In view of the linear upper and lower bound functions in Step 2, it suffices to show $Ee^{-\rho\tau^*_{2n}}X^i_{\tau^*_{2n}}I_{\{\tau^*_{2n}<\infty\}} \to 0$, $i = 1, 2$. Note that $Ee^{-\rho\tau^*_{2n}}X^i_{\tau^*_{2n}}I_{\{\tau^*_{2n}<\infty\}} = \lim_{N\to\infty} Ee^{-\rho(\tau^*_{2n}\wedge N)}X^i_{\tau^*_{2n}\wedge N} = Ee^{-(\rho-\mu_i)\tau^*_{2n}}$. It suffices to show the right-hand term above $Ee^{-(\rho-\mu_i)\tau^*_{2n}} \to 0$, $i = 1, 2$. Let $Y_t = \log(X^1_t/X^2_t)$. Then, $dY_t = \left(\mu_2 - \mu_1 - \frac{a_{22}}{2} + \frac{a_{11}}{2}\right)dt + (\sigma_{21} - \sigma_{11})dW^1_t + (\sigma_{22} - \sigma_{12})dW^2_t$. Note that $\{\tau^*_k\}$ can be defined in terms of Y_t hitting times

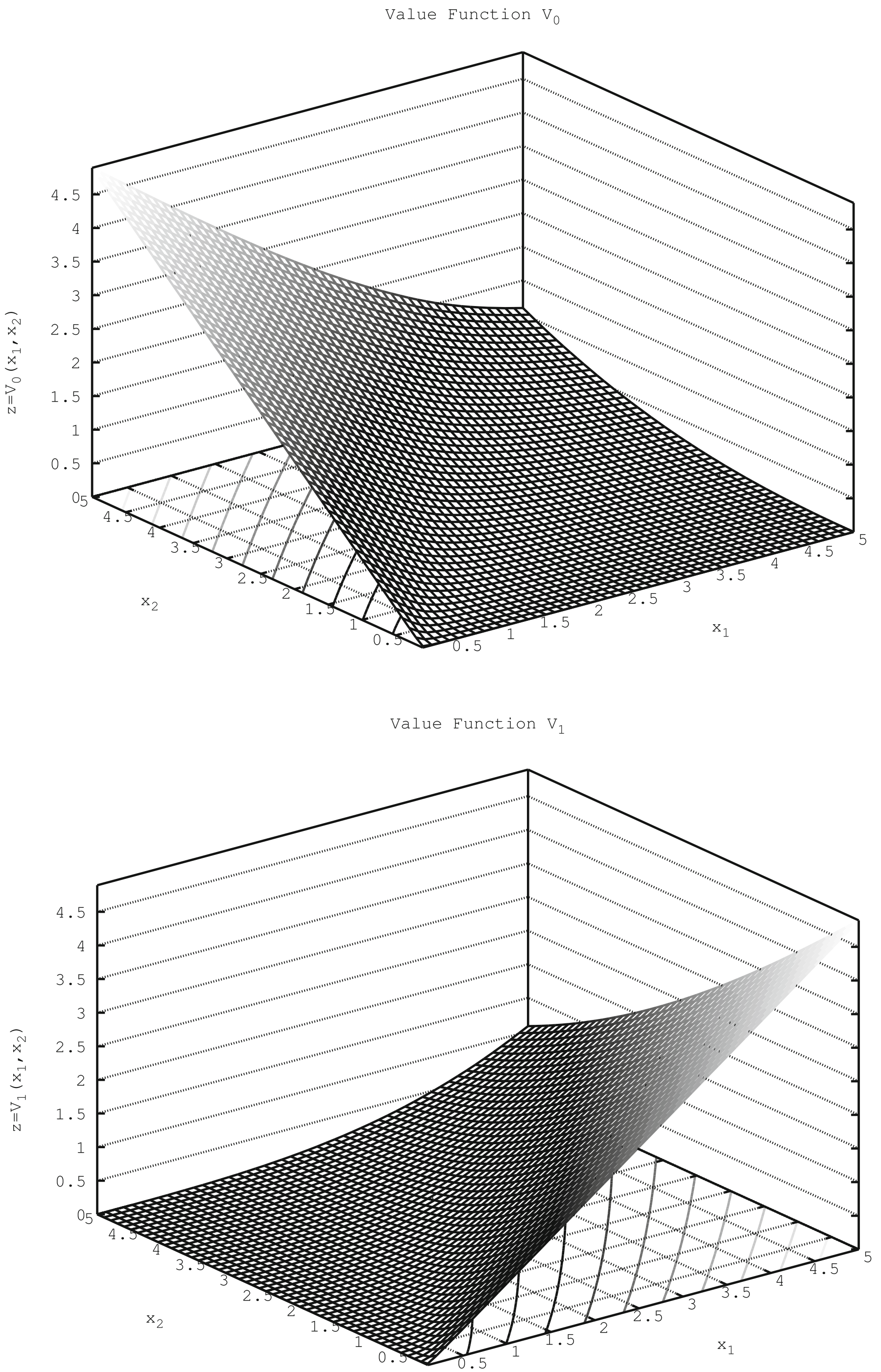

Fig. 3 Value functions

of $l_1 = \log k_1$ and $l_2 = \log k_2$. In view of this and Step 3, the desired convergence $Ee^{-(\rho-\mu_i)\tau_{2n}^*} \to 0$ follows. This completes the proof. □

6 A Numerical Example

In this section, we use the parameters of the TGT-WMT example, i.e., $\mu_1 = 0.2059$, $\mu_2 = 0.2459$, $\sigma_{11} = 0.3112$, $\sigma_{12} = 0.0729$, $\sigma_{21} = 0.0729$, $\sigma_{22} = 0.2943$. In addition, we take $K = 0.001$ and $\rho = 0.5$. Using (12), we can solve for k_1 and k_2 and obtain $k_1 = 1.03905$ and $k_2 = 1.28219$. The corresponding value functions are given in Fig. 3.

Dependence of (k_1, k_2) on parameters. Next, we vary one of the parameters at a time and examine the dependence of (k_1, k_2) on these parameters. First we consider how the pair (k_1, k_2) changes with μ_1. A larger μ_1 implies greater potential of growth in $\mathbf{S}^1$. It can be seen in Table 1 that both k_1 and k_2 decrease in μ_1 leading to more buying opportunities. Also, if we vary μ_2, the pair (k_1, k_2) increases in μ_2. This is because larger μ_2 means bigger growth potential in $\mathbf{S}^2$ which discourages establishing pairs position $\mathbf{Z}$ and encourages its early exit. In Table 2, we vary the volatility σ_{11} and σ_{22}. Larger volatility leads higher risk, which translates to smaller buying zone Γ_3. On the other hand, larger volatility gives more room for the price to move. This leads to smaller selling zone Γ_1 (Table 3).

Table 1 (k_1, k_2) with varying μ_1 and μ_2

μ_1	0.1059	0.1559	0.2059	0.2559	0.3059
k_1	1.38860	1.21356	1.03905	0.86272	0.68532
k_2	1.70104	1.49268	1.28219	1.07150	0.86008
μ_2	0.1459	0.1959	0.2459	0.2959	0.3459
k_1	0.75424	0.87372	1.03905	1.28131	1.67831
k_2	0.92168	1.07205	1.28219	1.59780	2.11803

Table 2 (k_1, k_2) with varying σ_{11} and σ_{22}

σ_{11}	0.2112	0.2612	0.3112	0.3612	0.4112
k_1	1.05320	1.04598	1.03905	1.02997	1.02008
k_2	1.26384	1.27295	1.28219	1.29364	1.30417
σ_{22}	0.1943	0.2443	0.2943	0.3443	0.3943
k_1	1.05147	1.04511	1.03905	1.03224	1.02469
k_2	1.26597	1.27399	1.28219	1.29133	1.30136

Table 3 (k_1, k_2) with varying $\sigma_{12} (= \sigma_{21})$

$\sigma_{12} (= \sigma_{21})$	−0.0271	0.0229	0.0729	0.1229	0.1729
k_1	1.00965	1.02318	1.03905	1.05546	1.07276
k_2	1.32062	1.30251	1.28219	1.26127	1.23904

Table 4 (k_1, k_2) with varying ρ

ρ	0.40	0.45	0.50	0.55	0.60
k_1	1.10935	1.06547	1.03905	1.02291	1.00997
k_2	1.41886	1.33396	1.28219	1.24591	1.22105

Table 5 (k_1, k_2) with varying K

K	0.0001	0.0005	0.001	0.002	0.003
k_1	1.07951	1.06318	1.03905	1.00787	0.98627
k_2	1.23819	1.25562	1.28219	1.31728	1.34231

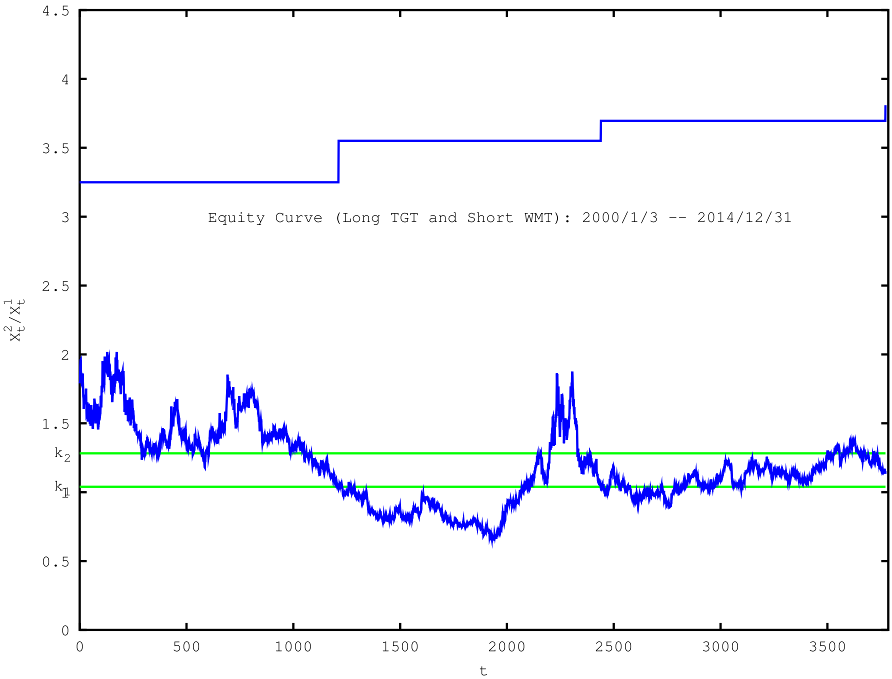

Fig. 4 $\mathbf{S}^1$=TGT, $\mathbf{S}^2$=WMT: the threshold levels k_1, k_2 and the corresponding equity curve

Next, we vary σ_{12} which equals σ_{21}. Note that this parameter dictates the correlation between X_t^1 and X_t^2. Larger σ_{12} leads to greater correlation, which encourages more buying opportunities (larger Γ_3) and more selling as well (larger Γ_1).

In Table 4, we vary the discount rate ρ. Larger ρ encourages quicker profits, which leads to more buying and shorter holding. This is confirmed in Table 4. It shows that larger ρ leads to a smaller k_2 and smaller $(k_2 - k_1)$.

Finally, we examine the dependence on K. Clearly, a larger K discourages trading transactions. This results smaller buying zone Γ_3 and smaller selling zone Γ_1 (Table 5).

Backtesting (TGT–WMT). We backtest our pairs-trading rule using the stock prices of TGT and WMT from 2000 to 2014. Using the parameters obtained in Example 2.1 based on the historical prices from 1985 to 1999, we found the pair $(k_1, k_2) =$

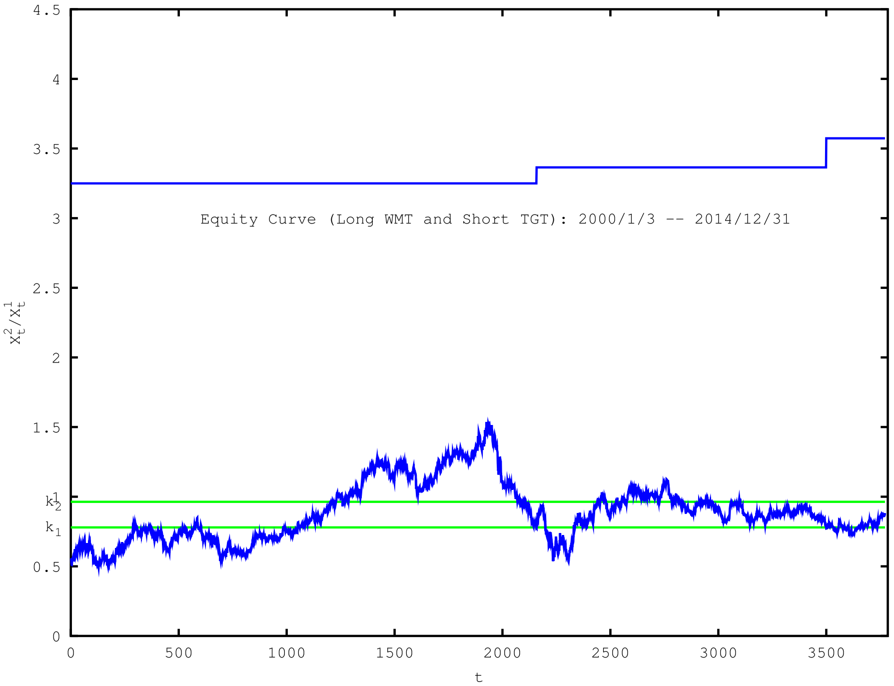

Fig. 5 $\mathbf{S}^1$=WMT, $\mathbf{S}^2$=TGT: the threshold levels k_1, k_2 and the corresponding equity curve

(1.03905, 1.28219). A pairs trading (long $\mathbf{S}^1$ and short $\mathbf{S}^2$) is triggered when (X_t^1, X_t^2) enters Γ_3. The position is closed when (X_t^1, X_t^2) enters Γ_1. Initially, we allocate trading the capital \$100 K. When the first long signal is triggered, buy \$50K TGT stocks and short the same amount of WMT. Such half-and-half capital allocation between long and short applies to all trades. In addition, each pairs transaction is charged \$5 commission. In Fig. 5, the corresponding ratio X_t^2/X_t^1, the threshold levels k_1 and k_2, and the corresponding equity curve are plotted. There are total 3 trades, and the end balance is \$155.914 K (Fig. 4).

We can also switch the roles of $\mathbf{S}^1$ and $\mathbf{S}^2$, i.e., to long WMT and short TGT by taking $\mathbf{S}^1$=WMT and $\mathbf{S}^2$=TGT. In this case, the new $(\tilde{k}_1, \tilde{k}_2) = (1/k_2, 1/k_1) = (1/1.28219, 1/1.03905)$. These levels and the corresponding equity curve are given in Fig. 5. Such trade leads to the end balance \$132.340 K. Note that both types of trades have no overlap, i.e., they do not compete for the same capital. The grand total profit is \$88254 which is a 88.25% gain. Note also that there are only 5 trades in the fifteen-year period leaving the capital in cash most of the time. This is desirable because the cash sitting in the account can be used for other types of shorter-term trading in between, at least drawing interest over time.

7 Conclusions

In this paper, we have studied the pairs-trading problem following geometric Brownian motions and obtained a closed-form solution. The major advantage of pairs trading is its risk neutral nature, i.e., it can be profitable regardless of general market directions. It would be interesting to examine how the method works for a larger selection of stocks. In addition, it would also be interesting to study the problem under more realistic models, e.g., GBM's with regime switching.

Acknowledgements We thank the anonymous referee and the editors for their valuable comments and suggestions, which led to much improvements of this paper. This research is supported in part by the Simons Foundation (235179) to (Qing Zhang).

References

1. Gatev, E., Goetzmann, W.N., Rouwenhorst, K.G.: Pairs trading: performance of a relative-value arbitrage rule. Revi. Financ. Stud. **19**, 797–827 (2006)
2. Vidyamurthy, G.: Pairs Trading: Quantitative Methods and Analysis. Wiley, Hoboken (2004)
3. Song, Q.S., Zhang, Q.: An optimal pairs-trading rule. Automatica **49**, 3007–3014 (2013)
4. McDonald, R., Siegel, D.: The value of waiting to invest. Q. J. Econ. **101**, 707–727 (1986)
5. Hu, Y., Øksendal, B.: Optimal time to invest when the price processes are geometric Brownian motions. Financ. Stoch. **2**, 295–310 (1998)
6. Davis, M.H.A., Norman, A.R.: Portfolio selection with transaction costs. Math. Oper. Res. **15**, 676–713 (1990)
7. Shreve, S.E., Soner, H.M.: Optimal investment and consumption with transaction costs. Ann. Appl. Probab. **4**, 609–692 (1994)
8. Zhang, Q.: Stock trading: an optimal selling rule. SIAM J. Control Optim. **40**, 64–87 (2001)
9. Guo, X., Zhang, Q.: Optimal selling rules in a regime switching model. IEEE Trans. Autom. Control **50**, 1450–1455 (2005)
10. Dai, M., Zhang, Q., Zhu, Q.: Trend following trading under a regime switching model. SIAM J. Financ. Math. **1**, 780–810 (2010)
11. Iwarere, S., Barmish, B.R.: A confidence interval triggering method for stock trading via feedback control. In: Proceedings of American Control Conference, Baltimore, MD, (2010)
12. Merhi, A., Zervos, M.: A model for reversible investment capacity expansion. SIAM J. Control Optim. **46**, 839–876 (2007)
13. Løkka, A., Zervos, M: Long-term optimal investment strategies in the presence of adjustment costs. SIAM J. Control Optim. **51**, 996–1034 (2013)
14. Zhang, H., Zhang, Q.: Trading a mean-reverting asset: buy low and sell high. Automatica **44**, 1511–1518 (2008)
15. Øksendal, B.: Stochastic Differential Equations, 6th edn. Springer-Verlag, New York (2003)