Journal of Economic Dynamics & Control 56 (2015) 1–19

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Costly arbitrage through pairs trading

Yaoting Lei [a], Jing Xu [b,*,1]

[a] Department of Mathematics, National University of Singapore, Block S17, 10 Lower Kent Ridge Road, Singapore 119076, Singapore
[b] Risk Management Institute, National University of Singapore, 21 Heng Mui Keng Terrace, I[3] Building 04-03, Singapore 119613, Singapore

ARTICLE INFO

Article history:
Received 20 November 2014
Received in revised form
2 March 2015
Accepted 23 April 2015
Available online 1 May 2015

JEL classification:
C32
G11
G12

Keywords:
Pairs trading
Risky arbitrage
Delta-neutral strategies
Transaction costs

ABSTRACT

We study the optimal trading policy of an arbitrageur who can exploit temporary mispricing in a market with two convergent assets. We build on the model of Liu and Timmermann (2013) and include transaction costs, which impose additional limits to the implementation of such convergence trade strategy. We show that the presence of transaction costs could reveal an endogenous stop-loss concern in a certain economy, which affects the optimal policy of the arbitrageur in significant ways. Using pairs of dual-listed Chinese stock shares as samples and a pairs trading strategy based on standard deviation of the spread as benchmark, we demonstrate the efficiency of the strategy implied by our model. Several extensions of our model are also discussed.

1. Introduction

Pairs trading is an investment strategy which has been used in the financial industry for decades. An arbitrageur seeks investment opportunity by exploring the relative mispricing between assets which have historically exhibited autoregressive behavior. When the spread diverges from its recent history, the arbitrageur takes a long position in the undervalued asset and finances this position by taking a short position in the overvalued one. If the market moves to a rational equilibrium and the spread converges in the future, the arbitrageur can close both positions to generate a profit.

Besides its popularity in the financial industry, pairs trading has drawn considerable academic attention. Gatev et al. (2006) investigate the empirical performance of a pairs trading strategy, which is based on the historical standard deviation of the spread, in the U.S. equity market. Andrade et al. (2005), Perlin (2009), Bolgün et al. (2010), and Broussard and Vaihekoski (2012), conduct similar empirical tests in Taiwan, Brazilian, Istanbul and Finland markets, respectively. Mori and Ziobrowski (2011) test the same strategy in the U.S. real estate market. Moving beyond the standard deviation, Elliott et al. (2005) propose a pairs trading model in which the spread is modeled as a mean reverting process and propose a trading strategy based on the model's prediction. Do et al. (2006) generalize the approach in Elliott et al. (2005) to the stochastic residual spread. Meanwhile, the concept of co-integration, which is proposed by Engle and Granger (1987), has been closely related to pairs trading. Being

* Corresponding author. Tel.: +65 6516 4163.
E-mail addresses: matleiy@nus.edu.sg (Y. Lei), matxuji@nus.edu.sg (J. Xu).

[1] We thank Tony He (the Editor) and anonymous referees for valuable comments and suggestions. We also thank Luis Goncalves-Pinto and the participants at the 3rd NUS Workshop on Risk & Regulation for helpful discussions which significantly improve the quality of this paper.

http://dx.doi.org/10.1016/j.jedc.2015.04.006
0165-1889/

imperfect substitutes, pairs often exhibit the co-integration effect. Chiu and Wong (2011), Tourin and Yan (2013), and Liu and Timmermann (2013) consider general portfolio choice problems with co-integrated assets.

One of the major limitations of the aforementioned studies is the absence of transaction costs. Being implemented in almost every market, transaction costs can negatively affect the market timing ability and can be an important limit to arbitrage. When trading pairs, the spread is usually not sizable and the transaction costs could reduce the profitability severely. Gatev et al. (2006) note that ignoring the transaction costs can lead to bad market entry decision near the maturity, because the convergence cannot sufficiently compensate the transaction costs bill. Do and Faff (2012) find that in the U.S. equity market, the net return of pairs trading falls to 19–38 bps per month, in contrast with 90 bps per month as documented in Gatev et al. (2006), after subtracting trading costs. Similarly, the optimal pairs trading strategies in Tourin and Yan (2013) and Liu and Timmermann (2013), are practically infeasible due to the heavy transaction costs incurred in continuous trading.

It is worth pointing out that Brennan and Schwartz (1990) and Dai et al. (2011) essentially study the pairs trading of index futures with transaction costs. However, they do not consider the co-integration effect, which can be a salient characteristic of the dynamics of pairs' prices and can have important implications for market timing.

Building on the model of Liu and Timmermann (2013) and Tourin and Yan (2013), we propose a pairs trading model which incorporates both co-integration and transaction costs. We follow Dai et al. (2010) and consider an arbitrageur who maximizes the expected discounted cash-flows within her finite investment horizon. We find that the type of error correction significantly affects the optimal strategy. Specifically, the optimal strategy could either exhibit pure arbitrage incentives or mixed incentives, depending on the signs of the correction coefficients. Most interestingly, the mixed incentives include an endogenous stop loss component, which is not revealed by alternative pairs trading models.

We investigate the performance of our model in the Chinese stock market which consists of the A share market in Shanghai and the H share market in Hong Kong. Some stocks issued by the same companies are dual-listed in both markets. This special market structure sets the stage for our empirical study. By comparing with a benchmark strategy based on the standard deviation of the price spread, we demonstrate the efficiency of the strategy implied by our model.

The rest of this paper proceeds as follows. In Section 2 we propose the co-integration model. In Section 3 we characterize and analyze the optimal trading strategy. The sensitivity analysis of some key parameters is also presented. In Section 4 we conduct empirical tests with a large sample of dual-listed Chinese shares. In Section 5 we consider two extensions of our model. Finally, in Section 6 we conclude.

2. A pairs trading model

2.1. The assets market

In the baseline model we assume a pair of two assets,[2] whose prices P_{1t} and P_{2t} are assumed to be co-integrated and evolve according to the following equations

$$\frac{dP_{1t}}{P_{1t}} = \mu_1\, dt + \sigma_1\, dW_{1t} - \lambda_1 z_t\, dt \tag{1}$$

$$\frac{dP_{2t}}{P_{2t}} = \mu_2\, dt + \sigma_2\, dW_{2t} + \lambda_2 z_t\, dt \tag{2}$$

where $\mu_i + (-1)^i \lambda_i z_t$ denotes the instantaneous expected rate of return of asset i, and σ_i is the instantaneous volatility, $i = = 1, 2$. W_{1t} and W_{2t} are standard Brownian motions defined on a filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, with constant correlation $\rho \in [-1, 1]$. $z_t = \ln(P_{1t}/P_{2t})$ is the difference between the logarithms of the two asset prices, which represents pricing errors.[3] The parameter λ_i captures the speed of convergence of asset i, and we assume that $\lambda_1 + \lambda_2 > 0$. This assumption implies that z_t is stationary and the logarithms of the prices are co-integrated with co-integrating vector $(1, -1)$. Note that z_t induces correlation between the two asset prices even if $\rho = 0$.

Let $y_{it} = \ln P_{it}, i = 1, 2$. Itô's lemma implies that

$$dy_{1t} = \left(\mu_1 - \lambda_1 (y_{1t} - y_{2t}) - \frac{1}{2}\sigma_1^2\right) dt + \sigma_1\, dW_{1t} \tag{3}$$

$$dy_{2t} = \left(\mu_2 + \lambda_2 (y_{1t} - y_{2t}) - \frac{1}{2}\sigma_2^2\right) dt + \sigma_2\, dW_{2t} \tag{4}$$

[2] In our model, the assets can be portfolios, i.e., linear combination of more fundamental assets. However, we will keep using the terminology “asset” in the context when it does not lead to any confusion.

[3] z_t is in fact equivalent to the error correction term in Engle and Granger (1987).

and, given that $z_t = y_{1t} - y_{2t}$, we then have

$$dz_t = \left(\mu_1 - \frac{1}{2}\sigma_1^2 - \mu_2 + \frac{1}{2}\sigma_2^2 - (\lambda_1 + \lambda_2) z_t\right) dt + \sigma_1\, dW_{1t} - \sigma_2\, dW_{2t} \tag{5}$$

Mean reversion of z_t captures the temporary nature of relative mispricing.[4] This leads to recurring arbitrage opportunities, according to which the price differential z_t only spends an infinitesimally small time at zero.[5]

Our asset market is similar to the one in Liu and Timmermann (2013), except that they consider the convergence trading problem as part of a more general portfolio choice problem which includes a market index and a riskless asset, while we focus exclusively on the optimal timing of pairs trades using a delta-neutral strategy. Moreover, different from Liu and Timmermann (2013), we assume that trading in our market incurs proportional transaction costs. In particular, asset i can be purchased at the ask price $(1+K_b^i)P_{it}$ and can be sold at the bid price $(1-K_s^i)P_{it}$, for some constants $K_b^i \in (0,\infty)$ and $K_s^i \in (0,1)$.

We effectively assume that the arbitrage opportunities are both risky and costly to implement. In particular, in the event of deviations in prices between our two co-integrated assets, idiosyncratic risks can lead prices to deviate further before they eventually converge. Moreover, transaction costs further limit the ability to explore price deviations between the two assets.[6]

2.2. The Arbitrageur's timing problem

We consider a risk neutral arbitrageur who is either fully invested in the pair of assets described in the previous section, or not invested at all. When fully invested, the arbitrageur is long on one of the assets and short on the other. The objective of this arbitrageur is to choose the time of entry when the prices of the two assets diverge, and the time of exit when the prices converge so as to maximize the discounted value of the cash-flows she obtains from this long/short strategy.

At any point in time, the arbitrageur can choose one of the following: (i) long asset 1 and short asset 2, (ii) short asset 1 and long asset 2, or (iii) stay out of the asset market. For ease of exposition we refer to case (i) as "long position", case (ii) as "short position" and case (iii) as "flat position". We denote by $\tau_n \in [0,T)$ the points in time at which the arbitrageur enters the asset market, and $\nu_n \in [0,T]$ the points at which she exits the asset market. When the arbitrageur enters the asset market, she can either take a long or a short position. Therefore, we further define the following two events: (i) A_n is the event in which the arbitrageur takes a long position at time τ_n and (ii) B_n is the event in which the arbitrageur takes a short position at time τ_n.

Similar to Dai et al. (2010), we consider the discounted value of the cumulative cash-flows that the arbitrageur obtains from trading on the pair of assets, which we denote by $C_j(0)$. This quantity depends on the initial state. As described above, there are three possible initial states, $j \in \{-1,0,1\}$. If the arbitrageur's initial state is to be flat, then this state is identified by $j=0$, and the discounted value of the cumulative cash-flows she obtains from her long/short strategy is given by

$$\begin{aligned} C_0(t) = &\sum_{n=1}^{\infty} \left[e^{-\beta(\nu_n - t)} (P_{1\nu_n}\alpha_s^1 - P_{2\nu_n}\alpha_b^2) - e^{-\beta(\tau_n - t)} (P_{1\tau_n}\alpha_b^1 - P_{2\tau_n}\alpha_s^2) \right] I_{\{A_n\}} I_{\{\tau_n < T\}} \\ &+ \sum_{n=1}^{\infty} \left[e^{-\beta(\nu_n - t)} (P_{2\nu_n}\alpha_s^2 - P_{1\nu_n}\alpha_b^1) - e^{-\beta(\tau_n - t)} (P_{2\tau_n}\alpha_b^2 - P_{1\tau_n}\alpha_s^1) \right] I_{\{B_n\}} I_{\{\tau_n < T\}} \end{aligned}$$

where $\alpha_s^i = 1 - K_s^i, \alpha_b^i = 1 + K_b^i, I_{\{S\}}$ is the usual indicator function of event S, i.e.,

$$I_{\{S\}} = \begin{cases} 1 & \text{if } S \text{ is true} \\ 0 & \text{if } S \text{ is false} \end{cases}$$

and the discount rate β measures the arbitrageur's time preference. Alternatively, if the initial state of the arbitrageur is to be long, then $j=1$ and the discounted value of the cumulative cash-flows from the pair trading strategy is as follows:

$$\begin{aligned} C_1(t) = &e^{-\beta(\nu_1 - t)} (P_{1\nu_1}\alpha_s^1 - P_{2\nu_1}\alpha_b^2) \\ &+ \sum_{n=2}^{\infty} \left[e^{-\beta(\nu_n - t)} (P_{1\nu_n}\alpha_s^1 - P_{2\nu_n}\alpha_b^2) - e^{-\beta(\tau_n - t)} (P_{1\tau_n}\alpha_b^1 - P_{2\tau_n}\alpha_s^2) \right] I_{\{A_n\}} I_{\{\tau_n < T\}} \\ &+ \sum_{n=2}^{\infty} \left[e^{-\beta(\nu_n - t)} (P_{2\nu_n}\alpha_s^2 - P_{1\nu_n}\alpha_b^1) - e^{-\beta(\tau_n - t)} (P_{2\tau_n}\alpha_b^2 - P_{1\tau_n}\alpha_s^1) \right] I_{\{B_n\}} I_{\{\tau_n < T\}} \end{aligned}$$

[4] The necessary and sufficient condition which leads z_t to revert to zero is $\mu_1 - \frac{1}{2}\sigma_1^2 = \mu_2 - \frac{1}{2}\sigma_2^2$ and it is essentially assumed in Liu and Timmermann (2013). However, we do not pose this constraint here since (1) the convergence of spread in the sampling period could be very slow; (2) neither Engle and Granger (1987) co-integration test nor the augmented Dicky–Fuller co-integration test requires zero mean.

[5] Like Liu and Timmermann (2013), we also study the case of nonrecurring arbitrage opportunities, in which case the difference in prices is also temporary but is eliminated the first time it achieves full convergence, which means $z_t = 0$. The prices of the two risky assets remain identical after they converge. Since our empirical tests are mostly related to the recurring case, we do not present the results of nonrecurring case here to save the space. The model and the results of this analysis are available upon request.

[6] For simplicity, we assume that the prices of these two assets are exogenous to the investor's decisions. For the study on price effect of convergence trades, the readers are referred to Kondor (2009).

If the initial position of the arbitrageur is to be short, then $j=-1$ and the discounted value of the cumulative cash-flows is represented by the following expression:

$$\begin{aligned} C_{-1}(t) &= e^{-\beta(\nu_1-t)}\left(P_{2\nu_1}\alpha_s^2 - P_{1\nu_1}\alpha_b^1\right) \\ &\quad + \sum_{n=2}^{\infty}\left[e^{-\beta(\nu_n-t)}\left(P_{1\nu_n}\alpha_s^1 - P_{2\nu_n}\alpha_b^2\right) - e^{-\beta(\tau_n-t)}\left(P_{1\tau_n}\alpha_b^1 - P_{2\tau_n}\alpha_s^2\right)\right] I_{\{A_n\}} I_{\{\tau_n<T\}} \\ &\quad + \sum_{n=2}^{\infty}\left[e^{-\beta(\nu_n-t)}\left(P_{2\nu_n}\alpha_s^2 - P_{1\nu_n}\alpha_b^1\right) - e^{-\beta(\tau_n-t)}\left(P_{2\tau_n}\alpha_b^2 - P_{1\tau_n}\alpha_s^1\right)\right] I_{\{B_n\}} I_{\{\tau_n<T\}} \end{aligned}$$

In all cases, any non-flat position is closed at maturity T.

We assume that the arbitrageur maximizes the expected value of $C_j(t)$, which implies the following value functions:

$$V_j(P_1,P_2,t) = \sup_{\tau,\nu} \mathbb{E}_t\left[C_j(t)\right]$$

where $j=-1,0,1$ and the cash-flows are discounted to t. Using dynamic programming, it follows that these value functions are governed by the following Hamilton–Jacobi–Bellman (HJB) equations:

$$\begin{cases} \min\left\{-\dfrac{\partial V_0}{\partial t} - \mathcal{L}V_0,\ V_0 - V_1 + P_1\alpha_b^1 - P_2\alpha_s^2,\ V_0 - V_{-1} + P_2\alpha_b^2 - P_1\alpha_s^1\right\} = 0 \\ \min\left\{-\dfrac{\partial V_{-1}}{\partial t} - \mathcal{L}V_{-1},\ V_{-1} - V_0 + P_1\alpha_b^1 - P_2\alpha_s^2\right\} = 0 \\ \min\left\{-\dfrac{\partial V_1}{\partial t} - \mathcal{L}V_1,\ V_1 - V_0 - P_1\alpha_s^1 + P_2\alpha_b^2\right\} = 0 \end{cases} \tag{6}$$

in the region $(t,P_1,P_2) \in [0,T) \times (0,\infty) \times (0,\infty)$, with terminal conditions

$$\begin{cases} V_0(P_1,P_2,T) = 0 \\ V_{-1}(P_1,P_2,T) = \alpha_s^2 P_2 - \alpha_b^1 P_1 \\ V_1(P_1,P_2,T) = \alpha_s^1 P_1 - \alpha_b^2 P_2 \end{cases} \tag{7}$$

where the differential operator $\mathcal{L}$ is defined as follows:

$$\begin{aligned} \mathcal{L}\,\cdot &= \frac{\partial\cdot}{\partial P_1}\left(\mu_1 - \lambda_1 \ln\frac{P_1}{P_2}\right)P_1 + \frac{\partial\cdot}{\partial P_2}\left(\mu_2 + \lambda_2 \ln\frac{P_1}{P_2}\right)P_2 - \beta\cdot \\ &\quad + \frac{1}{2}\sigma_1^2 P_1^2 \frac{\partial^2\cdot}{\partial P_1^2} + \frac{1}{2}\sigma_2^2 P_2^2 \frac{\partial^2\cdot}{\partial P_2^2} + \rho\sigma_1\sigma_2 P_1 P_2 \frac{\partial^2\cdot}{\partial P_1 \partial P_2} \end{aligned}$$

The linear nature of the price dynamics of our two assets allows us to reduce the dimensionality of the problem. In particular, the above value functions are homogeneous in (P_1,P_2), which means that there are functions $U_j(x,t), j=-1,0,1$, such that

$$U_j(x,t) = \frac{1}{P_2} V_j(P_1,P_2,t), \quad x = \frac{P_1}{P_2} \tag{8}$$

and thus the HJB equations can be reduced to

$$\begin{cases} \min\left\{-\dfrac{\partial U_0}{\partial t} - \mathcal{M}U_0,\ U_0 - U_1 + x\alpha_b^1 - \alpha_s^2,\ U_0 - U_{-1} + \alpha_b^2 - x\alpha_s^1\right\} = 0 \\ \min\left\{-\dfrac{\partial U_{-1}}{\partial t} - \mathcal{M}U_{-1},\ U_{-1} - U_0 + x\alpha_b^1 - \alpha_s^2\right\} = 0 \\ \min\left\{-\dfrac{\partial U_1}{\partial t} - \mathcal{M}U_1,\ U_1 - U_0 - x\alpha_s^1 + \alpha_b^2\right\} = 0 \end{cases} \tag{9}$$

in the domain $(t,x) \in [0,T) \times (0,\infty)$, with terminal conditions

$$\begin{cases} U_0(x,T) = 0 \\ U_{-1}(x,T) = \alpha_s^2 - \alpha_b^1 x \\ U_1(x,T) = \alpha_s^1 x - \alpha_b^2 \end{cases} \tag{10}$$

and where the operator $\mathcal{M}$ is defined as

$$\mathcal{M}\,\cdot = (\mu_1 - \mu_2 - (\lambda_1+\lambda_2)\ln x)x\frac{\partial\cdot}{\partial x} + \frac{1}{2}(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)x^2\frac{\partial^2\cdot}{\partial x^2} + (\mu_2 + \lambda_2 \ln x - \beta)\cdot$$

For each time $t \in [0,T)$, the system of redefined HJB equations (9) splits the spacial domain $(0,\infty)$ into the following four trading regions:

$$T_1(t) = \{x : U_0(t,x) = U_1(t,x) - \alpha_b^1 x + \alpha_s^2\}$$
$$T_2(t) = \{x : U_0(t,x) = U_{-1}(t,x) + \alpha_s^1 x - \alpha_b^2\}$$
$$T_3(t) = \{x : U_{-1}(t,x) = U_0(t,x) - \alpha_b^1 x + \alpha_s^2\}$$
$$T_4(t) = \{x : U_1(t,x) = U_0(t,x) + \alpha_s^1 x - \alpha_b^2\}$$

where $T_1(t)$ is the region in which the arbitrageur should enter the asset market with a long position and $T_2(t)$ is the region in which the arbitrageur should enter the asset market but instead with a short position. The exit strategy is determined by the regions $T_3(t)$ and $T_4(t)$. In particular, $T_3(t)$ is the region in which the arbitrageur should exit the asset market from a short position to flat position, and $T_4(t)$ determines the exit from a long position to flat position. Note that these regions can eventually intersect. For instance, within $T_3(t) \bigcap T_1(t)$, the arbitrageur should switch directly from a short position to a long position.

In other words, the optimal pair trading strategy can be simply described in the following example. Assume that the arbitrageur starts with flat position on the asset market. The first time the price ratio x hits the boundary or the interior of the region $T_1(t)$, then the arbitrageur should enter the asset market with a long position, and follow the strategy induced by U_1. Next, when x hits the boundary or the interior of the region $T_4(t)$, she should exit the asset market by closing the long position.[7]

The following proposition indicates that arbitrage opportunity vanishes as the maturity approaches:

Proposition 2.1. *For each $x = P_1/P_2 > 0$, there exist $\epsilon(x) > 0$ such that when $T - \epsilon(x) < t < T$, we have*

$$U_0(t,x) > U_1(t,x) - x\alpha_b^1 + \alpha_s^2, \tag{11}$$

$$U_0(t,x) > U_{-1}(t,x) - \alpha_b^2 + x\alpha_s^1. \tag{12}$$

The proof is placed in A.2. The economic intuition is that, in the remaining short time period, the price convergence is insufficient to compensate the transaction costs bill. As a result, the arbitrageur should not enter the asset market near the maturity. This fact is noted (but not resolved) in Gatev et al. (2006). Our model, however, endogenously avoids this kind of unprofitable trade.

2.3. Remarks on the assumptions

In our baseline model, we have made several simplified assumptions, including

1. *The assets prices have continuous paths*: This is a direct consequence of modelling the assets prices as diffusions. However, asset prices can exhibit abrupt jumps, and jump-diffusion models might be more appropriate. Including jumps in the model is straightforward and is discussed in Section 5.1.

2. *The co-integrating vector is assumed to be* $(1,-1)$: Generally, the co-integrating vector can be $(1,\theta)$ for some constant θ and the associated stationary residual is given by $z_t = \alpha + \gamma t + \log P_1 + \theta \log P_2$, cf., Duan and Pliska (2004). Our model, as well as that in Liu and Timmermann (2013), is a special case with $\alpha = \gamma = 0$ and $\theta = -1$. This restriction is essential in the dimensional reduction (8). Although extending to the general case is straightforward, the three dimensional system of variation inequalities (6) cannot be reduced and the computational demand substantially increases, which renders the tests with a large sample impractical. The extension to the general co-integration case is discussed in Section 5.2.

3. Numerical analysis

3.1. Optimal trading strategies

Without available analytic solution, we numerically solve (9), using the penalty method and finite differences scheme (cf. Forsyth and Vetzal, 2002; Dai et al., 2007). In this section we characterize the optimal trading strategy.

The baseline parameter values for our numerical calculation are summarized in Table 1. For λ_1 and λ_2, we follow Liu and Timmermann (2013) and consider two possibilities: $\lambda_1 > 0, \lambda_2 > 0$ or $\lambda_1 > 0, \lambda_2 < 0$.[8] We refer to the first as Economy 1 and to the second as Economy 2.

[7] In our base model with continuous sample paths, the trades should, theoretically, always take place on the boundaries. However, in practice, trades are necessarily executed in discrete time, therefore the trades could possibly take place on the interiors. In addition, if jumps are introduced to the model, the trades could also take place on the interiors.

[8] The case with $\lambda_1 < 0$ and $\lambda_2 > 0$ is similar due to symmetry.

Table 1
Baseline parameter values.
This table provides the baseline parameter values that are used in the numerical calculation, whenever applicable.

General environment	
Discount rate	$\beta = 0.04$
Characteristics of Asset 1	
Expected rate of return	$\mu_1 = 0.10$
Instantaneous volatility of return	$\sigma_1 = 0.20$
Speed of convergence	$\lambda_1 = 0.35$ (Economy 1), 1 (Economy 2)
Transaction cost for purchase	$K_b^1 = 0.5\%$
Transaction cost for sale	$K_s^1 = 0.5\%$
Characteristics of Asset 2	
Expected rate of return	$\mu_2 = 0.112$
Instantaneous volatility of return	$\sigma_2 = 0.24$
Speed of convergence	$\lambda_2 = 0.25$ (Economy 1), -0.4 (Economy 2)
Transaction cost for purchase	$K_b^2 = 0.5\%$
Transaction cost for sale	$K_s^2 = 0.5\%$
Interaction between Assets 1 and 2	
Correlation between the returns	$\rho = 0.8$

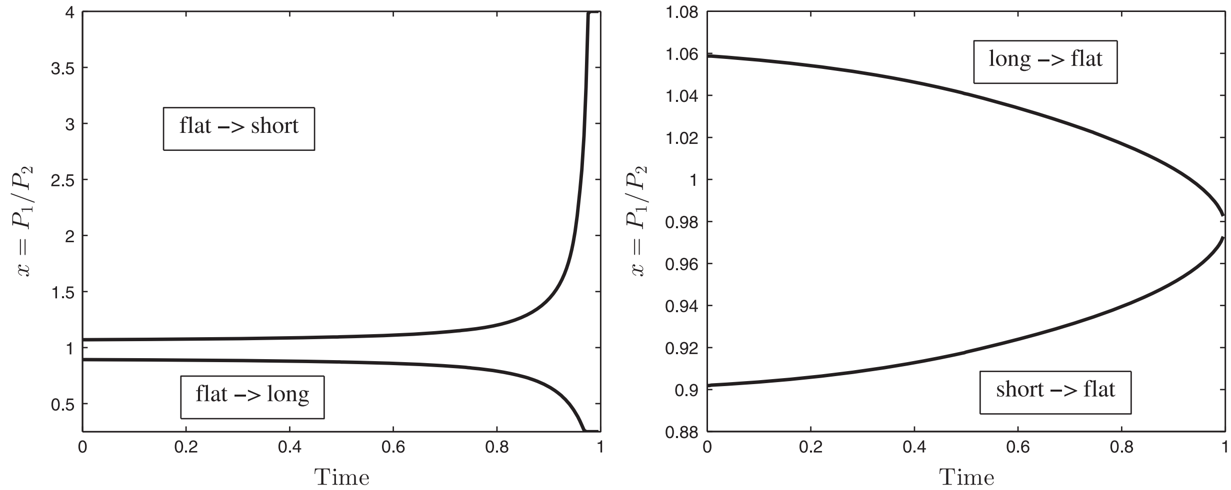

Fig. 1. Optimal trading boundaries in Economy 1. This figure depicts the optimal trading boundaries against time t in Economy 1. The baseline parameter values we use are as follows: $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, K_b = K_s = 0.5\%, \lambda_1 = 0.35, \lambda_2 = 0.25$.

Economy 1. Fig. 1 depicts the arbitrageur's optimal trading policy in Economy 1. The two time varying curves in the left panel determine the arbitrageur's optimal timing of entering the asset market. When the ratio $x = P_1/P_2$ lies above the upper curve, the arbitrageur should enter the asset market with a short position (as signified by flat $\rightarrow$ short). Symmetrically, if x lies below the lower curve, the arbitrageur should enter the asset market with a long position (as signified by flat $\rightarrow$ long). Similarly, the two time varying curves in the right panel determine the arbitrageur's optimal timing of exiting the asset market. The arbitrageur with a long position should close her position to lock in the profit when x lies above the upper curve (as signified by long $\rightarrow$ flat), and the arbitrageur with a short position should close her position when x lies below the lower curve (as signified by short $\rightarrow$ flat).

The time varying nature of trading boundaries is directly due to the presence of transaction costs and the finite horizon. The numerical result suggests that when time approaches the maturity, the upper curve in left panel approaches infinity and the lower curve approaches 0, which indicates vanishing future trading opportunity and is consistent with Proposition 2.1.

In summary, the optimal trading strategy in Economy 1 exhibits pure arbitrage incentives. When the spread reaches certain threshold, the arbitrageur takes a long position in the undervalued asset and finances her position by taking a short position in the overvalued one. She closes the position when the convergence is sufficient to cover the transaction costs and waits for the next opportunity of opening position.

Economy 2. Negative value of λ_2 in Economy 2 implies that when the prices diverge, P_2 tends to diverge even further, instead of instantaneous converging. As a result, the situation in Economy 2 is more complicated. Fig. 2 depicts the arbitrageur's optimal trading policy in Economy 2.

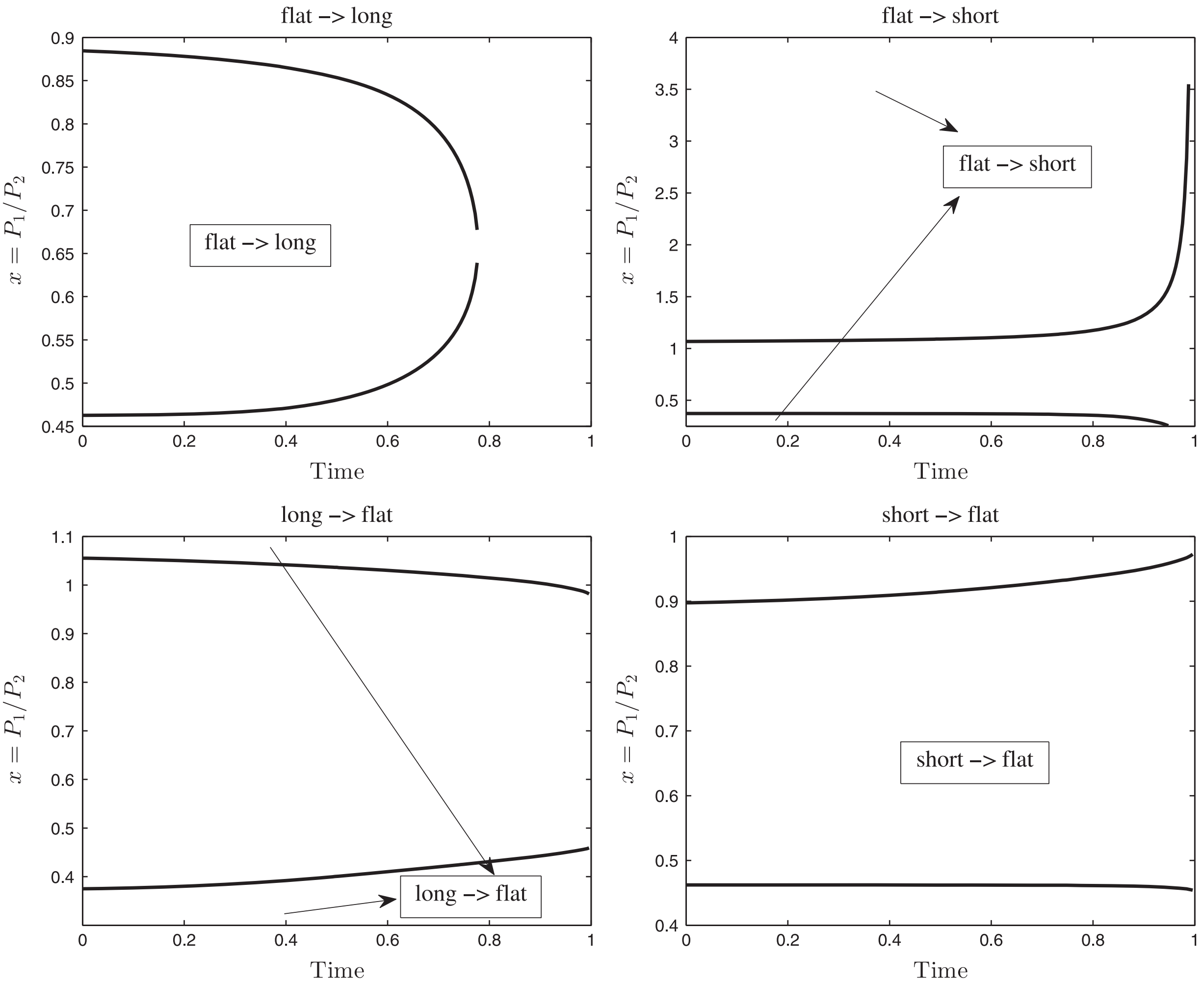

Fig. 2. Optimal trading boundaries in Economy 2. This figure depicts the optimal trading boundaries against time t in Economy 2. The baseline parameter values we use are as follows: $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, K_b = K_s = 0.5\%, \lambda_1 = 1, \lambda_2 = -0.4$.

The two time varying curves in the upper-left panel determine the optimal timing of entering the asset market with a long position. The arbitrageur should enter the market with a long position when the ratio x lies between the two curves. The existence of the lower curve indicates that the arbitrageur should not take long position on asset 1 when it is deeply undervalued. This phenomenon does not appear in Economy 1. Similarly, the two time varying curves in the upper-right panel determine optimal timing of entering the asset market with a short position. The arbitrageur should enter the market with a short position when the ratio x lies either above the upper curve or below the lower curve. This feature indicates that the arbitrageur should short asset 1 when it is sufficiently undervalued. The reason is that, by longing asset 2, the manager could take the advantage of the temporal divergence and gain more profit.

The curves in the lower-left panel and the lower-right panel of Fig. 2 determine when the arbitrageur should exit the asset market. The optimal actions are signified accordingly. An interesting observation is that, the lower-left panel shows that the co-integration model implies stop loss strategy in Economy 2. When asset 1 is deeply undervalued, the arbitrageur should immediately sell it instead of waiting for the convergence. The reason is that since the manager concerns about her cash-flows within the finite horizon, she will not let asset 2 diverge further. The stop loss component is absent in Economy 1 where the prices are mutually convergent.

3.2. Comparative statics

Now we explore how the optimal policies change with parameter values. We focus on the characteristics of the assets, such as the correlation between asset returns and the speed of convergence. The macroeconomic variables, such as the discount rate and transaction costs rate, are left unchanged.

Change in correlation: Figs. 3 and 4 plot the optimal trading boundaries against time in both economies, with different correlation parameter ρ. We observe that various transaction regions (including both entering and existing) are widened

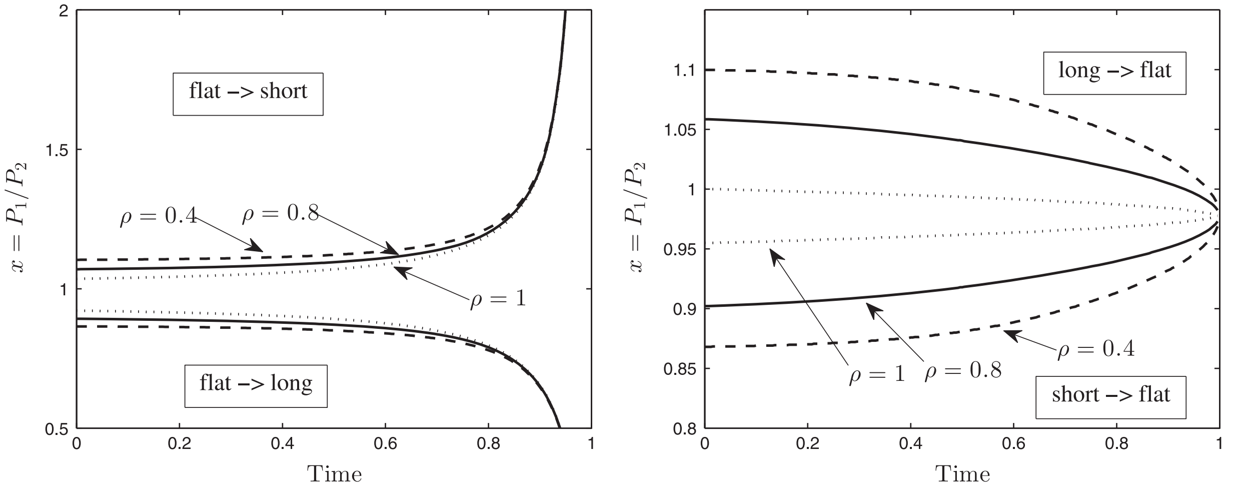

Fig. 3. Optimal trading bounds for various ρ in Economy 1. This figure depicts the optimal trading boundaries against time t in Economy 1 for different correlation ρ=0.4, 0.8 and 1. Other parameter values we use are $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, K_b = K_s = 0.5\%, \lambda_1 = 0.35, \lambda_2 = 0.25$.

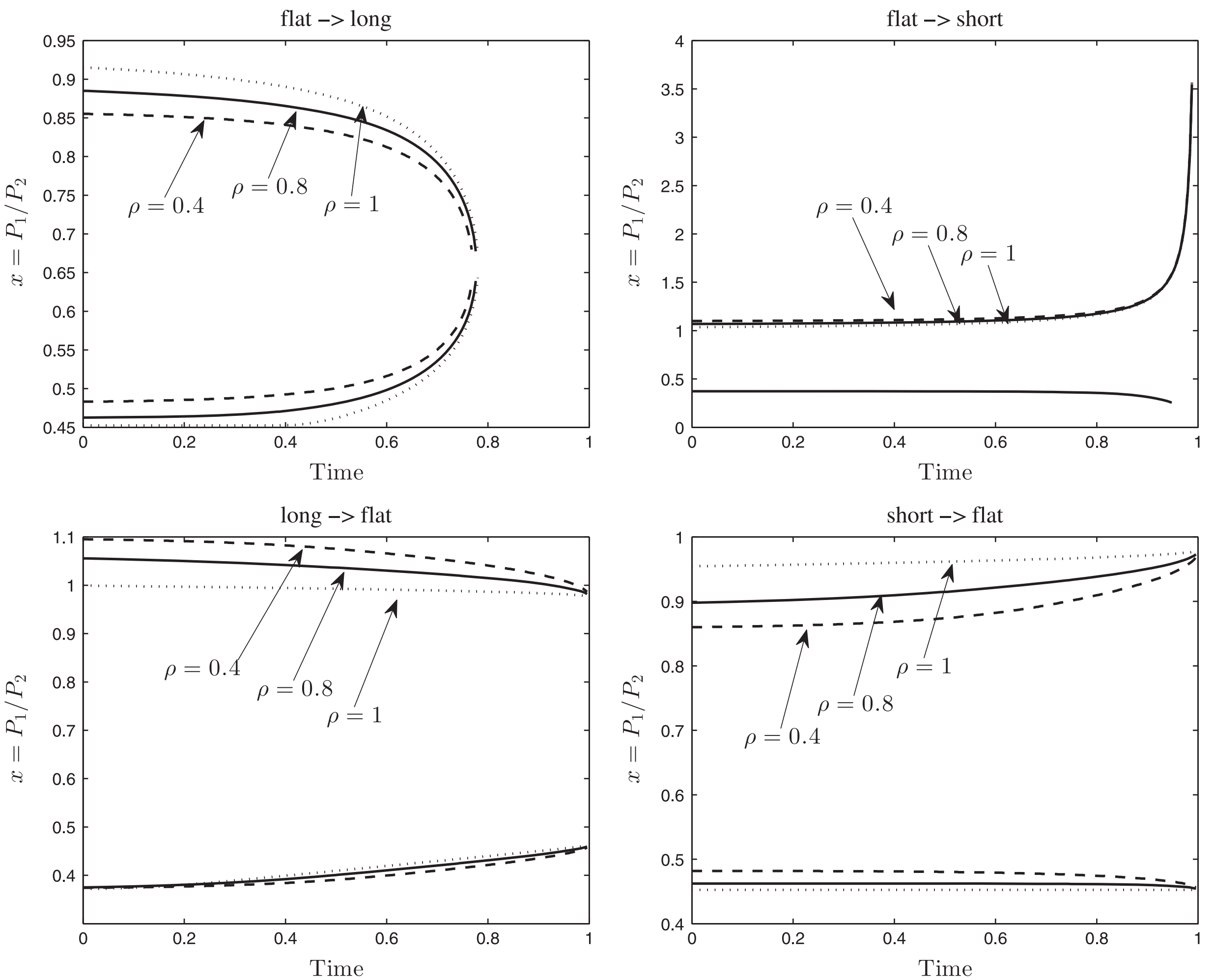

Fig. 4. Optimal trading boundaries for various ρ in Economy 2. This figure depicts the optimal trading boundaries against time t in Economy 2 for different correlation ρ=0.4, 0.8 and 1. Other parameter values we use are $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, K_b = K_s = 0.5\%, \lambda_1 = 1, \lambda_2 = -0.4$.

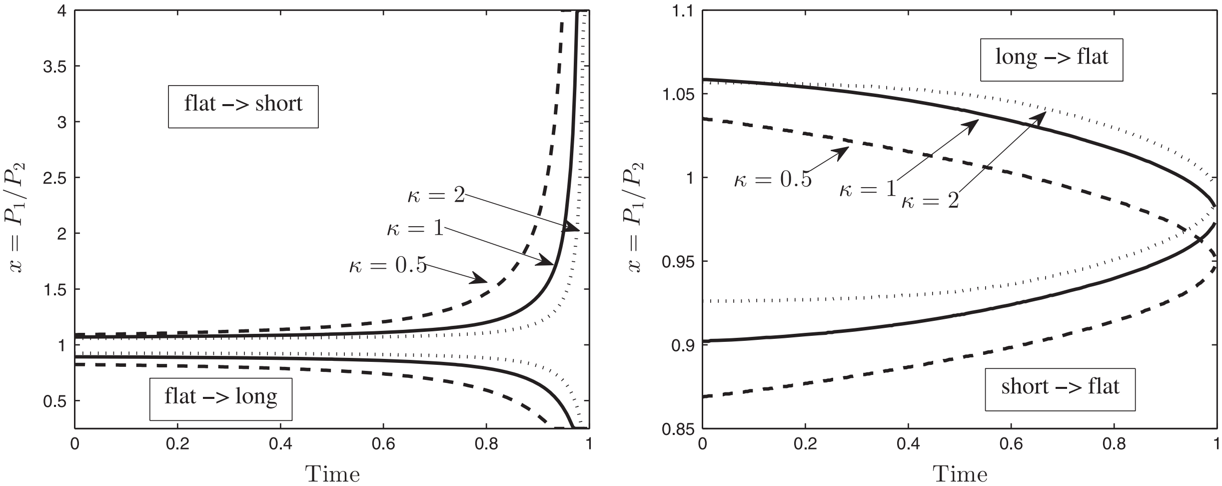

Fig. 5. Optimal trading boundaries for various κ in Economy 1 This figure depicts the optimal trading boundaries against time t in Economy 1 for different speed of convergence $\kappa = 0.5$, 1 and 2. Other parameter values we use are $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, K_b = K_s = 0.5\%$, $\lambda_1 = 0.35\kappa, \lambda_2 = 0.25\kappa$.

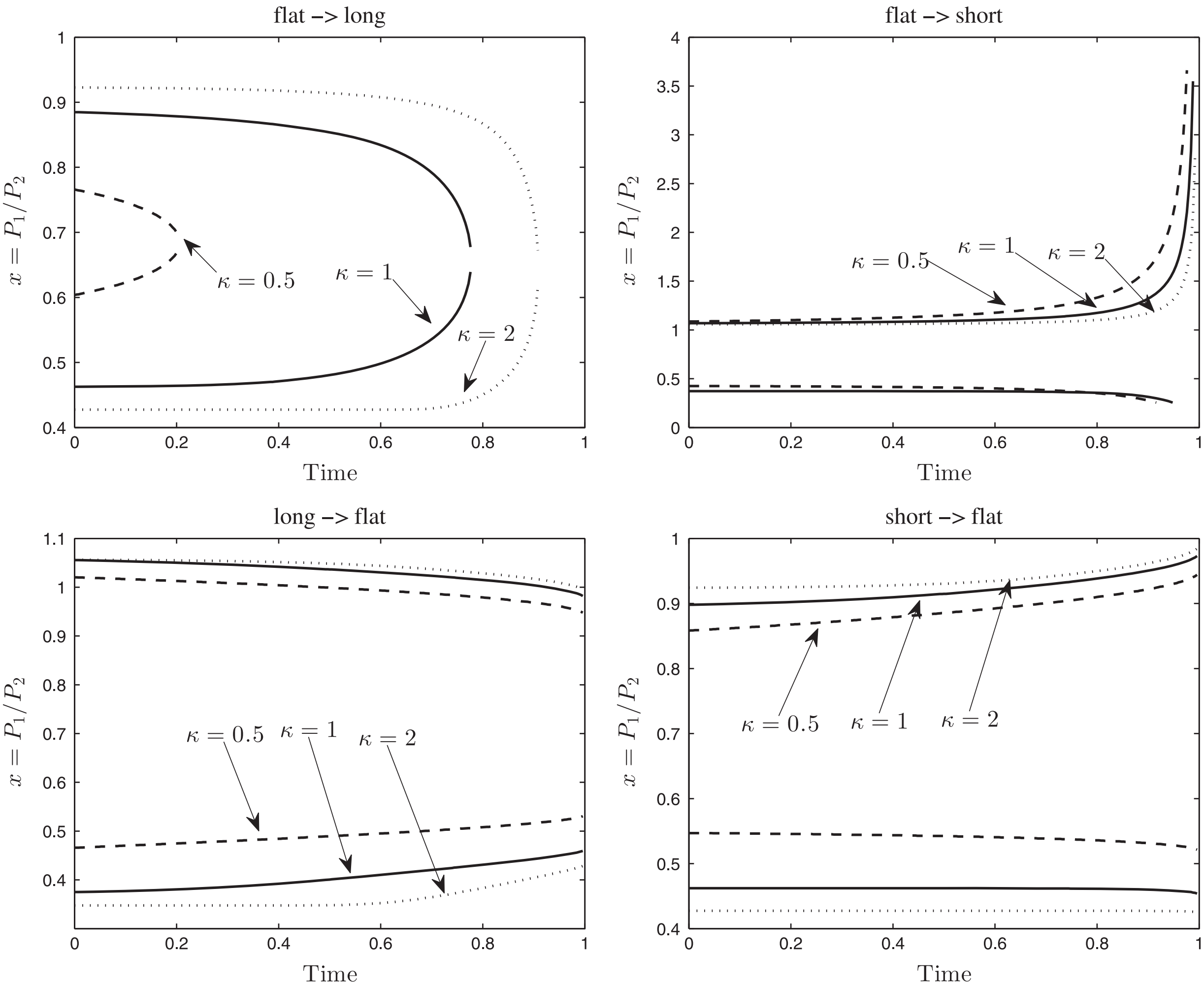

Fig. 6. Optimal trading boundaries for various κ in Economy 2. This figure depicts the optimal trading boundaries against time t in Economy 2 for different speed of convergence $\kappa = 0.5$, 1 and 2. Other parameter values we use are $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, K_b = K_s = 0.5\%$, $\lambda_1 = \kappa, \lambda_2 = -0.4\kappa$.

when the correlation increases. This pattern indicates that high correlation introduces more trading opportunities. Intuitively, lower correlation creates more idiosyncratic risk which cannot be hedged away by taking a long/short position, and hence reduces the arbitrage incentives.

Change in the speed of convergence: Figs. 5 and 6 plot the optimal trading boundaries against time in both economies, with different convergence speed κ.[9] It is observed that when the speed of convergence increases, the transaction regions in the left panel of Fig. 5, in the upper-left, and in the upper-right panel of Fig. 6, all widen. This pattern suggests that fast convergence creates more future arbitrage opportunities, and hence strengthens the arbitrageur's incentives of opening position.

However, the incentives of closing position are not uniformly strengthened. The right panel of Fig. 5, the lower-left panel of Fig. 6 and the lower-right panel of Fig. 6 reveal that the arbitrageur exhibits less incentives to close long position and more incentives to close short position when the speed of convergence increases.

4. Empirical tests in the Chinese stock market

Moving beyond the numerical examples presented in Section 3, we test our model with real market data in this section. We introduce the samples and explain our methodology first, and then present the results and discussion.

4.1. Samples and relevant data

Pairs formation is the first step of implementing pairs trading strategy. Fundamental analysis and distance metrics are two popular approaches to form the pairs (cf. Vidyamurthy, 2004). In this study we employ a combined approach. Specifically, we follow Liu and Timmermann (2013) and consider the dual-listed companies in China. These companies are listed in both Shanghai Stock Exchange (SSE) and Hong Kong Stock Exchange (HKEx), and their shares are separately traded in each exchange.[10] Since these stocks entitle their holders similar dividends and voting rights, there are economic reasons to expect that co-integration effect exists between them. At the same time, these stocks are usually traded at spread and provide arbitrage opportunities which are risky to implement.[11] Fig. 7 shows the time series plots of six bank shares' prices. It can be seen that though the pairs' prices tend to move together, the spread can be quite substantial at times.

Given a pair of dual-listed stocks, A_1 and A_2, we construct the portfolio P_1 and P_2 whose price processes are assumed to follow (1) and (2). We choose $\delta > 0$ which minimizes the L^2 distance between A_1 and δA_2 during the in sample period (see next section). We then choose $P_1 = A_1$ and $P_2 = \delta A_2$ as the pair, and each trade in the tests will involve 100 shares of P_1 and P_2.[12]

The sample stocks are selected from four major sectors: Transportations, Industrial, Utilities and Financials. Price data are typically available at daily frequency. A number of representative samples are picked from each sector to implement the pairs trading strategy. We list the samples' information in A.4 for reference.

The transaction costs currently prevailing in A share and H share market are about 0.7% and 0.36%, respectively, due to different market liquidity and taxation policies. Therefore we set $K_b^1 = K_s^1 = 0.7\%$ and $K_b^2 = K_s^2 = 0.36\%$.[13] In addition, the subjective discount rate β is set at 0.02 unless otherwise stated.

4.2. Methodology

We follow the procedure used in Gatev et al. (2006). Specifically, we construct the portfolio and estimate model parameters in-sample, then test the model implied strategy out-of-sample. For each sample, we denote by T_0 the starting time of available data.

The benchmark strategy: The benchmark strategy is analogous to the relative-value arbitrage strategy tested in Gatev et al. (2006), which is based on standard deviation of the spread. Specifically, we first calculate the in-sample historical standard deviation of the spread σ_h, using the data from T_0 to T_0+1. In the following testing period from T_0+1 to $T_0+1.5$ we open a position when the spread exceeds $2\sigma_h$ and close it when either price converges or the maturity is reached. This procedure generates the cash-flows in the first testing period. Then we move forward and re-calculate the historical standard deviation using the data from $T_0+0.5$ to $T_0+1.5$, and implement the derived strategy in the period from $T_0+1.5$ to T_0+2. The same procedure is repeated through the entire sampling period.

The co-integration strategy: To compare with the benchmark, we implement the co-integration pairs trading model in a similar manner. In each "in-sample + out-of-sample" period, we use the in-sample data to estimate the vector of

[9] Here κ is the multiplier attached to λ_1 and λ_2. In this study we set $\kappa = 0.5$, 1 and 2, respectively.

[10] The shares traded in SSE are called A shares, while the shares traded in HKEx are called H shares.

[11] There are several explanations for the existence of such spread. For example, Hong Kong investors generally enjoy more investment opportunities than mainland investors due to higher market maturity and looser financial regulation. Therefore Hong Kong investors exhibit less demand for the shares issued by mainland companies. At the same time, some companies employ different dividend policies in two markets, which also affect the prices of the same stock in different markets.

[12] A shares and H shares are quoted in Chinese yuan (CNY) and in Hong Kong dollar (HKD), respectively. We convert them to the U.S. dollar (USD) using the historical exchange rates, which are obtained from the website of the board of governors of the federal reserve system: http://www.federalreserve.gov/releases/h10/hist/default.htm.

[13] In reality, the transaction costs are typically time varying, due to the change of market liquidity and transaction policies. However, we assume the transaction costs are constant in the empirical study to simplify the calculation and to avoid the ambiguity.

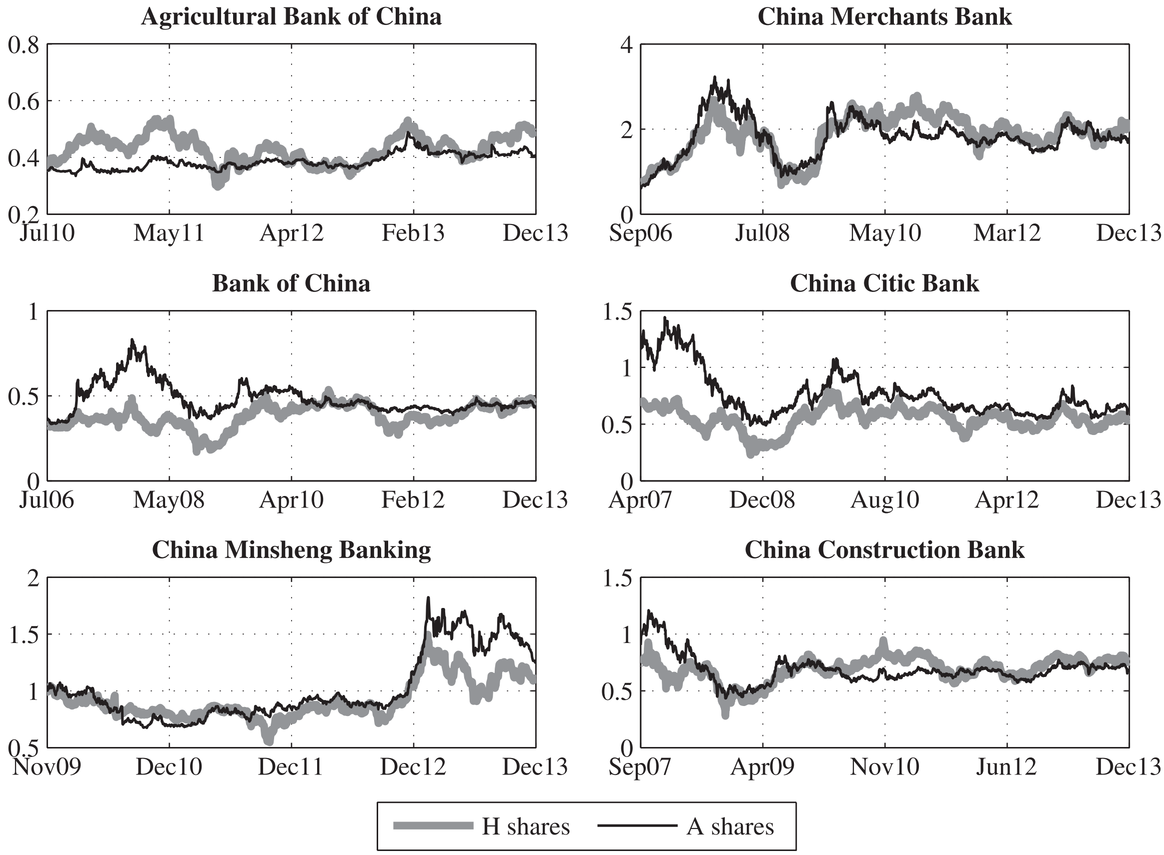

Fig. 7. Price series of six sample stocks. This figure depicts the historical prices of six sample stocks, all measured in the U.S. dollar. The thick gray lines are the H share prices in the Hong Kong Stock Exchange, while the black lines are the A share prices in the Shanghai Stock Exchange.

parameters $\Theta = (\mu_1, \sigma_1, \lambda_1, \mu_2, \sigma_2, \lambda_2, \rho)$ by a quasi-maximum likelihood method, which is briefly explained in A.3. The estimate Θ^* is then fed into the model to generate a trading strategy in the following half year, which is implemented with the out-of-sample data. The same procedure is again repeated through the entire sampling period.

Measuring the profitability and riskiness: In order to measure the profitability, we calculate the accumulated P&L at the end of sampling period for each sample. This quantity represents the net profit generated by the strategy, assuming zero initial investment. In order to measure the riskiness, we track the maximal loss of the arbitrageur's account during the entire testing period. Since the maximal loss is closely related to margins or collateral requirements, it serves as a reasonable risk indicator. In the calculation we assume that the cashes grow at constant rate of 2%, which is a typical value used in the literature. For a given strategy π, we denote the total P&L by $P\&L_\pi$, and the maximal loss by L_π. Regarding two trading strategies $\pi(A)$ and $\pi(B)$, we define $\pi(A)$ to be superior to $\pi(B)$ in the profit-only sense if $P\&L_{\pi(A)} > P\&L_{\pi(B)}$, and we define $\pi(A)$ to be superior to $\pi(B)$ in the profit-risk sense if $P\&L_{\pi(A)}/|L_{\pi(A)}| > P\&L_{\pi(B)}/|L_{\pi(B)}|$.

4.3. Results and discussion

In this section we report the main results of the empirical tests. It is found that the performance of pairs trading strategies could vary wildly across individual samples (see Tables 2 and 3). To smooth out the variations and ease the exposition, we aggregate the individual results in each sector to form the sectional performance. Specifically, we calculate the aggregate P&L and the aggregate maximal loss within each sector.[14] Tables 2 and 3 report the results.[15]

[14] Strictly speaking, at the portfolio level, the sum of maximal loss of individuals is not equal to the maximal loss of the entire portfolio during certain period, since these maximal losses for individuals occur at different points of time. We only use the total maximal loss as a simple approximation to assess the sectional riskiness.

[15] In Tables 2 and 3 we use abbreviations to indicate the issuing companies' names. The corresponding full names could be found in Table A1 in Appendix 2. The profits on individual stock reported in Tables 2 and 3 are normalized by the number of testing periods. For example, in the case of Air China (AC), using the co-integration strategy, the normalized cash profit 35.74 dollars implies that the total net profit during the entire trading period is $35.74 \times 12 = 428.88$ dollars, with 0 initial investment.

Table 2
Empirical performance.

This table provides the results on the empirical performance of the standard deviation strategy (Stdev. Strategy) and of the co-integration strategy (Co-int. Strategy). The column " of Periods" reports the number of testing periods of each sample. The column "Normalized Cash Profit (Maximal Loss)" reports the aggregate P&L generated by each strategy, normalized by the number of testing periods, and the maximal loss during the entire testing period in the parenthesis. The entries in the row labeled "PL Ratio" are the ratio of the total profit to the absolute value of the total loss. In the calculation we assume the cashes grow at constant interest rate $r=2\%$. The first trading dates are different across samples. For a certain sample, the first trading date is the day one year after the beginning of the sampling period of that sample. The sampling periods are presented in Table A1.

Sample company	of Periods	Normalized cash profit (maximal loss)			
		Stdev. strategy		Co-int. strategy	
A. Transportations					
AC	12	2.13	(−77.39)	35.74	(−86.05)
CEA	13	6.02	(−4.79)	−1.89	(−101.00)
CHS	10	6.27	(0.00)	3.46	(0.00)
COSCO	11	3.60	(0.00)	19.80	(0.00)
CSA	13	−5.07	(−190.17)	−0.45	(−212.83)
DLP	4	1.30	(−2.88)	5.86	(0.00)
GSR	12	4.14	(0.00)	11.96	(0.00)
Total		195.16	(−275.24)	817.69	(−399.89)
P-L ratio		0.71		2.04	
B. Industrial					
ACCL	13	67.42	(0.00)	82.48	(0.00)
AGS	13	2.30	(−152.95)	11.68	(−166.46)
CCE	9	9.06	(0.00)	−10.72	(−96.53)
COFS	10	9.75	(−110.86)	25.78	(−7.16)
CPC	13	0.53	(−138.06)	1.63	(−161.62)
CQIS	11	8.86	(0.00)	11.71	(0.00)
CSR	8	2.31	(−18.15)	1.69	(−27.67)
DFE	13	132.40	(0.00)	206.74	(0.00)
DTIPG	12	0.78	(−7.77)	13.76	(0.00)
GZAG	1	0.00	(0.00)	9.72	(0.00)
HDPI	13	1.28	(−17.51)	8.05	(−12.69)
HNPI	13	−1.59	(−64.00)	9.24	(−89.19)
JX	13	12.26	(−232.56)	22.67	(−318.32)
MASIS	13	1.15	(−49.29)	3.31	(−68.43)
NEED	13	1.75	(−39.86)	3.34	(−39.78)
NJPE	13	17.55	(0.00)	19.65	(0.00)
PC	10	2.06	(−15.44)	9.80	(−0.03)
SDXHP	13	9.97	(0.00)	11.48	(0.00)
SHEG	13	−1.67	(−333.74)	23.44	(−380.58)
SHPH	3	30.11	(0.00)	27.62	(−5.71)
SSHP	13	−3.32	(−99.53)	11.92	(−49.14)
TJCEPG	13	8.38	(0.00)	7.36	(0.00)
XJGWST	4	−2.29	(−15.30)	9.26	(0.00)
YZCM	13	33.33	(−43.64)	29.21	(−78.96)
ZJMG	9	6.11	(0.00)	16.93	(0.00)
ZTE	13	40.61	(−142.37)	96.29	(−324.20)
Total		4522.27	(−1481.03)	7829.74	(−1826.47)
P-L ratio		3.05		4.29	

For example, comparing the two strategies in the Transportations sector, we find that the co-integration strategy generates total profit of 817.69 dollars, which is much larger than the total profit, 195.16 dollars, generated by the benchmark standard deviation strategy. At the same time, the total maximal loss experienced by the co-integration strategy is 399.89 dollars. As a result, the co-integration strategy records a profit-to-loss ratio of 2.04. Similarly, the standard deviation strategy records a profit-to-loss ratio of 0.71. This comparison clearly indicates that the co-integration strategy outperforms the standard deviation strategy in this sector in both profit-only and profit-risk senses. The analysis of other tables can be analogously interpreted.

We summarize in Table 4 the preferred strategy, classified by the sector and the judging criteria. We find that in the profit-only sense, the co-integration strategy outperforms the benchmark in all four sectors. It turns out that the co-integration strategy could produce significantly higher net profit at the end of the trading period. Sometimes the improvement could be very substantial. For example, in Transportations and Financials sectors, the co-integration model generates profit that is three times larger than the profit generated by the benchmark strategy. Thus, if the arbitrageur could endure and carry the intermediate loss, the co-integration strategy will always generate the highest profit in the end. On the other hand, in terms of maximal loss, we find that the loss experienced by co-integration strategy and the benchmark strategy are roughly at comparable levels. Therefore, in the profit-risk sense, the co-integration strategy is still preferred in three sectors: Transportations, Industrial and Financials.

Table 3
Empirical performance (Continued).

This table provides the results on the empirical performance of the standard deviation strategy (Stdev. strategy) and of the co-integration strategy (Co-int. strategy). The column " of Periods" reports the number of testing periods of each sample. The column "Normalized Cash Profit (Maximal Loss)" reports the aggregate P&L generated by each strategy, normalized by the number of testing periods, and the maximal loss during the entire testing period in the parenthesis. The entries in the row labeled "PL Ratio" are the ratio of the total profit to the absolute value of the total loss. In the calculation we assume the cashes grow at constant interest rate $r=2\%$. The first trading dates are different across samples. For a certain sample, the first trading date is the day one year after the beginning of the sampling period of that sample. The sampling periods are presented in Table A1.

Sample company	of periods	Normalized cash profit (maximal loss)			
		Stdev. strategy		Co-int. strategy	
C. Financials					
ABC	4	1.19	(−0.39)	1.16	(−6.44)
BOC	12	0.23	(−16.07)	1.29	(−26.59)
BOCOM	11	7.89	(0.00)	12.94	(0.00)
CCB	10	0.41	(−18.80)	9.50	(0.00)
CITIC	11	7.53	(0.00)	11.22	(0.00)
CITICS	2	12.69	(0.00)	27.84	(0.00)
CLI	11	1.09	(−22.12)	11.73	(−125.74)
CMB	12	−5.08	(−90.65)	5.29	(−139.87)
CMSB	6	−2.98	(−38.34)	1.67	(−10.72)
CPIC	6	−3.05	(−18.29)	−0.09	(−58.12)
HTS	1	2.19	(0.00)	−10.04	(−10.04)
ICBC	12	1.16	(0.00)	5.92	(0.00)
NCLI	1	0.00	(0.00)	77.12	(0.00)
PAIC	11	13.65	(−205.88)	19.56	(−291.82)
Total		283.42	(−410.55)	944.92	(−669.34)
P-L ratio		0.69		1.41	
D. Utilities					
AHE	13	2.67	(−10.76)	4.87	(−36.26)
CCC	1	0.00	(0.00)	7.29	(0.00)
CML	6	1.84	(0.00)	0.95	(−8.71)
CRC	9	1.43	(−4.05)	2.44	(0.00)
CRG	10	3.54	(0.00)	0.79	(0.00)
JSE	13	7.70	(0.00)	19.17	(0.00)
SCE	6	−1.09	(−17.83)	1.01	(−24.04)
SZE	13	3.67	(−24.28)	1.98	(−49.82)
Total		235.14	(−56.91)	387.05	(−118.83)
P-L ratio		4.13		3.26	

Table 4
The preferred strategy.

This table provides the results on the preferred pairs trading strategy in our empirical tests, assessed by profit-only criteria or profit-risk criteria, respectively. The meanings of the abbreviations in the table are as follows: Co-int. Strat.: co-integration strategy; Stdev. Strat.: standard deviation strategy.

Sector/criteria	**Preferred strategy**	
	Profit-only	**Profit-risk**
Transportations	Co-int. Strat.	Co-int. Strat.
Industrial	Co-int. Strat.	Co-int. Strat.
Financials	Co-int. Strat.	Co-int. Strat.
Utilities	Co-int. Strat.	Stdev. Strat.

4.4. Robustness check against β

In this section we conduct a robustness check for the co-integration model. The unique subjective parameter in the co-integration model is the arbitrageur's discount rate β, which is set equal to the risk free interest rate in the baseline of the empirical tests. Since changing β would result in the changes of trading boundaries, one might question that whether this would significantly change the performance of the co-integration model. To address this issue, we conduct the otherwise identical empirical tests for various values of β, within a reasonable range. Specifically, besides the baseline $\beta=0.02$, we also set $\beta=0.001$ and $\beta=0.1$, to check the robustness against β.[16]

Table 5 shows that when we vary β from 0.001 to 0.1, the performance on both P&L and the maximal loss is almost unchanged. The only exception is that when $\beta=0.1$, the average P&L slightly increases by 0.96 dollars. The reason is that the

[16] To save the space, we report the results of the samples in the Transportations sector only. The results of the other sectors are available upon request.

trades are conducted in discrete time (daily in our case). Although changing β changes the trading boundaries, as long as the changes are not dramatic, the prices at which the trades are executed will not significantly change. Hence the overall performance remains stable. This indicates that within a reasonable range of β, the performance of the co-integration model remains robust.

4.5. The co-integration tests

Now we turn to the fundamental question: are these samples' prices actually co-integrated in statistical sense? To answer this question, we conduct the augmented Dicky–Fuller (ADF) co-integration test on the entire paths of each sample pair.[17] In particular, for a pair consists of assets A_1 and A_2, we test the null hypothesis

H_0: $\log A_1 - \log A_2$ is not stationary.

against the alternative hypothesis

H_1: $\log A_1 - \log A_2$ is stationary.

at significance level of 0.05.[18]

The results are shown in Table 6. It turns out that most samples pass the co-integration test in their entire sampling period. The highest percentage of pass is achieved in the Utilities sector, in which 87.50% of the samples pass the test; while the lowest percentage of pass is achieved in the Transportations sector, in which 57.14% of the samples pass the test. In the other two sectors, Industrial and Financials, 80.77% and 85.71% of samples pass the test, respectively. These testing results confirm that these dual-listed shares indeed exhibit significant co-integration effect. Although these results cannot be used as ex ante evidence in the testing exercise, they help set the ground of our modeling strategy at the first place.

5. Extensions

We consider two extensions in this section. First, we introduce jumps to the dynamics of the pair's prices. Second, we extend the co-integration relation. The purpose of this section is to demonstrate various possibilities of extending our model, thus we only present the result in Economy 1 and the discussion are mostly brief.

5.1. Jump-diffusion process

In a jump-diffusion model, the asset prices P_{1t} and P_{2t} are assumed to follow[19]

$$\frac{dP_{1t}}{P_{1t}} = \mu_1 dt + \sigma_1 dW_{1t} - \lambda_1 z_t dt + d\sum_{i=1}^{N_t}(\xi_i - 1) \tag{13}$$

$$\frac{dP_{2t}}{P_{2t}} = \mu_2 dt + \sigma_2 dW_{2t} + \lambda_2 z_t \, dt \tag{14}$$

where ξ_is, i.i.d. copies of some positive random variable ξ, are the percentage jump size. We assume $Q = \log \xi$ has probability density function $\Phi(q)$, and N_t is a Poisson process with constant intensity $\eta > 0$. In addition, we assume mutual independence between ξ_i, N_t and $W_t = (W_{1t}, W_{2t})$.

With jumps in asset return, the value functions $V_j(P_1, P_2, t), j = -1, 0, 1$, satisfy the system of HJB equations (6) with the operator $\mathcal{L}$ replaced by $\mathcal{L}^1$:

$$\mathcal{L}^1 V_j = \mathcal{L}V_j + \eta \int_{-\infty}^{\infty} \left[V_j(P_1 e^q, P_2, t) - V_j(P_1, P_2, t)\right] \Phi(q)\, dq$$

Dimensional reduction (8) is still available, and U_j satisfies the system of reduced HJB equations (9) with the operator $\mathcal{M}$ replaced by $\mathcal{M}^1$:

$$\mathcal{M}^1 U_j = \mathcal{M}U_j + \eta \int_{-\infty}^{\infty} \left[U_j(xe^q, t) - U_j(x, t)\right] \Phi(q)\, dq$$

In an illustrative example, we assume that Q is double-exponentially distributed as in Kou (2002), with density function

$$\Phi(q) = p_1 \theta_1 e^{-\theta_1 q} I_{\{q \ge 0\}} + p_2 \theta_2 e^{\theta_2 q} I_{\{q < 0\}}$$

where $p_1 \ge 0$, $p_2 = 1 - p_1 \ge 0$, $\theta_1 > 1$, and $\theta_2 > 0$.[20]

[17] We do not conduct such tests in the empirical tests of trading strategies because, as pointed out in Liu and Timmermann (2013), the co-integration tests can only have low power in short samples. The moving window of one year sampling period may be insufficient to conduct meaningful co-integration tests.

[18] Note that although we use various relative weight δ to construct the portfolios in the empirical tests, the ADF test is invariant to the choice of δ.

[19] For simplicity, we only consider jumps in the return of asset 1. Including jumps in the returns of both assets is straightforward.

[20] The model can be further generalized to the mixed-exponential jump case as discussed in Cai and Kou (2011).

Table 5
Robustness check, sector: transportations.
This table provides the results on the robustness check of the co-integration strategy in the Transportations sector, against subjective discount rate β. We provide the trading results for $\beta = 0.001, 0.02$ and 0.1, respectively. The numbers outside parenthesis are the normalized P&L, and the numbers in the parenthesis are the maximal loss during the entire testing period.

Sample company/β	**Normalized cash profit (maximal loss)**					
	0.001		0.02		0.1	
AC	35.74	(−86.05)	35.74	(−86.05)	35.43	(−86.05)
CEA	−1.89	(−101.00)	−1.89	(−101.00)	−2.05	(−101.00)
CHS	3.46	(0.00)	3.46	(0.00)	3.46	(0.00)
COSCO	19.80	(0.00)	19.80	(0.00)	19.80	(0.00)
CSA	−0.45	(−212.83)	−0.45	(−212.83)	0.11	(−212.83)
DLP	5.86	(0.00)	5.86	(0.00)	5.86	(0.00)
GSR	11.96	(0.00)	11.96	(0.00)	11.91	(0.00)
Total	817.69	(−399.89)	817.69	(−399.89)	818.65	(−399.89)

Table 6
The co-integration tests.
This table provides the results of the ADF co-integration tests. We report the results on sectional basis. In each sector, we report the number of samples in that sector, the number of samples that passes the ADF co-integration test, and the percentage of samples that passes the ADF test. The ADF co-integration test is conducted on the entire historical path of each sample at the significance level of 0.05.

Sector	**Number of samples**	**Number of passes**	**Percentage of passes (%)**
Transportations	7	4	57.14
Industrial	26	21	80.77
Utilities	8	7	87.50
Financials	14	12	85.71

A numerical example of the optimal trading boundaries is presented in Fig. 8. We present the case with/without jumps together to make a comparison. Overall, the nature of trading boundaries is not changed by introducing jumps to the model. With the parameter values used, $E[\xi-1] = -0.0208$ and the jump in return is on average negative. As a result, the left panel shows that the arbitrageur is more likely to open short position, and is less likely to open long position, compared to the no jump case. Symmetrically, the right panel shows that the arbitrageur is more likely to close long position, and less likely to close short position.

5.2. General co-integration relation

In this section we generalize the pricing error $z_t = \ln P_{1t} - \ln P_{2t}$ in (1) and (2) to

$$z_t = \alpha + \gamma t + \ln P_{1t} + \theta \ln P_{2t}$$

With this extension, the value functions $V_j(P_1, P_2, t), j = -1, 0, 1$, are still governed by (6) and (7), except that the differential operator $\mathcal{L}$ is modified to

$$\mathcal{L}^2 \cdot = \frac{\partial \cdot}{\partial P_1}[\mu_1 - \lambda_1(\alpha+\gamma t + \ln P_1 + \theta \ln P_2)]P_1 + \frac{\partial \cdot}{\partial P_2}[\mu_2 + \lambda_2(\alpha + \gamma t + \ln P_1 + \theta \ln P_2)]P_2 - \beta \cdot \\ + \frac{1}{2}\sigma_1^2 P_1^2 \frac{\partial^2 \cdot}{\partial P_1^2} + \frac{1}{2}\sigma_2^2 P_2^2 \frac{\partial^2 \cdot}{\partial P_2^2} + \rho \sigma_1 \sigma_2 P_1 P_2 \frac{\partial^2 \cdot}{\partial P_1 \partial P_2}$$

However, the variable reduction (8) is no longer available, due to the nonlinear dynamics.

The trading boundaries at each time point are characterized by two variables $(x, P_2) = (P_1/P_2, P_2)$, not $x = P_1/P_2$ alone. Fig. 9 shows a snapshot of the optimal trading boundary at time $t = 0$. The left panel shows the optimal timing of entry. If the point (x, P_2) lies below/above the solid/dashed line, the arbitrageur should enter the asset market with a long/short position, respectively. The right panel shows the optimal timing of exit. If the point (x, P_2) lies below/above the solid/dashed line, the arbitrageur with short/long position should close her position and exit the asset market. Note that with general co-integration relation, the decision making will involve P_2 as well as P_1/P_2.

6. Concluding remarks

Our paper contributes to the literature with the study on the optimal trading strategy with co-integrated assets and transaction costs. We formulate the optimal stopping problem and obtain the optimal trading strategy. We show that the type of error correction (economy 1 or 2) has interesting effects on the optimal strategy. The optimal strategy could either

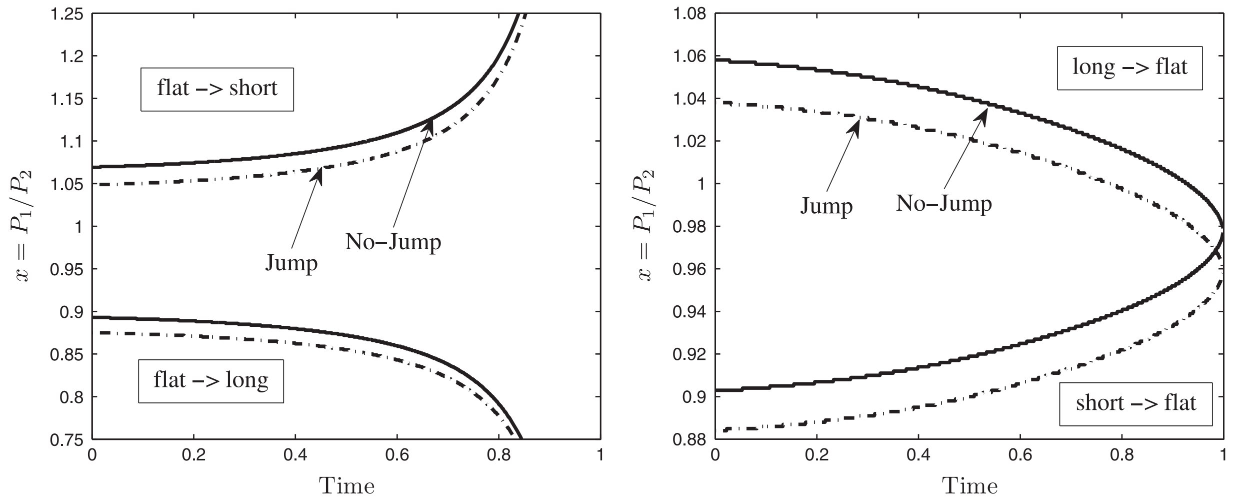

Fig. 8. Optimal trading boundaries, jump-diffusion process. This figure depicts the optimal trading boundaries against time t in a jump-diffusion model in Economy 1. The baseline parameter values we use are as follows: $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, K_b = K_s = 0.5\%, \lambda_1 = 0.35, \lambda_2 = 0.25, p_1 = 0.3, \theta_1 = 50, \theta_2 = 25, \eta = 0.5$.

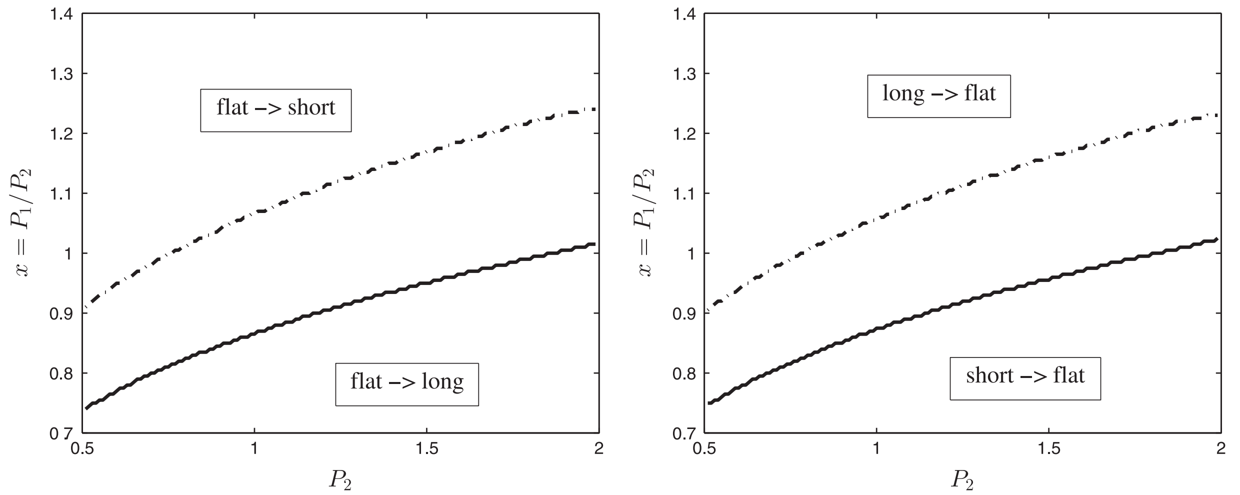

Fig. 9. Optimal trading boundaries at $t = 0$, General Co-integration Relation. This figure depicts the optimal trading boundaries at time $t = 0$. The baseline parameter values we use are as follows: $\beta = 0.04, \sigma_1 = 0.2, \sigma_2 = 0.24, \mu_1 = 0.10, \mu_2 = 0.112, \rho = 0.8, \alpha = 0.02, \gamma = 0, \theta = -1.2, K_b = K_s = 0.5\%, \lambda_1 = 0.35, \lambda_2 = 0.25$.

exhibit pure arbitrage incentives or mixed incentives. The mixed incentives consist of an endogenous stop loss component, which is novel and not revealed by alternative pairs trading models, to the authors' best knowledge.

We test the co-integration strategy with dual-listed Chinese shares. We find that the co-integration strategy generally outperforms the benchmark standard deviation strategy. The co-integration strategy yields much higher net profit, typically at the cost of moderate increase in the risk.

Finally, our study can be further extended in many other directions. Some possibilities are

- Including other market fractions, such as the margin requirement, the stock loan fee, and the portfolio constraint, which are incurred in real trades.
- In crisis time, the prices of a pair could exhibit abnormal divergence rather than convergence. Incorporating this abnormality is relevant to the risk management purpose.
- Including different risk appetite of the arbitrageur hence endogenously set control for the downside risk.

Appendix A

In this appendix, we present the penalty approximation of the system of HJB equations (9), proof of Proposition 2.1, the quasi-maximal likelihood estimate (QMLE) of the co-integration model (3) and (4), and the samples' information.

A.1. Penalty approximation of the system of HJB equations (9)

According to Dai et al. (2007), and Forsyth and Vetzal (2002), the penalty approximation of the system of HJB equations (9) is given by

$$\begin{cases} \dfrac{\partial U_0}{\partial t}+\mathcal{M}U_0+K\left(-U_0+U_1-x\alpha_b^1+\alpha_s^2\right)^+ +K\left(-U_0+U_{-1}-\alpha_b^2+x\alpha_s^1\right)^+ =0 \\ \dfrac{\partial U_{-1}}{\partial t}+\mathcal{M}U_{-1}+K\left(-U_{-1}+U_0-x\alpha_b^1+\alpha_s^2\right)^+ =0 \\ \dfrac{\partial U_1}{\partial t}+\mathcal{M}U_1+K\left(-U_1+U_0+x\alpha_s^1-\alpha_b^2\right)^+ =0 \end{cases} \tag{A.1}$$

where $a^+ = \max\{a,0\}$ and K is a large penalty parameter. As $K\to\infty$, the solution to (A.1) converges to the solution to (9). In the numerical method, the standard finite difference method is applied to (A.1).

A.2. Proof of Proposition 2.1

Proof. On one hand, given the terminal conditions (10), simple algebra yields that (11) and (12) hold when $t=T$. On the other hand, standard argument suggests that $U_j, j=-1,0,1$, are the continuous viscosity solution of system (9) and (10). The desired result follows by the continuity of the functions $U_j, j=-1,0,1$.□

A.3. QMLE of the co-integration model (3)–(4)

We consider the Euler approximation of (3) and (4)

$$y_1(t_k)-y_1(t_{k-1})=\left(\mu_1-\lambda_1\left(y_1(t_{k-1})-y_2(t_{k-1})\right)-\frac{1}{2}\sigma_1^2\right)\Delta t_k+\sigma_1\sqrt{\Delta t_k}Z_1$$

$$y_2(t_k)-y_2(t_{k-1})=\left(\mu_2+\lambda_2\left(y_1(t_{k-1})-y_2(t_{k-1})\right)-\frac{1}{2}\sigma_2^2\right)\Delta t_k+\sigma_2\sqrt{\Delta t_k}Z_2$$

where $\Delta t_k = t_k - t_{k-1}$, and Z_1 and Z_2 are two standard normal random variables with correlation ρ. Therefore, given $(y_1(t_{k-1}), y_2(t_{k-1})) = (\log P_1(t_{k-1}), \log P_2(t_{k-1}))$, the transition density to $(y_1(t_k), y_2(t_k)) = (\log P_1(t_k), \log P_2(t_k))$ is

$$f(y_1(t_k), y_2(t_k)|y_1(t_{k-1}), y_2(t_{k-1}); \Theta) = |\det(J_k)|^{-1}\phi(z_1, z_2; \Sigma)$$

where

$$z_1=\frac{y_1(t_k)-y_1(t_{k-1})-\left(\mu_1-\lambda_1\left(y_1(t_{k-1})-y_2(t_{k-1})\right)-\frac{1}{2}\sigma_1^2\right)\Delta t_k}{\sigma_1\sqrt{\Delta t_k}}$$

$$z_2=\frac{y_2(t_k)-y_2(t_{k-1})-\left(\mu_2+\lambda_2\left(y_1(t_{k-1})-y_2(t_{k-1})\right)-\frac{1}{2}\sigma_2^2\right)\Delta t_k}{\sigma_2\sqrt{\Delta t_k}}$$

$$\Sigma=\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

$\phi(z_1, z_2; \Sigma)$ is the probability density function of bivariate normal distribution with zero mean and covariance matrix Σ, and $\det(J) = \sigma_1\sigma_2\Delta t_k$ is the Jacobian determinant. The parameters Θ are then estimated by maximizing the likelihood of the entire path of returns

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}}\left\{\prod_{k=1}^{n} f(y_1(t_k), y_2(t_k)|y_1(t_{k-1}), y_2(t_{k-1}); \Theta)\right\}$$

subject to $\lambda_1+\lambda_2>0$.

A.4. The samples

We list the information on the samples we used in the empirical test here. The stock price data are obtained from Yahoo Finance and Netease Finance Channel. We present the samples' names, tickers (as A shares or H shares), and the sampling periods in Table A1. We assume the sampling period starts from Jan 2006 whenever applicable. In case that the samples were listed later than Jan 2006, we use the actual listing dates as the beginning of sampling period.

Table A1
The samples' information.
This table summarizes the information of the samples we use in the empirical test. A/H Ticker is the ticker symbol used in the Shanghai Security Exchange/the Hong Kong Stock Exchange. The column "Abbr." reports the abbreviations of each sample.

Company	Abbr.	A Ticker	H Ticker	Sampling period
A. Transportations				
Air China Ltd.	AC	601111.SS	00753.HK	08/18/2006–12/31/2013
China Eastern Airlines Co., Ltd.	CEA	600115.SS	00670.HK	01/04/2006–12/31/2013
China Shipping Ltd.	CHS	601866.SS	02866.HK	12/12/2007–12/31/2013
China COSCO Holdings Co., Ltd.	COSCO	601919.SS	01919.HK	06/26/2007–12/31/2013
China Southern Airlines Co., Ltd.	CSA	600029.SS	01055.HK	01/04/2006–12/31/2013
Port of Dalian Co., Ltd.	DLP	601880.SS	02880.HK	12/06/2010–12/31/2013
Guangshen Railway Co., Ltd.	GSR	601333.SS	00525.HK	12/22/2006–12/31/2013
B. Industrial				
Aluminum Corp. Of China Ltd.	ACCL	601600.SS	02600.HK	01/04/2006–12/31/2013
Angang Steel Co., Ltd.	AGS	000898.SZ	00347.HK	01/04/2006–12/31/2013
China Coal Energy Co., Ltd.	CCE	601898.SS	01898.HK	02/01/2008–12/31/2013
China Oilfield Services Ltd.	COFS	601808.SS	02883.HK	09/28/2007–12/31/2013
China Petroleum & Chemical Corp.	CPC	600028.SS	00386.HK	01/04/2006–12/31/2013
Chongqing Iron and Steel Co., Ltd.	CQIS	601005.SS	01053.HK	02/28/2007–12/31/2013
China South Locomotive & Rolling Stock Corp., Ltd.	CSR	601766.SS	01766.HK	08/21/2008–12/31/2013
Dongfang Electric Corp., Ltd.	DFE	600875.SS	01072.HK	01/04/2006–12/31/2013
Datang International Power Generation Co., Ltd.	DTIPG	601991.SS	00991.HK	12/20/2006–12/31/2013
Guangzhou Automobile Group Co., Ltd.	GZAG	601238.SS	02238.HK	03/29/2012–12/31/2013
Huadian Power International Corp., Ltd.	HDPI	600027.SS	01071.HK	01/04/2006–12/31/2013
Huaneng Power International, Inc.	HNPI	600011.SS	00902.HK	01/04/2006–12/31/2013
Jiangxi Copper Co., Ltd.	JXC	600362.SS	00358.HK	01/04/2006–12/31/2013
Maanshan Iron & Steel Co., Ltd.	MASIS	600808.SS	00323.HK	01/04/2006–12/31/2013
Northeast Electric Development Co., Ltd.	NEED	000585.SZ	00042.HK	01/04/2006–12/31/2013
Nanjing Panda Electronics Co., Ltd.	NJPE	600775.SS	00553.HK	01/04/2006–12/31/2013
PetroChina Co., Ltd.	PC	601857.SS	00857.HK	11/05/2007–12/31/2013
Shandong Xinhua Pharmaceutical Co., Ltd.	SDXHP	000756.SZ	00719.HK	01/04/2006–12/31/2013
Shanghai Electric Group Co., Ltd.	SHEG	601727.SS	02727.HK	01/04/2006–12/31/2013
Shanghai Pharmaceuticals Holding Co., Ltd.	SHPH	601607.SS	02607.HK	05/20/2011–12/31/2013
Sinopec Shanghai Petrochemical Co., Ltd.	SSHP	600688.SS	00338.HK	01/04/2006–12/31/2013
Tianjin Capital Environmental Protection Group Co., Ltd.	TJCEPG	600874.SS	01065.HK	01/04/2006–12/31/2013
Xinjiang GoldWind Science & Technology Co., Ltd.	XJGWST	002202.SZ	02208.HK	10/08/2010–12/31/2013
Yanzhou Coal Mining Co., Ltd.	YZCM	600188.SS	01171.HK	01/04/2006–12/31/2013
Zijin Mining Group Co., Ltd.	ZJMG	601899.SS	02899.HK	04/25/2008–12/31/2013
Zhongxing Telecommunication Equipment Corp.	ZTE	000063.SZ	00763.HK	01/04/2006–12/31/2013
C. Financials				
Agricultural Bank of China	ABC	601288.SS	01288.HK	07/16/2010–12/31/2013
Bank of China, Ltd.	BOC	601988.SS	03988.HK	07/17/2006–12/31/2013
Bank of Communications Co., Ltd.	BOCOM	601328.SS	03328.HK	05/15/2007–12/31/2013
China Construction Bank	CCB	601939.SS	00939.HK	09/25/2007–12/31/2013
China CITIC Bank Co., Ltd.	CITIC	601998.SS	00998.HK	04/27/2007–12/31/2013
CITIC Securities Co., Ltd.	CITICS	600030.SS	06030.HK	10/10/2011–12/31/2013
China Life Insurance Co., Ltd.	CLI	601628.SS	02628.HK	01/09/2009–12/31/2013
China Merchants Bank	CMB	601398.SS	01398.HK	09/22/2006–12/31/2013
China Minsheng Banking Co., Ltd.	CMSB	600016.SS	01988.HK	11/26/2009–12/31/2013
China Pacific Insurance (Group) Co., Ltd.	CPIC	601601.SS	02601.HK	12/23/2009–12/31/2013
Haitong Securities Co., Ltd.	HTS	600837.SS	06837.HK	04/27/2012–12/31/2013
Industrial & Commercial Bank of China	ICBC	601398.SS	01398.HK	10/27/2006–12/31/2013
New China Life Insurance Co., Ltd.	NCLI	601336.SS	01336.HK	01/30/2012–12/31/2013
Ping An Insurance (Group) Co. of China Ltd.	PAIC	601318.SS	02318.HK	03/01/2007–12/31/2013
D. Utilities				
Anhui Expway Co., Ltd.	AHE	600012.SS	00995.HK	01/04/2006–12/31/2013
China Communications Co., Ltd.	CCC	601800.SS	01800.HK	06/08/2012–12/31/2013
China Metal Lurgical Co., Ltd.	CML	601618.SS	01618.HK	09/24/2009–12/31/2013
China Railway Construction Co., Ltd.	CRC	601186.SS	01186.HK	03/13/2008–12/31/2013
China Railway Group Ltd.	CRG	601390.SS	00390.HK	12/07/2007–12/31/2013
Jiangsu Expway Co., Ltd.	JSE	600377.SS	00177.HK	01/04/2006–12/31/2013
Sichuan Expway Co., Ltd.	SCE	601107.SS	00107.HK	07/27/2009–12/31/2013
Shenzhen Expway Co., Ltd.	SZE	600548.SS	00548.HK	01/09/2006–12/31/2013

References

Andrade, S., Di Pietro, V., Seasholes, M., 2005. Understanding the profitability of pairs trading. Working Paper, UC Berkeley, Northwestern University.
Bolgün, K.E., Kurun, E., Güven, S., 2010. Dynamic pairs trading strategy for the companies listed in the Istanbul stock exchange. Int. Rev. Appl. Financ. Issues Econ. 2 (1), 37–57.

Brennan, M.J., Schwartz, E.S., 1990. Arbitrage in stock index futures. J. Bus. 63 (1), S7–S31.
Broussard, J.P., Vaihekoski, M., 2012. Profitability of pairs trading strategy in an illiquid market with multiple share classes. J. Int. Financ. Mark. Inst. Money 22 (5), 1188–1201.
Cai, N., Kou, S.G., 2011. Option pricing under a mixed-exponential jump diffusion model. Manag. Sci. 57 (11), 2067–2081.
Chiu, M.C., Wong, H.Y., 2011. Mean-variance portfolio selection of cointegrated assets. J. Econ. Dyn. Control 35 (8), 1369–1385.
Dai, M., Kwok, Y.K., You, H., 2007. Intensity-based framework and penalty formulation of optimal stopping problems. J. Econ. Dyn. Control 31 (12), 3860–3880.
Dai, M., Zhang, Q., Zhu, Q.J., 2010. Trend following trading under a regime switching model. SIAM J. Financ. Math. 1 (1), 780–810.
Dai, M., Zhong, Y., Kwok, Y.K., 2011. Optimal arbitrage strategies on stock index futures under position limits. J. Futur. Mark. 31 (4), 394–406.
Do, B., Faff, R., 2012. Are pairs trading profits robust to trading costs? J. Financ. Res. 35 (2), 261–287.
Do, B., Faff, R., Hamza, K., 2006. A new approach to modeling and estimation for pairs trading. In: Proceedings of 2006 Financial Management Association European Conference.
Duan, J.-C., Pliska, S.R., 2004. Option valuation with co-integrated asset prices. J. Econ. Dyn. Control 28 (4), 727–754.
Elliott, R.J., van der Hoek, J., Malcolm, W.P., 2005. Pairs trading. Quant. Finance 5 (3), 271–276.
Engle, R.F., Granger, C.W., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica 55 (2), 251–276.
Forsyth, P., Vetzal, K., 2002. Quadratic convergence for valuing American options using a penalty method. SIAM J. Sci. Comput. 23 (6), 2095–2122.
Gatev, E., Goetzmann, W.N., Rouwenhorst, K.G., 2006. Pairs trading: performance of a relative-value arbitrage rule. Rev. Financ. Stud. 19 (3), 797–827.
Kondor, P., 2009. Risk in dynamic arbitrage: the price effects of convergence trading. J. Finance 64 (2), 631–655.
Kou, S.G., 2002. A jump-diffusion model for option pricing. Manag. Sci. 48 (8), 1086–1101.
Liu, J., Timmermann, A., 2013. Optimal convergence trade strategies. Rev. Financ. Stud. 26 (4), 1048–1086.
Mori, M., Ziobrowski, A.J., 2011. Performance of pairs trading strategy in the US REIT market. Real Estate Econ. 39 (3), 409–428.
Perlin, M.S., 2009. Evaluation of pairs-trading strategy at the Brazilian financial market. J. Deriv. Hedge Funds 15 (2), 122–136.
Tourin, A., Yan, R., 2013. Dynamic pairs trading using the stochastic control approach. J. Econ. Dyn. Control 37 (10), 1972–1981.
Vidyamurthy, G., 2004. Pairs Trading: Quantitative Methods and Analysis, vol. 217. John Wiley & Sons, New York.