International Journal of Financial Engineering
Vol. 3, No. 3 (2016) 1650021 (20 pages)

DOI: 10.1142/S2424786316500213

World Scientific
www.worldscientific.com

Trading VIX futures under mean reversion with regime switching

Jiao Li
APAM Department, Columbia University
New York, NY 10027, United States

Received: 26 May 2016; Revised: 11 June 2016; Accepted: 12 June 2016
Published: 3 November 2016

Abstract

This paper studies the optimal VIX futures trading problems under a regime-switching model. We consider the VIX as mean reversion dynamics with dependence on the regime that switches among a finite number of states. For the trading strategies, we analyze the timings and sequences of the investor's market participation, which leads to several corresponding coupled system of variational inequalities. The numerical approach is developed to solve these optimal double stopping problems by using projected-successive-over-relaxation (PSOR) method with Crank–Nicolson scheme. We illustrate the optimal boundaries via numerical examples of two-state Markov chain model. In particular, we examine the impacts of transaction costs and regime-switching timings on the VIX futures trading strategies.

Keywords: Optimal stopping; mean reversion; futures trading; regime switching; variational inequality.

JEL Classifications: C41; G11; G13

1. Introduction

The Chicago Board Options Exchange (CBOE) Volatility Index, commonly known as the VIX, is widely used measure of the market volatility. It was introduced by the CBOE back in 1993, and its calcualtion is based on the implied volatility of S&P 500 index options. Empirically, the VIX is shown to be negatively correlated to the market index. One broadly accepted explanation is that investors buy put options on S&P500 for protection against market turmoil. This increases the option prices and implied volatilities, and thus, the value of the VIX. As a result, the VIX is also called the investor fear gauge (Whaley, 2000).

Email address: jl4170@columbia.edu

While VIX itself is not traded, investors can gain exposure to the index by trading VIX futures. VIX futures, which are cash settled on the VIX index level, were first traded on the CBOE Futures Exchange in 2004, and since then the market has been growing constantly. Recent market data illustrate the liquidity and popularity of the VIX futures market. During 2015, the average daily volume (ADV) in the VIX futures was over 200,000 contracts, with 51.6 million VIX futures contracts traded in total.

Each VIX futures comes with a fixed term, ranging from 1 to 9 months, but futures holders do not need to keep the position through the expiration date, and can choose when to close out the position. Furthermore, before entering the market, the investor can opt to start a long or short position, followed by closing it at later time before expiration.

In this paper, we investigate the VIX futures trading under Cox–Ingersoll–Ross (CIR) model with regime switching. The CIR dynamics is well studied in many empirical studies (Grübichler and Longstaff, 1996; Wang and Daigler, 2011). Zhang and Zhu (2006) analyze the empirical validity of the CIR model by first estimating its parameters from VIX historical data and provide a futures pricing formula. Later Dotsis *et al.* (2007) add jumps to a CIR diffusion. However, studies of Mencía and Sentana (2013) and Sircar and Papanicolaou (2014) show that VIX implied volatility data can be better reproduced by incorporating regime switching. Leung *et al.* (2016) display two characteristically different term structures observed in the VIX futures market. These markedly different regimes offer a very interesting testing ground for analyzing the performance of different models for volatility derivatives. It is necessary to allow the key parameters of the VIX to respond to the general market movements. The regime-switching model is one of such formulations, where the model parameters depend on the market regime that switches among a finite number of states. The market regime could reflect the state of the market, the general mood of investors, and other economic factors. Elliott *et al.* (2008) evaluate the risk measures for derivatives via a partial differential equation (PDE) approach when the underlying asset dynamics are associated with regime switching. The applications on regime switching for derivatives pricing and optimal stopping problems have been well studied in the literature (Guo, 2001; Buffington and Elliott, 2002; Le and Wang, 2010).

Moreover, we introduce the investor's optimal strategies to participate the trading. In the first strategy, an investor is expected to establish the long position when the price is sufficiently low, and then exits when the price is high. The opposite is expected for the second strategy. In both cases, the presence of transaction costs expands the waiting region, indicating the investor's desire for better prices. In addition, the waiting region expands drastically near expiry since transaction costs discourage entry when futures is very close to maturity. Finally, the main feature of our trading problem approach is to combine these two related problems and analyze the optimal strategy when an investor has the freedom to choose between either a *long-short* or a *short-long* position. Among our results, we find that when the investor has the right to choose, she delays market entry to wait for better prices compared to the individual standalone problems.

In an earlier study, Brennan and Schwartz (1990) propose the optimal strategies for stock index futures with position limits, where the Brownian bridge process is applied for the the basis dynamics. Dai *et al.* (2011) extend the optimal strategies with the additional flexibility by allowing the investor to switch between long position and short position directly. In this paper, we extend the model introduced in Leung *et al.* (2016) by incorporating regime switching into the CIR model for the VIX. We provide a link between the futures pricing problem under the risk-neutral measure and the trading problem conducted under the historical measure. In contrast to Brennan and Schwartz (1990) and Dai *et al.* (2011), we do not *a priori* assume the existence of a stochastic basis that may or may not be consistent with the futures prices, but consider the long and short strategies that take advantage of the temporal price difference of futures. Similar trading strategies have been studied by Leung and Li (2015), Leung *et al.* (2015), Leung *et al.* (2014), Leung and Li (2016), Leung and Shirai (2015) and Stepanek (2015), among others. Moreover, the strategy studied herein can be automated. For a comprehensive study on algorithmic trading, we refer the reader to Cartea *et al.* (2015). The ideas can be applied to other derivatives, such as swaps (Leung and Yamazaki, 2013; Leung and Liu, 2012).

With regime switching, the optimal trading strategies are determined from a system of m variational inequalities, where m is the number of possible regimes of the market. We apply an implicit-explicit finite difference method with projected-successive-over-relaxation (PSOR) to solve for the optimal trading boundaries under all regimes. While there are a number of studies on optimal stopping problems with regime switching as mentioned above, there are very few that discuss the numerical methods and solutions. In a related study, Khaliq and Liu (2009) develop a penalty method method for pricing the regime-switching American option.

The rest of the paper is organized as follows. In Sec. 2, we formulate the optimal stopping and trading strategies with regime switching. In Sec. 3, we develop an implicit scheme based on PSOR method. In Sec. 4, we present the numerical results for optimal trading problems and provide financial interpretations.

2. Optimal Timing to Trade Futures with Regime Switching

We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ is the historical measure. Let ξ be a continuous-time irreducible finite-state Markov chain with state space

$E = \{1, 2, \ldots, m\}$. The generator matrix of ξ is denoted by Q, which has constant entries q_{ij} for $i, j \in E$, such that $q_{ij} \geq 0$ for $i \neq j$ and $\sum_{j \in E} q_{ij} = 0$ for each $i \in E$. This Markov chain represents the changing regime of the financial market, and it influences the dynamics of the index.

Let B be a standard Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and assume it is independent of ξ. The VIX, denoted by S, is assumed to follow the CIR process regime switching:

$$dS_t = \mu(\xi_t)(\theta(\xi_t) - S_t)dt + \sigma(\xi_t)\sqrt{S_t}dB_t, \tag{2.1}$$

where for each $i \in E$, the coefficients $\mu(i)$, $\theta(i)$ and $\sigma(i)$ are known constants, with $\mu(i)$, $\theta(i)$, $\sigma(i) > 0$. Note that $\mu(i)$, $\theta(i)$ and $\sigma(i)$ depend on the Markov chain, representing the mean reversion rate, the long-run mean and the volatility of the VIX at regime i.

To price futures, we assume a re-parameterized CIR model for the risk-neutral VIX dynamics. Due to the additional uncertainty described by regime switching, we note that introducing a Markov chain results in an incomplete market. Hence the equivalent martingale measure is not unique. Elliott *et al.* (2008) employ the regime-switching Esscher transform to determine an equivalent martingale pricing measure. We do not include the argument in our paper and instead assume that the risk-neutral probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ is given. Thus, under the risk-neutral measure $\mathbb{Q}$, the VIX follows

$$dS_t = \tilde{\mu}(\xi_t)(\tilde{\theta}(\xi_t) - S_t)dt + \sigma(\xi_t)\sqrt{S_t}dB_t^{\mathbb{Q}}, \tag{2.2}$$

where $\tilde{\mu}(i)$, $\tilde{\theta}(i)$, $\sigma(i) > 0$ and $B^{\mathbb{Q}}$ is a $\mathbb{Q}$-standard Brownian motion which is also independent of Markov chain. For convenience, we may use the subscript notation for these constants, e.g. $\mu_i \equiv \mu(i)$, $\theta_i \equiv \theta(i)$, and $\sigma_i \equiv \sigma(i)$. In both SDEs, (2.1) and (2.2), we require $2\mu_i\theta_i \geq \sigma_i^2$ and $2\,\tilde{\mu}_i\tilde{\theta}_i \geq \sigma_i^2$ (Feller condition) so that the CIR process stays strictly positive at all times. The two Brownian motions are related by

$$dB_t^{\mathbb{Q}} = dB_t + \frac{\mu(\xi_t)(\theta(\xi_t) - S_t) - \tilde{\mu}(\xi_t)(\tilde{\theta}(\xi_t) - S_t)}{\sigma(\xi_t)\sqrt{S_t}}dt, \tag{2.3}$$

such as change of measure preserves the CIR model, up to different parameter values across two measures.

We consider a futures contract written on the S with maturity $T < \infty$ and futures price at time $0 \leq t \leq T$ when $S_t = s$ and regime $\xi_t = i$. The price of a futures contract is given by

$$f(t, s, i) = \mathbb{E}^{\mathbb{Q}}\{S_T | S_t = s, \xi_t = i\}, \quad t \leq T, \tag{2.4}$$

where $\mathbb{E}^{\mathbb{Q}}$ is the expectation operator with respect to the risk-neutral measure $\mathbb{Q}$. We can show that $f_i(t,s) \equiv f(t,s,i)$, $i = 1,\ldots,m$, satisfy the following partial differential equation (Yao *et al.*, 2006),

$$\begin{cases} \dfrac{\partial f_i}{\partial t} + \tilde{\mu}_i(\tilde{\theta}_i - s)\dfrac{\partial f_i}{\partial s} + \dfrac{\sigma_i^2 s}{2}\dfrac{\partial^2 f_i}{\partial s^2} + \displaystyle\sum_{j\neq i} \tilde{q}_{ij}(f_j - f_i) = 0, & (t,s) \in [0,T) \times \mathbb{R}_+, \\ f_i(T,s) = s, & s \in \mathbb{R}_+, \end{cases} \tag{2.5}$$

where $\tilde{q}_{ij}$ is the transition probability under measure $\mathbb{Q}$. We note that (2.5) involves m interconnected PDEs due to the regime switching introduced in the VIX dynamics. One can efficiently compute the futures prices in all regimes by finite difference methods (see Sec. 3).

2.1. Optimal double stopping approach

Let us consider the scenario in which an investor has a long position in a futures contract with expiration date T. With a long position in the futures, the investor can hold it till maturity, but can also close the position early by taking an opposite position at the prevailing market price. At maturity, the two opposite positions cancel each other. This motivates us to investigate the best time to close.

If the investor selects to close the long position at time $\tau \leq \hat{T}$, then she will receive the market value of the futures on the expiry date, denoted by $f(\tau, S_\tau, \xi_\tau)$, minus the transaction cost $c \geq 0$. To maximize the expected discounted value, evaluated under the investor's historical probability measure $\mathbb{P}$ with a constant subjective discount rate $r \geq 0$, the investor solves the optimal stopping problem

$$\mathcal{V}(t,s,i) = \sup_{\tau \in \mathcal{T}_{t,\hat{T}}} \mathbb{E}\{e^{-r(\tau-t)}(f(\tau, S_\tau, \xi_\tau) - c) | S_t = s, \xi_t = i\}, \tag{2.6}$$

where $\mathcal{T}_{t,\hat{T}}$ is the set of all stopping times, with taking values between t and $\hat{T}$, where $\hat{T} \in (0,T]$ is the trading deadline, which can equal but not exceed the futures' maturity. Throughout this chapter, we continue to use the notation $\mathbb{E}\{\cdot\}$ to indicate the expectation taken under the historical probability measure $\mathbb{P}$.

The value function $\mathcal{V}(t,s,i)$ represents the expected liquidation value associated with the long futures position in i-state. Prior to taking the long position in f_i, the investor, with zero position, can select the optimal timing to start the trade, or not to enter at all. This leads us to analyze the timing option inherent in the trading problem. Precisely, at time $t \leq \hat{T}$, the investor faces the optimal entry timing problem

$$\mathcal{J}(t,s,i) = \sup_{\nu\in\mathcal{T}_{t,\hat{T}}} \mathbb{E}\{e^{-r(\nu-t)}(\mathcal{V}(\nu,S_\nu,\xi_\nu) - (f(\nu,S_\nu,\xi_\nu)+\hat{c}))^+ | S_t = s, \xi_t = i\}, \tag{2.7}$$

where $\hat{c} \geq 0$ is the transaction cost, which may differ from c. In other words, the investor seeks to maximize the expected difference between the value function $\mathcal{V}(\nu,S_\nu,\xi_\nu)$ associated with the long position and the prevailing futures price $f(\nu,S_\nu,\xi_\nu)$. The value function $\mathcal{J}(\nu,S_\nu,\xi_\nu)$ represents the maximum expected value of the trading opportunity embedded in the futures. We refer this "long to open, short to close" strategy as the *long-short* strategy.

Alternatively, an investor may well choose to short a futures contract with the speculation that the futures price will fall. By taking a short futures position, the investor can either close it out later by establishing a long position, or hold it until the expiry which will result in a cash settlement. Given an investor who has a unit short position in the futures contract, the objective is to minimize the expected discounted cost to close out this position at/before maturity. The optimal timing strategy is determined from

$$\mathcal{U}(t,s,i) = \inf_{\tau\in\mathcal{T}_{t,\hat{T}}} \mathbb{E}\{e^{-r(\tau-t)}(f(\tau,S_\tau,\xi_\tau)+\hat{c}) | S_t = s, \xi_t = i\}. \tag{2.8}$$

If the investor begins with a zero position, then she can decide when to enter the market by solving

$$\mathcal{K}(t,s,i) = \sup_{\nu\in\mathcal{T}_{t,\hat{T}}} \mathbb{E}\{e^{-r(\nu-t)}((f(\nu,S_\nu,\xi_\nu)-c) - \mathcal{U}(\nu,S_\nu,\xi_\nu))^+ | S_t = s, \xi_t = i\}. \tag{2.9}$$

We call this "short to open, long to close" strategy as the *short-long* strategy.

When an investor contemplates entering the market, she can either long or short first. Therefore, on top of the timing option, the investor has an additional choice between the long-short and short-long strategies. Hence, the investor solves the market entry timing problem:

$$\mathcal{P}(t,s,i) = \sup_{\varsigma\in\mathcal{T}_{t,\hat{T}}} \mathbb{E}\{e^{-r(\varsigma-t)}\max\{\mathcal{A}(\varsigma,S_\varsigma,\xi_\varsigma),\mathcal{B}(\varsigma,S_\varsigma,\xi_\varsigma)\} | S_t = s, \xi_t = i\}, \tag{2.10}$$

with two alternative rewards upon entry defined by

$$\begin{aligned}
\mathcal{A}(\varsigma,S_\varsigma,\xi_\varsigma) &:= (\mathcal{V}(\varsigma,S_\varsigma,\xi_\varsigma) - (f(\varsigma,S_\varsigma,\xi_\varsigma)+\hat{c}))^+ \quad \text{(long-short)},\\
\mathcal{B}(\varsigma,S_\varsigma,\xi_\varsigma) &:= ((f(\varsigma,S_\varsigma,\xi_\varsigma)-c) - \mathcal{U}(\varsigma,S_\varsigma,\xi_\varsigma))^+ \quad \text{(short-long)}.
\end{aligned}$$

Accordingly, the corresponding inputs associated with the optimal stopping problem in (2.10) is given by taking the maximum between the above two rewards.

2.2. Variational inequalities and optimal trading strategies

Given the CIR dynamics of the VIX, the value functions defined in (2.6)–(2.10) all satisfy the same governing differential equation in their respective continuation regions and their values equal to the corresponding rewards in their exercise regions. In order to solve for the optimal trading strategies, we need to study the coupled systems of variational inequalities respectively. To this end, we first define the operators:

$$\mathcal{L}_i\{\cdot\} := -r\cdot + \frac{\partial\cdot}{\partial t} + \mu_i(\theta_i - s)\frac{\partial\cdot}{\partial s} + \frac{\sigma_i^2 s}{2}\frac{\partial^2\cdot}{\partial s^2} + \sum_{j\neq i} q_{ij}(\cdot_j - \cdot_i), \tag{2.11}$$

corresponding to CIR model. For convenience, we adopt the subscript notation for these value functions, e.g. $\mathcal{V}_i(t,s) \equiv \mathcal{V}(t,s,i)$, $\mathcal{J}_i(t,s) \equiv \mathcal{J}(t,s,i)$, $\mathcal{U}_i(t,s) \equiv \mathcal{U}(t,s,i)$, $\mathcal{K}_i(t,s) \equiv \mathcal{K}(t,s,i)$ and $\mathcal{P}_i(t,s) \equiv \mathcal{P}(t,s,i)$.

The optimal exit and entry problems $\mathcal{J}_i$ and $\mathcal{V}_i$ associated with the *long-short* strategy are solved from the following pair of variational inequalities:

$$\max\{\mathcal{L}_i\mathcal{V}_i(t,s), (f_i(t,s) - c) - \mathcal{V}_i(t,s)\} = 0, \tag{2.12}$$

$$\max\{\mathcal{L}_i\mathcal{J}_i(t,s), (\mathcal{V}_i(t,s) - (f_i(t,s) + \hat{c}))^+ - \mathcal{J}_i(t,s)\} = 0, \tag{2.13}$$

for $(t,s) \in [0,\hat{T}] \times \mathbb{R}_+$. Similarly, the reverse *short-long* strategy can be determined by numerically solving the variational inequalities satisfied by $\mathcal{U}_i$ and $\mathcal{K}_i$:

$$\min\{\mathcal{L}_i\mathcal{U}_i(t,s), (f_i(t,s) + \hat{c}) - \mathcal{U}_i(t,s)\} = 0, \tag{2.14}$$

$$\max\{\mathcal{L}_i\mathcal{K}_i(t,s), ((f_i(t,s) - c) - \mathcal{U}_i(t,s))^+ - \mathcal{K}_i(t,s)\} = 0. \tag{2.15}$$

To determine the optimal timing to enter the futures market, we solve the variational inequality

$$\max\{\mathcal{L}_i\mathcal{P}_i(t,s), \max\{\mathcal{A}_i(t,s), \mathcal{B}_i(t,s)\} - \mathcal{P}_i(t,s)\} = 0. \tag{2.16}$$

In other words, we have to solve the variational inequality (2.12)–(2.15), and then use the solution as inputs to the variational inequality (2.16).

3. Numerical Implementation with Regime Switching

The numerical solution of the system of variational inequalities can be obtained by applying a finite-difference scheme in all regimes with the use of the projected-successive-over-relaxation (PSOR) method.[1] The solution of the resulting

[1]We refer to Chapter 9 of Wilmott *et al.* (1995) for a detailed discussion on the projected SOR method.

equations for value functions are solved by the successive over relaxation (SOR) method. In each SOR iterative step in finding the numerical approximation of the value functions, we simply take the maximum value between the approximated function value and compensated futures price. The futures price can be pre-computed by (2.5) conveniently. Similar numerical schemes with regime switching can be found in Leung (2010). More details of our numerical scheme are described in the following.

First, we define the generic differential operator

$$\mathcal{L}_i\{\cdot\} := -r\cdot + \frac{\partial\cdot}{\partial t} + \varphi_i(s)\frac{\partial\cdot}{\partial s} + \frac{\sigma_i^2(s)}{2}\frac{\partial^2\cdot}{\partial s^2} + \sum_{j\neq i} q_{ij}(\cdot_j - \cdot_i), \tag{3.1}$$

then the variational inequalities (2.12)–(2.16) admit the same form as the following variational inequality problem:

$$\begin{cases} \mathcal{L}_i g_i(t,s) \leq 0, g_i(t,s) \geq h_i(t,s), & (t,s) \in [0,\hat{T}) \times \mathbb{R}_+, \\ (\mathcal{L}_i g_i(t,s))(h_i(t,s) - g_i(t,s)) = 0, & (t,s) \in [0,\hat{T}) \times \mathbb{R}_+, \\ g_i(\hat{T},s) = h_i(\hat{T},s), & s \in \mathbb{R}_+. \end{cases} \tag{3.2}$$

Here, $g_i(t,s)$ represents the value functions $\mathcal{V}_i(t,s)$, $\mathcal{J}_i(t,s)$, $-\mathcal{U}_i(t,s)$, $\mathcal{K}_i(t,s)$, or $\mathcal{P}_i(t,s)$. The function $h_i(t,s)$ represents $f_i(t,s)-c$, $(\mathcal{V}_i(t,s)-(f_i(t,s)+\hat{c}))^+$, $-(f_i(t,s)+\hat{c})$, $(f_i(t,s)-c)-\mathcal{U}_i(t,s))^+$, or $\max\{\mathcal{A}_i(t,s),\mathcal{B}_i(t,s)\}$. The futures price $f_i(t,s)$, with $\hat{T}\leq T$, is given by (2.4).

We now consider the discretization of the partial differential equation $\mathcal{L}_i g_i(t,s)=0$, over an uniform grid with discretizations in time ($\delta t = \frac{\hat{T}}{N}$), and space ($\delta s = \frac{S\max}{M}$). Applying the Crank–Nicolson method for s-derivatives and backward difference for t-derivatives on the resulting equation leads the finite difference equation:

$$\begin{aligned} &-\frac{r}{2}(g_i^{m,n}+g_i^{m,n-1}) + \frac{g_i^{m,n}-g_i^{m,n-1}}{\delta t} + \frac{\varphi_i^m}{2}\Bigg(\frac{g_i^{m+1,n}-g_i^{m-1,n}}{2\delta s} \\ &\quad + \frac{g_i^{m+1,n-1}-g_i^{m-1,n-1}}{2\delta s}\Bigg) + \frac{(\sigma_i^m)^2}{2}\Bigg(\frac{g_i^{m+1,n}-2g_i^{m,n}+g_i^{m-1,n}}{2(\delta s)^2} \\ &\quad + \frac{g_i^{m+1,n-1}-2g_i^{m,n-1}+g_i^{m-1,n-1}}{2(\delta s)^2}\Bigg) \\ &\quad + \frac{q_{ii}}{2}(g_i^{m,n}+g_i^{m,n-1}) + \sum_{j\neq i}\frac{q_{ij}}{2}(g_j^{m,n}+g_j^{m,n-1}) = 0, \end{aligned} \tag{3.3}$$

where $q_{ii} = -\sum_{j\neq i} q_{ij}$ is used. For convenience, we may use the subscript notation for these constants, e.g. $g_i^{m,n} \equiv g_i(n\delta t, m\delta s)$, $h_i^{m,n} \equiv h_i(n\delta t, m\delta s)$, $\varphi_i^m = \varphi_i(m\delta s)$

and $\sigma_i^m = \sigma_i(m\delta s)$. We implement by explicitly treating the regime coupling terms. Replace $g_j^{m,n-1}$ with $g_j^{m,n}$ for $j \neq i$ and obtain

$$\begin{aligned}
&-\alpha_i^m g_i^{m-1,n-1} + (1-\beta_i^m) g_i^{m,n-1} - \gamma_i^m g_i^{m+1,n-1} \\
&\quad = \alpha_i^m g_i^{m-1,n} + (1+\beta_i^m) g_i^{m,n} + \gamma_i^m g_{m+1,n} + \delta t \sum_{j\neq i} q_{ij} g_j^{m,n},
\end{aligned} \tag{3.4}$$

where

$$\begin{cases}
\alpha_i^m = \dfrac{\delta t}{4\delta s}\left(\dfrac{(\sigma_i^m)^2}{\delta s} - \varphi_i^m\right), \\
\beta_i^m = -\dfrac{\delta t}{2}\left((r - q_{ii}) + \dfrac{(\sigma_i^m)^2}{(\delta s)^2}\right), \\
\gamma_i^m = \dfrac{\delta t}{4\delta s}\left(\dfrac{(\sigma_i^m)^2}{\delta s} + \varphi_i^m\right),
\end{cases} \tag{3.5}$$

for $i, j \in E$, $m = 1, 2, \ldots, M-1$ and $n = 1, 2, \ldots, N-1$. The system to be solved backward in time is

$$\mathbf{M_i^1 g_i^{n-1}} = \mathbf{r_i^n}, \tag{3.6}$$

where the right-hand side is

$$\mathbf{r_i^n} = \mathbf{M_i^2 g_i^n} + \delta t \sum_{j\neq i} q_{ij} \mathbf{g_j^n} + \alpha_i^1 \begin{bmatrix} g_i^{0,n-1} + g_i^{0,n} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \gamma_i^{M-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ g_i^{M,n-1} + g_i^{M,n} \end{bmatrix}, \tag{3.7}$$

and

$$\mathbf{M_i^1} = \begin{bmatrix}
1-\beta_i^1 & -\gamma_i^1 & & & & \\
-\alpha_i^2 & 1-\beta_i^2 & -\gamma_i^2 & & & \\
 & -\alpha_i^3 & 1-\beta_i^3 & -\gamma_i^3 & & \\
 & & \ddots & \ddots & \ddots & \\
 & & & -\alpha_i^{M-2} & 1-\beta_i^{M-2} & -\gamma_i^{M-2} \\
 & & & & -\alpha_i^{M-1} & 1-\beta_i^{M-1}
\end{bmatrix}, \tag{3.8}$$

$$\mathbf{M_i^2} = \begin{bmatrix} 1+\beta_i^1 & \gamma_i^1 & & & & \\ \alpha_i^2 & 1+\beta_i^2 & \gamma_i^2 & & & \\ & \alpha_i^3 & 1+\beta_i^3 & \gamma_i^3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \alpha_i^{M-2} & 1+\beta_i^{M-2} & \gamma_i^{M-2} \\ & & & & \alpha_i^{M-1} & 1+\beta_i^{M-1} \end{bmatrix}, \tag{3.9}$$

$$\mathbf{g_i^n} = \left[g_i^{1,n}, g_i^{2,n}, \ldots, g_i^{M-1,n} \quad \right]^T. \tag{3.10}$$

We note that the dimension of $\mathbf{M_i^1}$ is independent of m, which is the number of regimes. Thus, the scheme (3.6) can be computed in parallel.

Remark 1. The futures price function (2.5) can be computed via solving the linear system (3.6) by replacing $g_i^{m,n}$ with $f_i^{m,n}$, where $f_i^{m,n} \equiv f(t,s,i)$, and setting $r=0$.

Since the investor can establish her position at anytime before the expiry, the value functions $g_i(t,s)$ must satisfy the constraint

$$g_i(t,s) \geq h_i(t,s), \quad s \geq 0, \quad 0 \leq t \leq \hat{T}, \quad i \in E, \tag{3.11}$$

where the discrete scheme can be written as

$$g_i^{m,n} \geq h_i^{m,n}, \quad 0 \leq m \leq M, \quad 0 \leq n \leq M, \quad i \in E. \tag{3.12}$$

Hence, at each time step $n \in \{1, 2, \ldots, N-1\}$, we need to solve

$$\begin{cases} \mathbf{M_i^1 g_i^{n-1}} \geq \mathbf{r_i^n}, \\ \mathbf{g_i^{n-1}} \geq \boldsymbol{h}_i^{n-1}, \\ (\mathbf{M_i^1 g_i^{n-1}} - \mathbf{r_i^n})^T (\boldsymbol{h}_i^{n-1} - \mathbf{g_i^{n-1}}) = 0. \end{cases} \tag{3.13}$$

To guarantee the constraint, our algorithm enforces the constraint explicitly as follows:

$$g_{i,new}^{m,n-1} = \max\{g_{i,old}^{m,n-1}, h_i^{m,n-1}\}. \tag{3.14}$$

The projected SOR method is used to solve the linear system. Note that the constraint is enforced at the same time as the iterate $g_{i,(k+1)}^{m,n-1}$ is calculated; the effect of the constraint is immediately felt in the calculation of $g_{i,(k+1)}^{m+1,n-1}$, $g_{i,(k+1)}^{m+2,n-1}$, etc. Thus, at each time step n, the PSOR algorithm is to iterate (on k) the equations

$$g_{i,(k+1)}^{1,n-1} = \max\left\{ h_i^{1,n-1}, g_{i,(k)}^{1,n-1} + \frac{\omega}{1-\beta_i^1}[r_i^{1,n} - (1-\beta_i^1) g_{i,(k)}^{1,n-1} + \gamma_i^1 g_{i,(k)}^{2,n-1}] \right\},$$

$$
\begin{aligned}
g_{i,(k+1)}^{2,n-1} &= \max\left\{h_i^{2,n-1}, g_{i,(k)}^{2,n-1} + \frac{\omega}{1-\beta_i^2}[r_i^{2,n} + \alpha_i^2 g_{i,(k+1)}^{1,n-1} - (1-\beta_i^2)g_{i,(k)}^{2,n-1}\right. \\
&\quad \left. + \gamma_i^2 g_{i,(k)}^{3,n-1}]\right\}, \\
&\vdots \\
g_{i,(k+1)}^{M-1,n-1} &= \max\left\{h_i^{M-1,n-1}, g_{i,(k)}^{M-1,n-1}\right. \\
&\quad \left. + \frac{\omega}{1-\beta_i^{M-1}}[r_i^{M-1,n} + \alpha_i^{M-1} g_{i,(k+1)}^{M-2,n-1} - (1-\beta_i^{M-1})g_{i,(k)}^{M-1,n-1}]\right\},
\end{aligned} \tag{3.15}
$$

where k is the iteration counter and ω is the overrelaxation parameter. The iterative scheme starts from an initial point $\mathbf{g}_{i,(0)}^n$ and proceeds until a convergence criterion is met, such as $||\mathbf{g}_{\mathbf{i},(k+1)}^{\mathbf{n-1}} - \mathbf{g}_{\mathbf{i},(k)}^{\mathbf{n-1}}|| < \epsilon$, where ϵ is a tolerance parameter. The optimal boundary $S_f(t)$ can be identified by locating the boundary that separates the regions where $g_i(t,s) = h_i(t,s)$, or $g_i(t,s) > h_i(t,s)$.

4. Optimal Trading Strategies

In this section, we provide numerical examples to further interpret the optimal trading strategies. The regime-switching CIR model is capable of generating futures curves of different term structures. Figure 1 displays the futures prices in

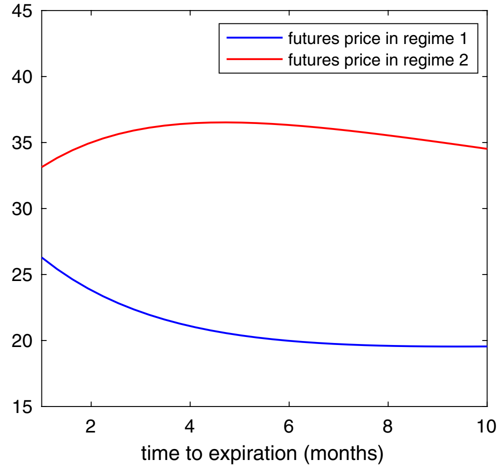


Fig. 1. The futures prices in 2 regimes. Parameters: $S_0 = 30, \sigma_1 = 5.33, \sigma_2 = 6.42, \tilde{\theta}_1 = 18.16, \tilde{\theta}_2 = 40.36$, $\tilde{\mu}_1 = 4.55$, $\tilde{\mu}_2 = 4.59, q_{12} = -q_{11} = 0.1, q_{21} = -q_{22} = 0.5$.

two different regimes. The CIR model with regime switching generates a convex curve in regime 1 (blue), and a concave curve in regime 2 (red). In Fig. 2, we illustrate the cases of optimal boundaries for futures trading under the CIR model in a two-regime market. As Fig. 2(a) shows, optimal boundaries divide the space into three disjoint regions in each regime, which can be specified as the long region (region below "$\mathcal{J}$"), short region (region above "$\mathcal{V}$") and the waiting region (region between "$\mathcal{J}$" and "$\mathcal{V}$"). The subscripts of value functions index to the regimes. Assuming the investor pre-commits to "long-short" strategy, it is optimal

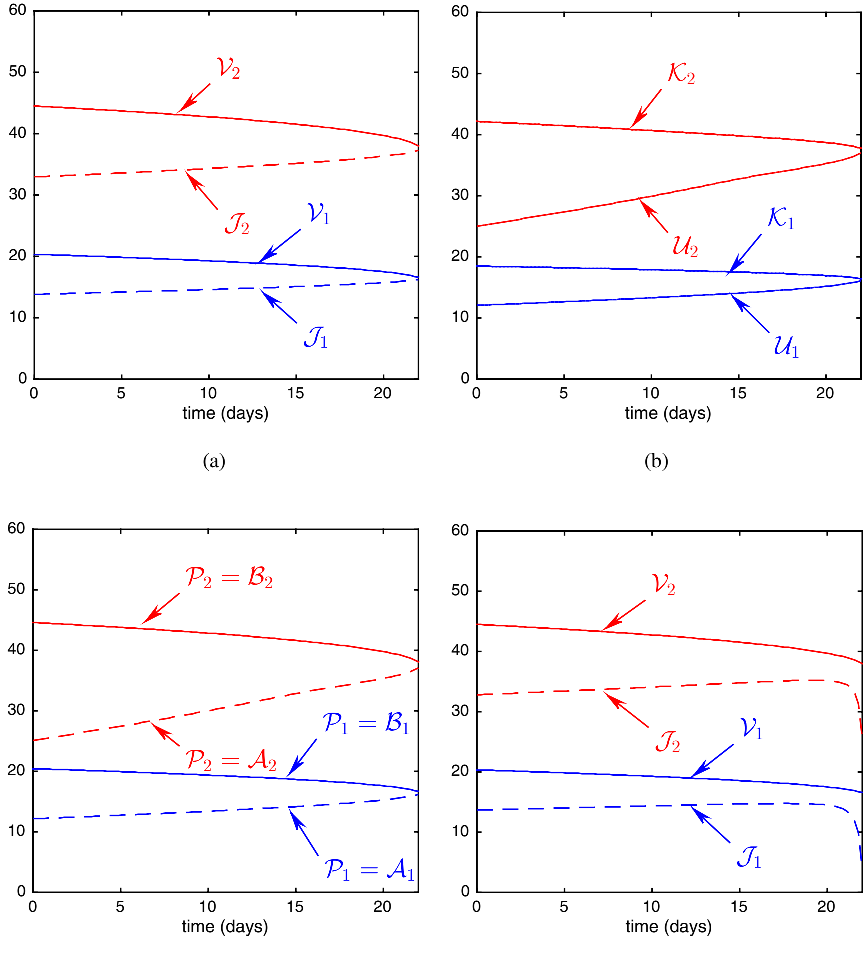


Fig. 2. Optimal long-short boundaries for futures trading with 2-state regime-switching model. Left panel: $c = \hat{c} = 0$. Right panel: $c = \hat{c} = 0.01$. Common parameters: $\hat{T} = \frac{22}{252}$, $T = \frac{66}{252}$, $r = 0.05, \sigma_1 = 5.33, \sigma_2 = 6.42, \theta_1 = 17.58, \theta_2 = 39.5, \tilde{\theta}_1 = 18.16, \tilde{\theta}_2 = 40.36, \mu_1 = 8.57, \mu_2 = 9$, $\tilde{\mu}_1 = 4.55$, $\tilde{\mu}_2 = 4.59, q_{12} = -q_{11} = 0.1, q_{21} = -q_{22} = 0.5$.

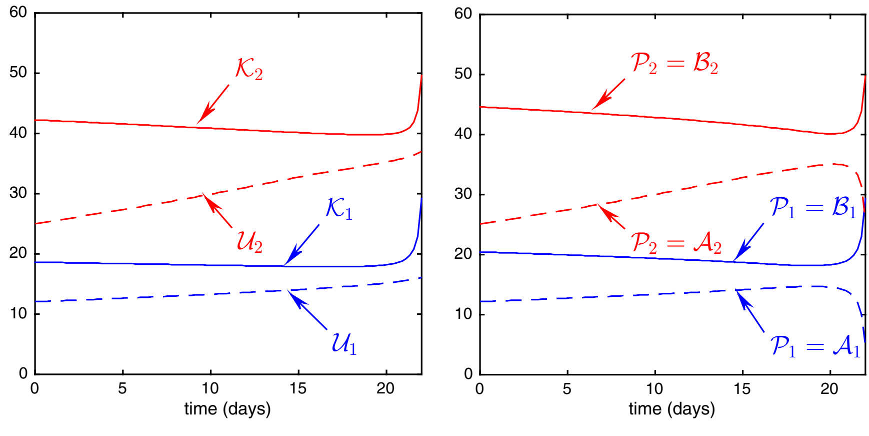


Fig. 2. (*Continued*)

to take a long position first if the VIX is in the long region, and then exit the market when the VIX goes up to hit optimal boundary "$\mathcal{J}$". While if the investor adopts the "short-long" strategy, she will first short a futures and subsequently close out with a long position by the "$\mathcal{K}$" and "$\mathcal{U}$" boundaries in Fig. 2(b). In either case, our strategies confirm the intuition: "buy low and sell high".

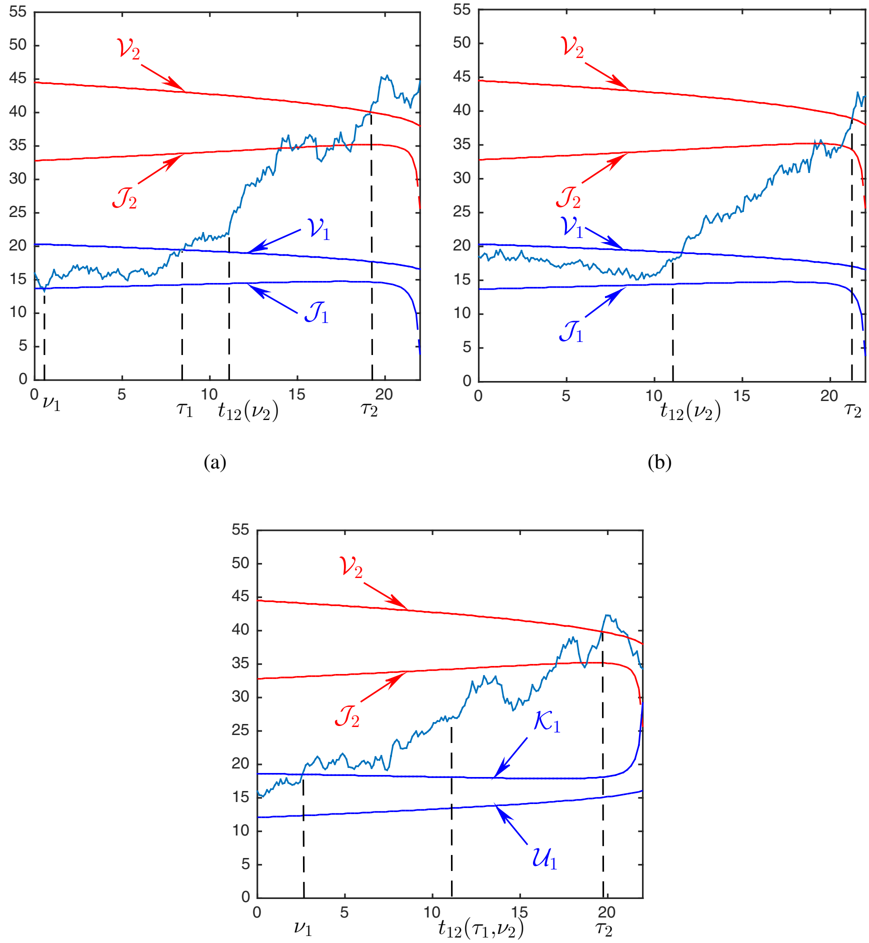
Figure 2 also displays the optimal boundaries with different transaction costs. Without transaction costs (see left panel of Fig. 2), the waiting region shrinks since the investor tends to enter and exit the market eariler, resulting in more rapid trades. In the presence of transaction costs (see right panel of Fig. 2), the waiting region is widen in order to save on transaction costs. It should also be noted that as the transaction cost increases, the long boundary decreases and the short boundary increases, making the investor trade less frequently. In particular, the fast divergence near expiry indicates that the investor should not enter the market after a critical time. The intuition is that a rational investor will never initiate a position if she does not have enough time to recover at least the transaction costs.

The "$\mathcal{P}$" boundaries in Fig. 2 indicates the optimal value of VIX at which the investor should open a position. The boundary labeled as "$\mathcal{P} = \mathcal{A}$" (resp. "$\mathcal{P} = \mathcal{B}$") indicates the critical value at which the investor enters the market by taking a *long* (resp. *short*) futures position. The investor should choose the "short-first" strategy if the VIX lies in the area above the "$\mathcal{P} = \mathcal{B}$" boundary, whereas choose the "long-first" strategy if the VIX is lower than the "$\mathcal{P} = \mathcal{A}$" boundary. The area between the two boundaries is the region where the investor should wait for an better enter opportunity. This confirms our intuition —take a long position when the VIX is low while take a short position when the VIX is high. Similarly, the waiting region expands significantly near expiry to cover transaction costs. In other words, the investor will not enter the market unless the VIX is either very low or very high.

Once the investor finished choosing entry strategy, she could resort the exit strategy to corresponding optimal boundaries. For example, if the investor starts by longing a futures, then the optimal exit timing to close her position is represented by the optimal boundary "$\mathcal{V}$" in Fig. 2. However, if the investor's initial position is short, then she will hold the short position until the VIX hits "$\mathcal{U}$" boundary.

To better explain our strategies, here we assume a two-state process and interpret the two regimes as a *low-mean* regime (regime 1) and a *high-mean* regime (regime 2). The regime-switching timing t_{12} indicates the shift from the *low-mean* regime to the *high-mean* regime. As shown in Fig. 3(a), the investor finishes the long-low-short-high tradings in both regimes. In the *low-mean* regime, the investor chooses to long a futures position at time ν_1 and then closes the position at τ_1. When the regime switches at t_{12}, the VIX locates in the long region in *high-mean* regime. The investor should long one position immediately at t_{12} (ν_2), and liquidate the position later at τ_2 according to the optimal boundaries in regime 2. Figure 3(b) shows another scenario. It is optimal for the investor to wait before the switch happens, since the VIX stays in the waiting region in regime 1. Then the investor would better to long a futures at ν_2 and short at τ_2 as the previous case. Another example is shown in Fig. 3(c). The VIX goes up to hit the short boundary "$\mathcal{K}_1$" at ν_1 in regime 1. The investor chooses to short a position and she speculates that the price will decrease. However, the regime switches at time t_{12} (τ_1, ν_2). The investor should immediately close her position and start to long one position according to the optimal boundaries "$\mathcal{J}_2$" and "$\mathcal{V}_2$". The investor might need to face the loss in the position switching process. In other words, there is a positive probability of losses in a finite time period. We see that the regime-switching timing t_{12} plays a key role for investor's trading decision.

The introduction of regime switching adds considerable complexity to the optimal trading strategies. The last three examples in Fig. 3 are the simplest cases since we assume that the regime switching only happens once at t_{12}. Figures 4 and 5 show another approach to understand our models without specifying the regime-switching times. The process used here still has two states. Optimal boundaries separate the space into five regions. Tables 1 and 2 indicate the exercises in different regimes by given VIX. Figure 4 shows the case of "*low-mean* regime vs *high-mean* regime". This regime-switching case captures some sudden changes of market (such as financial crisis in 2008), which might cause extremely high volatility. Once the regime switches, the investor is expected to take the long position aggressively. We can see the long region is as large as "$S_1^1 + S_2^1 + S_3^1$" in *high-mean* regime. It shows that the investor has to adjust her strategies based on



(a) (b)

(c)

Fig. 3. Simulated CIR paths and exercise times under 2-state regime-switching model. (a) The investor enters at ν_1 and exits at τ_1 in regime 1; enters again at ν_2 and exits at τ_2 in regime 2. (b) The investor enters at ν_2 and exits at τ_2 in regime 2 without any adjustments in regime 1. (c) The investor takes a short position at ν_1 in regime 1 but switches to a long position at $t_{12}(\tau_1, \nu_2)$, and then liquilate her position at τ_2 in regime 2. The parameters are the same as those in Fig. 2.

the different regimes, even though the VIX still stays in the same region. Figure 5 shows two regimes with relatively closer switching means. The optimal boundaries are higher in regime 2 than in regime 1, which means that the investor intends to enter and exit the market earlier in regime 1. Moreover, in regime 1, the waiting region is "$S_2^1 + S_3^1$", while in regime 2, the waiting region is "$S_3^1 + S_4^1$". "S_3^1" is the common waiting region in both regimes. It implies that the investor might participate in tradings more frequently by adopting 2-regime model than single

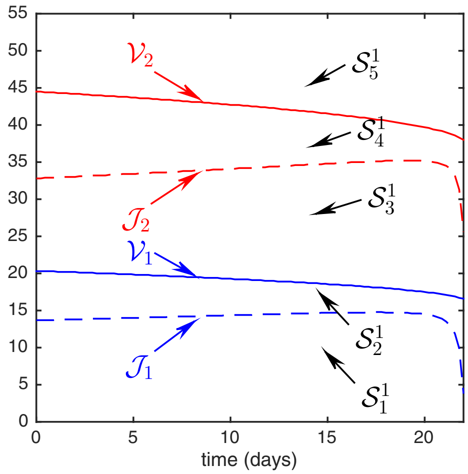


Fig. 4. Optimal boundaries for futures trading under the CIR model with 2-state regime switching. Parameters: $\hat{T} = \frac{22}{252}, T = \frac{66}{252}, r = 0.05, \sigma_1 = 5.33, \sigma_2 = 6.42, \theta_1 = 17.58, \theta_2 = 39.5, \tilde{\theta}_1 = 18.16, \tilde{\theta}_2 = 40.36, \mu_1 = 8.57, \mu_2 = 9, \tilde{\mu}_1 = 4.55, \tilde{\mu}_2 = 4.59, c = \hat{c} = 0.01, q_{12} = -q_{11} = 0.1, q_{21} = -q_{22} = 0.5$.

regime model. In other words, the presence of regime switching will impact investor's trading strategies.

From the perspective of an investor with no position, she is interested in determining the best time to enter the market. We study the *optimal timing premium*

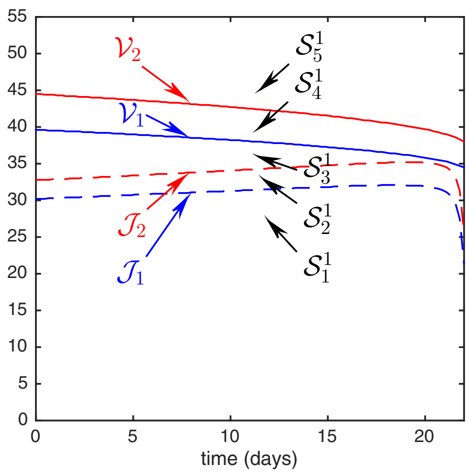


Fig. 5. Optimal boundaries for futures trading under the CIR model with 2-state regime switching. Parameters: $\hat{T} = \frac{22}{252}, T = \frac{66}{252}, r = 0.05, \sigma_1 = 5.33, \sigma_2 = 6.42, \theta_1 = 35.6, \theta_2 = 39.5, \tilde{\theta}_1 = 35.96, \tilde{\theta}_2 = 40.36, \mu_1 = 8.57, \mu_2 = 9, \tilde{\mu}_1 = 4.55, \tilde{\mu}_2 = 4.59, c = \hat{c} = 0.01, q_{12} = -q_{11} = 0.1, q_{21} = -q_{22} = 0.5$.

Table 1.

	In regime 1	In regime 2
S_1^1	Long	Long
S_2^1	Wait	Long
S_3^1	Short	Long
S_4^1	Short	Wait
S_5^1	Short	Short

Table 2.

	In regime 1	In regime 2
S_1^1	Long	Long
S_2^1	Wait	Long
S_3^1	Wait	Wait
S_4^1	Short	Wait
S_5^1	Short	Short

(Leung and Ludkovski, 2011; Leung and Liu, 2012; Leung and Ludkovski, 2012), which plays a vital role in the optimal strategies. This premium expresses the benefit of waiting to enter as compared to initialize the position immediately. Precisely, the premium is defined as

$$L(t,s,i) := \mathcal{P}(t,s,i) - \max\{\mathcal{A}(t,s,i), \mathcal{B}(t,s,i)\}. \tag{4.1}$$

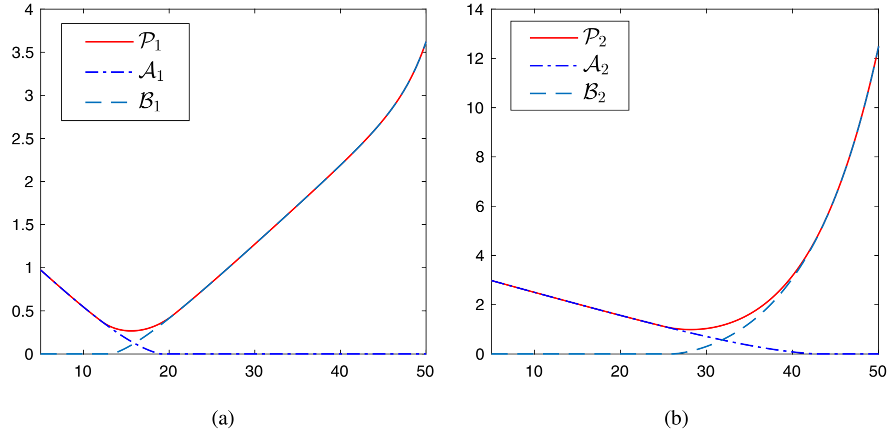


Fig. 6. The value functions $\mathcal{P}$, $\mathcal{A}$, and $\mathcal{B}$ in 2 regimes are plotted against the VIX at time 0. The parameters are the same as those in Fig. 2.

As we can see in (4.1), the optimal stopping time for $L(t, s, i)$ maximizes the expected discounted value from establishing the VIX futures position. Figure 6 shows that $\mathcal{P}$ dominates $\mathcal{A}$ and $\mathcal{B}$. We also note that $\mathcal{P} = \mathcal{A}$ when the VIX is low and $\mathcal{P} = \mathcal{B}$ when the VIX is high in each regime.

5. Conclusion

We extend the optimal VIX futures trading problems under a regime-switching model. This model allows the investor to capture the structural changes on the market. Numerical method is developed to solve these coupled system of variational inequalities that govern the value functions. Accounting for the timing options as well as the option to choose between a long or short position, we find that it is optimal to delay market entry, as compared to the case of committing to either go long or short *a priori*. By introduce of regime-switching mechanism, it is noted that investor should modify her trading strategies correspondingly to regime-switching timings. The strategies and numerical method introduced in this paper can also be applied to other derivatives.

Acknowledgment

The author would like to thank Professor Tim Leung for many helpful and insightful comments and suggestions.

References

Brennan, MJ and ES Schwartz (1990). Arbitrage in stock index futures, *Journal of Business*, 63(1), S7–S31.

Buffington, J and RJ Elliott (2002). American options with regime switching, *International Journal of Theoretical and Applied Finance*, 5, 497–514.

Cartea, A, S Jaimungal and J Penalva (2015). *Algorithmic and High-Frequency Trading*. Cambridge: Cambridge University Press.

Dai, M, Y Zhong and YK Kwok (2011). Optimal arbitrage strategies on stock index futures under position limits, *Journal of Futures Markets*, 31(4), 394–406.

Dotsis, G, D Psychoyios and G Skiadopoulos (2007). An empirical comparison of continuous-time models of implied volatility indices, *Journal of Banking & Finance*, 31(12), 3584–3603.

Elliott, RJ, TK Siu and L Chan (2008). A pde approach for risk measures for derivatives with regime switching, *Annals of Finance*, 4(1), 55–74.

Grübichler, A and F Longstaff (1996). Valuing futures and options on volatility, *Journal of Banking and Finance*, 20(6), 985–1001.

Guo, X (2001). An explicit solution to an optimal stopping problem with regime switching, *Journal of Applied Probability*, 464–481.

Khaliq, AQ and R Liu (2009). New numerical scheme for pricing american option with regime-switching, *International Journal of Theoretical and Applied Finance*, 12(03), 319–340.

Le, H and C Wang (2010). A finite time horizon optimal stopping problem with regime switching, *SIAM Journal on Control and Optimization*, 48(8), 5193–5213.

Leung, T (2010). A Markov-modulated stochastic control problem with optimal multiple stopping with application to finance, in *49th IEEE Conference on Decision and Control (CDC)*, pp. 559–566.

Leung, T, J Li, X Li and Z Wang (2016). Speculative futures trading under mean reversion, *Asia-Pacific Financial Markets*, pp. 1–24.

Leung, T and X Li (2015). Optimal mean reversion trading with transaction costs and stop-loss exit, *International Journal of Theoretical & Applied Finance*, 18(3), 15500.

Leung, T and X Li (2016). *Optimal Mean Reversion Trading: Mathematical Analysis and Practical Applications*. Modern Trends in Financial Engineering. Singapore: World Scientific.

Leung, T, X Li and Z Wang (2014). Optimal starting–stopping and switching of a CIR process with fixed costs, *Risk and Decision Analysis*, 5(2), 149–161.

Leung, T, X Li and Z Wang (2015). Optimal multiple trading times under the exponential OU model with transaction costs, *Stochastic Models*, 31(4), 554–587.

Leung, T and P Liu (2012). Risk premia and optimal liquidation of credit derivatives, *International Journal of Theoretical & Applied Finance*, 15(8), 1250059.

Leung, T and M Ludkovski (2011). Optimal timing to purchase options, *SIAM Journal on Financial Mathematics*, 2(1), 768–793.

Leung, T and M Ludkovski (2012). Accounting for risk aversion in derivatives purchase timing, *Mathematics & Financial Economics*, 6(4), 363–386.

Leung, T and Y Shirai (2015). Optimal derivative liquidation timing under path-dependent risk penalties, *Journal of Financial Engineering*, 2(1), 1550004.

Leung, T and K Yamazaki (2013). American step-up and step-down credit default swaps under Lévy models, *Quantitative Finance*, 13(1), 137–157.

Mencía, J and E Sentana (2013). Valuation of VIX derivatives, *Journal of Financial Economics*, 108(2), 367–391.

Sircar, R and A Papanicolaou (2014). A regime-switching Heston model for VIX and S&P 500 implied volatilities, *Quantitative Finance*, 14(10), 1811–1827.

Stepanek, C (2015). Comparison of commodity future pricing approaches with cointegration techniques, *Journal of Financial Engineering*, 2(01), 1550002.

Wang, Z and RT Daigler (2011). The performance of VIX option pricing models: Empirical evidence beyond simulation, *Journal of Futures Markets*, 31(3), 251–281.

Whaley, RE (2000). The investor fear gauge, *The Journal of Portfolio Management*, 26(3), 12–17.

Wilmott, P, S Howison and J Dewynne (1995). *The Mathematics of Financial Derivatives: A Student Introduction*, 1st edn. Cambridge: Cambridge University Press.

Yao, DD, Q Zhang and XY Zhou (2006). A regime-switching model for European options, in *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems*, pp. 281–300. Springer.

Zhang, JE and Y Zhu (2006). VIX futures, *Journal of Futures Markets*, 26(6), 521–531.