Optimal Convergence Trade Strategies*

Jun Liu† Allan Timmermann‡

UC San Diego and SAIF UC San Diego and CREATES

October 4, 2012

Abstract

Convergence trades exploit temporary mispricing by simultaneously buying relatively underpriced assets and selling short relatively overpriced assets. This paper studies optimal convergence trades under both recurring and non-recurring arbitrage opportunities represented by continuing and 'stopped' cointegrated price processes and considers both fixed and stochastic (Poisson) horizons. We demonstrate that conventional long-short delta neutral strategies are generally suboptimal and show that it can be optimal to simultaneously go long (or short) in two mispriced assets. We also find that the optimal portfolio holdings critically depend on whether the risky asset position is liquidated when prices converge. Our theoretical results are illustrated using parameters estimated on pairs of Chinese bank shares that are traded on both the Hong Kong and China stock exchanges. We find that the optimal convergence trade strategy can yield economically large gains compared to a delta neutral strategy.

Key words: convergence trades; risky arbitrage; delta neutrality; optimal portfolio choice

---

*The paper was previously circulated under the title Optimal Arbitrage Strategies. We thank the Editor, Pietro Veronesi, and an anonymous referee for detailed and constructive comments on the paper. We also thank Alberto Jurij Plazzi, Jun Pan, seminar participants at the CREATES workshop on Dynamic Asset Allocation, Blackrock, Shanghai Advanced Institute of Finance, and Remin University of China for helpful discussions and comments. Antonio Gargano provided excellent research assistance. Timmermann acknowledges support from CREATES, funded by the Danish National Research Foundation.

†UCSD and Shanghai Advanced Institute of Finance (SAIF). mail: Rady School of Management, Otterson Hall, 4S148, 9500 Gilman Dr, #0553, La Jolla, CA 92093-0553; ph: (858) 534-2022; email: junliu@ucsd.edu.

‡Rady School of Management, Otterson Hall, 4S146, 9500 Gilman Dr, #0553, La Jolla, CA 92093-0553; ph: (858) 534-0894; email: atimmerm@ucsd.edu.

1 Introduction

Convergence trades are arguably one of the most important strategies used to exploit mispricing in financial markets. In a classic convergence trade, two assets trade at different prices but have the same (or similar) payoff with certainty at a future date. Familiar examples of convergence trades include merger arbitrage (risk arbitrage), pairs trading (relative value trades), on-the-run/off-the-run bond trades, relative pricing of tranched structured securities, and arbitrage between the same stocks trading in different markets; see Bondarenko (2003), Hasbrouck (2003), and Hogan et al. (2004).

Industry practice as well as academic studies conventionally assume that convergence trades are based on delta neutral long-short positions, so that market exposure gets eliminated; see Shleifer and Vishny (1997), Mitchell and Pulvino (2001), Lehmann (2002), Liu and Longstaff (2004), Liu, Peleg and Subrahmanyam (2010), and Jurek and Yang (2007). However, our analysis demonstrates that the delta neutral arbitrage strategy is not the most efficient way to exploit temporary mispricing. Specifically, we show that the optimal convergence trading strategy that maximizes expected utility generally does not involve holding a delta neutral position. If the investor really prefers a market neutral portfolio, this can be better obtained by combining the optimal individual asset portfolio with the market index. Such a "market layover" strategy can potentially improve the performance of the combined portfolio.

The basic message of our analysis is that there is a trade-off between diversification and arbitrage. Delta neutral convergence trades are designed to explore long-term arbitrage opportunities but, in so doing, also create exposure to idiosyncratic risk which opens up diversification opportunities. By focusing on long-term arbitrage, delta neutral strategies do not take full advantage of the short-term risk-return trade-off and diversification benefits. By placing arbitrage opportunities in the context of a portfolio maximization problem, our optimal convergence strategy accounts for both arbitrage opportunities and diversification benefits.

We obtain several surprising new results. First, we show that it can be optimal to take the same 'side' of both risky assets (i.e., be long in both assets or short in both assets at the same time) even when prices eventually converge. This type of result, in which the sign of the optimal asset position can differ at short and long investment horizons, only occurs in a multiperiod model and will not happen in the static setting. A second surprising finding is how much of a difference it can make whether the arbitrage opportunities are recurring or non-recurring, particularly for small levels of mispricing. A third surprising finding is that the optimal convergence trade in some special cases involves holding only one asset and disregarding the second asset.

To model convergence trades, we follow earlier studies in the literature, e.g., Alexander (1999), Gatev, Goetzmann and Rouwenhorst (2006) and Jurek and Yang (2007), in assuming that individual asset prices contain a random walk component, but that pairs of asset prices can be cointegrated. This setup offers a tractable, yet flexible model that provides closed-form solutions in the case with recurring arbitrage opportunities. Cointegration between pairs of asset prices gives rise to a mean reverting error correction term which represents an expected excess return over and above the risk premium implied by the CAPM. This expected excess return is similar to a conventional 'alpha' component except that it is time-varying and has an expected value of zero in the long run. Such time variation in alpha reflects both absolute mispricing−abnormal expected returns over and above the CAPM benchmark values−and relative mispricing reflecting the relative prices of the two assets. At short horizons, time is too scarce for the arbitrage mechanism to be effective, and absolute mispricing dominates. At longer horizons, relative mispricing plays a key role as price differentials can be expected to revert to zero and the optimal portfolio is long in the (relatively) underpriced asset and short in the (relatively) overpriced asset.

Under recurring arbitrage opportunities, the optimal portfolio may switch from being long in one asset to shorting this asset if it changes from being (relatively) underpriced to being overpriced. Arguably, this misses the important point that investors close out their positions when prices converge and profit opportunities diminish. To deal with this issue, we modify the setup to allow for a 'stopped' cointegration process in which investors close out their position in the pair of risky assets when prices have converged and mispricing has disappeared. This case with non-recurring arbitrage opportunities gives rise to a set of very different boundary conditions when solving for the optimal portfolio weights.[1]

Comparing the cases with recurring and non-recurring arbitrage opportunities, we show that the optimal holdings in the risky assets can be very different, particularly when the price differential is small. Specifically, while risky stock holdings go to zero as the mispricing goes to zero under recurring arbitrage opportunities, under non-recurring arbitrage opportunities, risky asset positions are bounded away from zero when mispricing is non-zero and only get eliminated at zero. In practice, this can lead to quite different optimal portfolio holdings for the two cases.

We next compare the optimal unconstrained and delta neutral trading strategies. To illustrate the economic loss from adopting the delta neutral strategy, we consider a model whose parameters are calibrated to a new data set on Chinese banking shares. Stocks of some Chinese

[1] Our analysis does not account for funding risk. Even with a textbook arbitrage and a logarithmic utility maximizer, explicit modeling of funding risk is very involved, see, e.g., Liu and Longstaff (2004). Qualitatively, investors should hold smaller positions when funding risk is a concern while quantitatively it is difficult to combine both risky arbitrage and funding risk.

companies are traded simultaneously on the Hong Kong stock exchange as H shares and on the Chinese stock exchanges as A shares. A and H shares carry the same dividends and control rights, but can trade at very different prices. Due to trading restrictions on Chinese investors, H shares are more likely to be fairly priced while A shares are more likely to be mispriced. In this case the delta neutral long-short strategy is suboptimal and we find that the optimal convergence trade can generate economically significant gains over the arbitrage strategy for some of the banks.

In summary, the key contributions of our paper are as follows. First, we derive in closed form the optimal convergence trading strategy under the assumption that asset prices are cointegrated and arbitrage opportunities are recurring. We show that the delta neutral strategy is, in general, suboptimal and the optimal arbitrage strategy is determined by both relative mispricing (risky arbitrage) and absolute mispricing. Second, we extend the setup to allow for a stopped cointegrated price process in which the investor's position in pairs of risky assets is liquidated once prices converge. This can lead to optimal trading strategies that are quite different from those assuming recurring arbitrage opportunities. Third, we provide analytical solutions for optimal portfolio holdings when the holding period is stochastic as governed by a Poisson termination process. Fourth, we use a calibration exercise to pairs of Chinese banking shares to demonstrate that the loss incurred from following the delta neutral strategy can be economically significant.

The paper is organized as follows. Section 2 specifies our model for how asset prices evolve. Section 3 introduces the investor's portfolio choice problem. Sections 4 and 5 provide solutions for the optimal unconstrained and delta neutral strategies, respectively, separately considering the cases with recurring and non-recurring arbitrage opportunities. Section 6 analyses the case with a Poisson termination process. Section 7 conducts an empirical analysis of pairs of Chinese bank stocks simultaneously traded on the stock exchanges in China and Hong Kong. Section 8 concludes. Proofs are contained in the Appendix.

2 Convergence Trade and Cointegration

We assume that there is a riskless asset which pays a constant rate of return, r. A risky asset trading at the price P_{mt} represents the market index. This follows a geometric random walk process,

$$\frac{dP_{mt}}{P_{mt}} = (r + \mu_m)\, dt + \sigma_m dB_t, \tag{1}$$

where the market risk premium, μ_m, and market volatility, σ_m, are both constant and B_t is a standard Brownian motion. The market index is fairly priced. Papers such as Dumas, Kurshev and Uppal (2009) and Brennan and Wang (2006) assume that the market index is subject to pricing errors. We make no such assumptions here and instead concentrate on mispricing in (pairs of) individual asset prices.

In addition to the risk-free asset and the market index, we assume the presence of two risky assets whose prices P_{it}, $i = 1, 2$, evolve according to the equations

$$\frac{dP_{1t}}{P_{1t}} = (r + \beta\mu_m)\, dt + \beta\sigma_m dB_t + \sigma dZ_t + bdZ_{1t} - \lambda_1 x_t dt, \tag{2}$$

$$\frac{dP_{2t}}{P_{2t}} = (r + \beta\mu_m)\, dt + \beta\sigma_m dB_t + \sigma dZ_t + bdZ_{2t} + \lambda_2 x_t dt, \tag{3}$$

where λ_1, λ_2, β, b, and σ are constant parameters. Z_t and Z_{it} are standard Brownian motions, and B_t, Z_t, and Z_{it} are all mutually independent for $i = 1, 2$.[2] x_t represents pricing errors in our model and is the difference between the logarithms of the two asset prices, $p_{it} = \ln P_{it}$,

$$x_t = p_{1t} - p_{2t} = \ln\left(\frac{P_{1t}}{P_{2t}}\right). \tag{4}$$

We make the key assumption that $\lambda_1 + \lambda_2 > 0$. This implies that x_t is stationary and the logarithms of the prices are cointegrated with cointegrating vector $(1, -1)$. Following Engle and Granger (1987), we refer to x_t as the error-correction term. For simplicity, we use the CAPM as the benchmark, but our results will continue to hold for other asset pricing models.

Other statistical processes could be used to capture temporary deviations from equilibrium prices, including non-linear relations or fractional cointegration, to name a few. Our stylized model is meant to capture essential features of pricing errors while maintaining analytical tractability and allowing us to characterize the optimal trading strategy in closed form.

To make our analysis tractable we assume, unlike Xiong (2001), that the processes generating asset prices are exogenous with respect to the investor's decisions, and thus consider the optimal trades of a "small" investor with no market impact. Xiong considers investors with logarithmic utility and shows that while convergence traders normally reduce price volatility, they can actually amplify unfavorable shocks in situations where they are forced to liquidate their positions.[3]

Our setup captures the idea that two assets with identical payoffs can trade at different prices. Examples include pairs of stocks that have the same claim to dividends and identical voting rights but are traded in different markets and two stocks with the same payoffs such as the target and acquirer stocks in a merger. Specifically, the shares of Shell and Royal-Dutch traded at different prices despite being claims on the same underlying assets. If the time of convergence of the two prices was known with certainty, there would be a riskless arbitrage opportunity and

[2] The presence of a common nonstationary factor is consistent with the equilibrium asset pricing model analyzed by Bossaerts and Green (1989).

[3] See also Kondor (2009) for an approach that endogenizes the price process.

investors would have shorted the overpriced stock in the same amount as they would have been long in the underpriced stock. In reality, however, while the two stock prices can be expected to converge over time, the date where this would occur is not known ex ante, and so this is an example of risky arbitrage, i.e. a self-financing trading strategy with a strictly positive payoff today but a zero *expected* future cumulative payoff.

In equations (2)-(3), $\beta\sigma_m dB_t$ represents exposure to the market risk while $\sigma dZ_t + bdZ_{it}$ represents idiosyncratic risks. It is standard to assume that idiosyncratic risks are independent across different stocks with the market risk representing the only source of correlation among different assets. In our case, both assets are claims on similar fundamentals and so the presence of common idiosyncratic risk, dZ_t, is to be expected. The two asset prices are correlated both because of their exposure to the same market-wide risk factor (dB_t) and common idiosyncratic risk (dZ_t) but also due to the mean reverting error correction term (x_t) which will induce correlation between the two asset prices even in the absence of the two former components.

2.1 Alpha and Absolute Mispricing

The expected stock return in equations (2)-(3) is $(r+\beta\mu_m)dt-\lambda_1 x_t dt$ and $(r+\beta\mu_m)dt+\lambda_2 x_t dt$ respectively. If $\lambda_1 = \lambda_2 = 0$, expected returns satisfy the CAPM relation, $(r+\beta\mu_m)dt$; in this sense (i.e., that the CAPM correctly specifies the expected return), there is no mispricing in either asset and only the market index and the riskless asset will be held. Neither of the individual risky assets are held because of their additional idiosyncratic risk which goes without any associated extra expected return.

If either or both λ's are non-zero, expected returns have an extra term which represents deviations from the CAPM relation. When $-\lambda_1 x_t > 0$, asset one has a higher expected return than is justified by its risk, i.e., a positive alpha, and so asset one is underpriced. Conversely, when $-\lambda_1 x_t < 0$, asset one has a lower expected return than justified by its risk, i.e., a negative alpha, and is overpriced. Similarly, when $\lambda_2 x_t > 0$, asset two has a positive alpha and thus is underpriced; when $\lambda_2 x_t < 0$, asset two has a negative alpha and is overpriced. Therefore, $-\lambda_1 x_t$ and $\lambda_2 x_t$ represent mispricing "alphas" and they capture each asset's absolute mispricing which we know must exist; after all, convergence trades involve two assets with the same payoff trading at different prices.

Different combinations of the λ_1 and λ_2 price adjustment parameters are likely to reflect the (relative) liquidity of the two assets. For example, if two high-volume stocks are both traded in liquid markets, it is likely that their prices adjust equally rapidly and so $\lambda_1 = \lambda_2$ holds as a good approximation. A good point in case is the Royal Dutch and Shell shares traded on the Amsterdam and London stock exchanges, respectively, as considered by Jurek and Yang

(2007). Conversely, in the case of the Chinese stocks traded as both H-shares in Hong Kong and as A-shares in China, we might expect that stock prices adjust more rapidly in the Hong Kong market where there are fewer market frictions than in China, and so the λ-value in Hong Kong is expected to be greater than the $\lambda-$value in China, a conjecture that we corroborate empirically in Section 7. Idiosyncratic liquidity shocks to pairs of shares that lead one price to be relatively high while the other becomes relatively low can then be represented by λ_1 and λ_2 values that have the same sign but are of different magnitudes. As a third case, suppose two Chinese banking shares are hit by a common (industry-wide) liquidity shock that leads both assets to become underpriced, but that one asset is more underpriced than the other. This case can be captured by letting λ_1 and λ_2 be of different signs. For example, if $x_t > 0$ and $\lambda_1 < 0$, $\lambda_2 > 0$, then both assets are underpriced relative to the market. From the stability condition $\lambda_1 + \lambda_2 > 0$, the second asset must be more underpriced than the first asset and so is expected to revert back to its equilibrium price more rapidly.

2.2 Cointegration and Relative Mispricing

The variable $x_t = \ln(P_{1t}/P_{2t})$ is the difference in the logarithm of the prices of two assets that should be identical and so represents relative mispricing. If $\lambda_1 + \lambda_2 > 0$, equations (2)-(4) constitute a continuous-time cointegrated system with $-\lambda_1 x_t$ and $\lambda_2 x_t$ as the error correction terms.

Even though both asset prices are almost geometric Brownian motions, the difference between the two is stationary because of the error correction term which captures relative mispricing between the two assets. The dynamics of this term satisfies

$$dx_t = -\lambda_x x_t dt + b_x dZ_{xt}, \tag{5}$$

where

$$\begin{aligned} \lambda_x &= \lambda_1 + \lambda_2, \\ b_x dZ_{xt} &= b dZ_{1t} - b dZ_{2t}, \end{aligned} \tag{6}$$

and $b_x = \sqrt{2}b$. The assumption that the mean reversion coefficient, λ_x, is positive ensures that x_t is stationary. Mean reversion in x_t captures the temporary nature of any mispricing.

The error correction term produces mean reversion that keeps mispricing stationary and pricing errors "small" compared to either of the individual integrated price processes p_{1t}, p_{2t}. This ensures that, in the words of Chen and Knez (1995), "closely integrated markets should assign to similar payoffs prices that are close".

We consider two cases. In the first case ("recurring arbitrage opportunities"), the price differential, x_t, only spends an infinitesimally small time at zero, is characterized at all times by the dynamics in Eq. (5) and so follows a stationary process. In this case, stock prices are always described by Eqs. (2)-(3). In the second case ("non-recurring arbitrage opportunities"), any price difference is temporary and gets permanently eliminated the first time the two prices converge and $x_t = 0$. In this case, the price dynamics is subject to the additional restriction that $x_{\tau+\Delta} = 0$ for all $\Delta \geq 0$, where $\tau = \min(t : x_t = 0)$ is a stopping time. In this case, prices remain identical after they converge. As we show later, optimal portfolio weights are different for these cases.[4]

To summarize, absolute mispricing is determined in the short run by the conditional alphas while relative mispricing is determined in the long run by cointegration between asset prices.

3 Portfolio Choice Problem

Denote the investor's allocation to the market portfolio by ϕ_{mt} while the weights on the individual risky assets are given by ϕ_{it}, $i = 1, 2$. In the absence of intermediate consumption, the investor's wealth, W_t, evolves according to the process

$$\begin{aligned} dW_t &= W_t \left(rdt + \phi_{mt} \left(\frac{dP_{mt}}{P_{mt}} - rdt \right) + \phi_{1t} \left(\frac{dP_{1t}}{P_{1t}} - rdt \right) + \phi_{2t} \left(\frac{dP_{2t}}{P_{2t}} - rdt \right) \right) \\ &= W_t \Big(rdt + (\phi_{mt} + \beta(\phi_{1t} + \phi_{2t}))(\mu_m dt + \sigma_m dB_t) + \phi_{1t}(\sigma dZ_t + bdZ_{1t} - \lambda_1 x_t dt) \\ &\quad + \phi_{2t}(\sigma dZ_t + bdZ_{2t} + \lambda_2 x_t dt) \Big). \end{aligned}$$

We assume that the investor maximizes the expected value of a power utility function defined over terminal wealth, W_T

$$\max_{\{\phi_{mt}\}_{t=0}^T, \{\phi_{1t}\}_{t=0}^T, \{\phi_{2t}\}_{t=0}^T} \frac{1}{1-\gamma} E_0 \left[W_T^{1-\gamma} \right]. \tag{7}$$

If the investor is a hedge fund, T can be viewed as the fund's lifetime.

The investor's value function is given by

$$J(t, x, W) = \frac{1}{1-\gamma} \mathrm{E}_t [W_T^{*(1-\gamma)}], \tag{8}$$

[4] Under the first scenario, any mispricing is stationary over time. Conversely, mispricing in Liu and Longstaff (2004) and Liu, Peleg, and Subrahmanyam (2010) is expressed in terms of a Brownian bridge and a generalized Brownian bridge. These specifications are not stationary and are useful to describe cases where mispricing will be zero for sure at some future date. For example, on the settlement date of a futures contract, the difference between the spot and futures price has to be zero even though the individual spot and futures prices follow non-stationary processes (Brenner and Kroner, 1995).

where W_T^* is the wealth at time T obtained by the optimal trading strategy with $W_t = W$ and $x_t = x$ at time t.

Using standard results, it follows that when prices are described by diffusion processes such as those in equations (1)-(3), J satisfies the HJB equation

$$\max_{\tilde{\phi}_m, \phi_1, \phi_2} J_t + (-\lambda_x x) J_x + \frac{1}{2} b_x^2 J_{xx} + (r + \tilde{\phi}_m \mu_m + (-\lambda_1 \phi_1 + \lambda_2 \phi_2) x) W J_W$$

$$+ b^2 (\phi_1 - \phi_2) W J_{xW} + \frac{1}{2} (\tilde{\phi}_m^2 \sigma_m^2 + (\phi_1 + \phi_2)^2 \sigma^2 + (\phi_1^2 + \phi_2^2) b^2) W^2 J_{WW} = 0. \quad (9)$$

Here $\tilde{\phi}_m = \phi_m + \beta(\phi_1 + \phi_2)$, noting that maximizing over (ϕ_m, ϕ_1, ϕ_2) is equivalent to maximizing over $(\tilde{\phi}_m, \phi_1, \phi_2)$. Exploiting homogeneity, the value function should take the following form

$$J(t, x, W) = \frac{W^{1-\gamma}}{1-\gamma} e^{u(t,x)}. \quad (10)$$

Expressed in terms of $u(t, x)$, the first-order conditions for $\tilde{\phi}_m$, ϕ_1 and ϕ_2 are

$$\begin{aligned} \tilde{\phi}_m \sigma_m^2 (-\gamma) + \mu_m &= 0, \\ -\lambda_1 x + b^2 u_x + ((\phi_1 + \phi_2)\sigma^2 + \phi_1 b^2)(-\gamma) &= 0, \\ \lambda_2 x - b^2 u_x + ((\phi_1 + \phi_2)\sigma^2 + \phi_2 b^2)(-\gamma) &= 0. \end{aligned}$$

Solving these equations, the optimal portfolio weights take the form

$$\begin{aligned} \phi_{mt}^* &= \frac{\mu_m}{\gamma \sigma_m^2} - \beta(\phi_{1t}^* + \phi_{2t}^*), \\ \begin{pmatrix} \phi_{1t}^* \\ \phi_{2t}^* \end{pmatrix} &= \frac{1}{\gamma(2\sigma^2 + b^2) b^2} \begin{pmatrix} \sigma^2 + b^2 & -\sigma^2 \\ -\sigma^2 & \sigma^2 + b^2 \end{pmatrix} \begin{pmatrix} -\lambda_1 x + b^2 u_x \\ \lambda_2 x - b^2 u_x \end{pmatrix}. \end{aligned} \quad (11)$$

The first term in the expression for the market portfolio weight, ϕ_{mt}^*, is the standard mean-variance portfolio weight and thus depends on the market's Sharpe ratio divided by the investor's coefficient of risk aversion and market volatility. The second term offsets the market exposure of the individual assets which is linear in the portfolio weights, ϕ_{1t}^* and ϕ_{2t}^*, and proportional to their beta.

Turning to the expression for ϕ_{1t}^* and ϕ_{2t}^*, the first term in the bracket in Eq. (11), which depends explicitly on λ_1 and λ_2, is the mean-variance term; the second term, which is proportional to u_x, is the intertemporal hedging term. Note that parameters associated with the market index such as β, μ_m, and σ_m, do not affect ϕ_{1t}^*, ϕ_{2t}^*. This is because the individual assets' market exposure is hedged using the market index. In contrast, asset-specific parameters such as the volatility of the common and independent idiosyncratic risk components (σ, b), their sensitivity to the mispricing component (λ_1 and λ_2), the size of the mispricing ($x = \ln(P_1/P_2)$) in addition to the investor's attitude to risk (γ) and investment horizon (through u_x), determine optimal asset holdings.

Substituting the optimal portfolio weights back into the HJB equation, the following PDE is obtained for $u(t, x)$

$$\begin{aligned} 0 &= u_t - \frac{\lambda_x}{\gamma} x u_x + b^2 u_{xx} + \frac{1}{\gamma} b^2 u_x^2 \\ &+ \left(r + \frac{1}{2\gamma} \left(\mu_m^2/\sigma_m^2 + \frac{\lambda_1^2(\sigma^2 + b^2) + 2\lambda_1\lambda_2\sigma^2 + \lambda_2^2(b^2 + \sigma^2)}{b^2(b^2 + 2\sigma^2)} x^2 \right) \right) (1 - \gamma). \end{aligned} \tag{12}$$

The terminal condition is

$$u(T, x) = 1. \tag{13}$$

Note that the PDE for u in (12) and the boundary condition in (13) are quite general in the sense that they hold under both recurring and non-recurring arbitrage opportunities. In contrast, the boundary condition for $u(t, x)$ depends on what happens when x reaches zero and the two prices converge.

4 Optimal Investment Strategies

This section separately considers cases with a continuing cointegrated price process (recurring arbitrage opportunities), which gives closed-form solutions for the optimal portfolio weights, versus a stopped cointegrated price process (non-recurring arbitrage opportunities) for which the optimal portfolio weights have to be solved numerically.

4.1 Continuing Cointegrated Price Process

We first consider the case where stock prices continue to be described by the cointegrated process even after the price difference reaches zero. In this case, with probability 1, the prices will diverge again. This case is relevant for Royal Dutch and Shell's stock prices as well as for A- and H-share prices of Chinese stocks. In this case, the PDE for $u(t, x)$ is satisfied for all x and can be solved in closed form.

Lemma 1 *Suppose asset prices evolve according to equations (1)-(3) and the investor has constant relative risk aversion preferences. Then the value function in equation (10) is characterized by*

$$u(t, x) = A(t) + \frac{1}{2} C(t) x^2,$$

where the $A(t)$ and $C(t)$ functions only depend on time, t, and are given in the Appendix.

Substituting the specific form of $u(.)$ in Lemma 1 into equation (11), the optimal portfolio weights can be obtained.

Proposition 1 *Under the assumptions of Lemma 1, the optimal weights on the market portfolio, ϕ^*_{mt}, and the individual assets, $(\phi^*_{1t}, \phi^*_{2t})$, are given by*

$$\phi^*_{mt} = \frac{\mu_m}{\gamma\sigma_m^2} - (\phi^*_{1t} + \phi^*_{2t})\beta,$$

$$\begin{pmatrix} \phi^*_{1t} \\ \phi^*_{2t} \end{pmatrix} = \frac{1}{\gamma(2\sigma^2 + b^2)b^2} \begin{pmatrix} \sigma^2 + b^2 & -\sigma^2 \\ -\sigma^2 & \sigma^2 + b^2 \end{pmatrix} \begin{pmatrix} -\lambda_1 + b^2 C(t) \\ \lambda_2 - b^2 C(t) \end{pmatrix} \ln\left(\frac{P_{1t}}{P_{2t}}\right).$$

This result follows from equation (11) and Lemma 1 which implies that $u_x = C(t)x$. We shall use Proposition 1 to compute optimal portfolio weights given a set of estimates for the parameters $\{\lambda_1, \lambda_2, b^2, \sigma^2\}$.

4.2 Stopped Cointegrated Process

Next consider the case where the price difference, x_t, stays at zero when it reaches zero so that it follows a "stopped" cointegrated process. This case is relevant for "one-shot" arbitrages such as risk arbitrage in mergers and acquisitions. Alternatively, if the investor decides to close out the position once prices converge, we can also view the price process as a stopped cointegrated process.

Formally, let τ be defined by

$$\tau = \min\{t : x_t = 0\}.$$

The log-price differential now follows a stopped AR(1) process $X_\tau = x_{t\wedge\tau}$.

When $t < \tau$, the value function $u(t, x)$ satisfies the same partial differential equation in (12) as the continuing cointegrated process. When $t \geq \tau$, prices have converged, the investment opportunity in the individual stocks is gone, and the investor will hold only the market index. This is a standard Merton problem and so produces utility at τ of

$$\frac{W_\tau^{1-\gamma}}{1-\gamma} e^{(1-\gamma)(r + \frac{\mu_m^2}{2\gamma\sigma_m^2})(T-\tau)}. \tag{14}$$

Equation (14) implies that

$$u(\tau, 0) = (1-\gamma)(r + \frac{\mu_m^2}{2\gamma\sigma_m^2})(T-\tau). \tag{15}$$

When $x \to \infty$, x is unlikely to reach zero before time T, so the boundary condition becomes

$$u(t, x) \to A(t) + \frac{1}{2}C(t)x^2. \tag{16}$$

where $A(t) + \frac{1}{2}C(t)x^2$ is the $u(t,x)$ function for the continuing cointegrated process introduced in Lemma 1. In this case, there is no closed-form solution for the value function $u(t,x)$, but we can solve for $u(t,x)$ numerically and obtain the optimal portfolio weights using equation (11).

The most surprising feature of the portfolio weights under a stopped cointegrated price process is that they approach a non-zero limit as $x \to 0$. Under recurring arbitrage opportunities, the individual portfolio weights are proportional to the log-price difference, x, and so approach zero as $x \to 0$. The position in these two assets unwinds gradually as $x \to 0$. In contrast, if x is stopped at $x = 0$, the individual portfolio weights approach a non-zero limit when $x \to 0$. In this case, the investor has finite positions in both assets as $x \downarrow 0$; to unwind the position as required for this case, a large portfolio adjustment has to be made at $x = 0$.

To establish intuition for this finding, note that in the case of recurring arbitrage opportunities, the risk premium, which is proportional to x, can become negative. In fact, when $x = 0$, the probability that x becomes negative in the next instant equals the probability that x becomes positive. As a consequence, when x equals zero, it is optimal to hold zero in the two risky stocks: holding individual stocks would only add idiosyncratic risk without any additional increase in expected returns since there is no mispricing.

Conversely, in the case with non-recurring arbitrage opportunities, the investment opportunity disappears the first time x equals zero. Because the x process cannot fall below zero, there is not the same downside risk as in the case with recurring arbitrage opportunities. Suppose now that x is very small but just a little above zero. Then the distribution of future returns is similarly truncated since the worst that can happen to the risk premium is that it becomes zero (when $x = 0$). Since the future risk premium is finite but positive and there is limited downside risk, the agent chooses optimally to hold a finitely positive amount in the risky assets.

To illustrate our results, we use parameter values obtained from an empirical analysis of pairs of Chinese banking stocks traded simultaneously as A-shares on the Chinese stock exchange and as H-shares on the Hong Kong stock exchange. Details of the analysis are provided in Section 7. We focus on two pairs of error-correction parameter values, namely $(\lambda_1, \lambda_2) = (0.29, 0.31)$ and $(\lambda_1, \lambda_2) = (0.52, -0.35)$, corresponding to the estimated parameters for Agricultural Bank of China and China Citic Bank, respectively. We let x vary from zero to 0.2, corresponding to asset one being relatively overpriced.[5] For the first set of parameters, $(\lambda_1, \lambda_2) = (0.29, 0.31)$, the positive error correction term leads to a decrease in the price of asset one and an increase in the price of asset two. For the second set of parameters, $(\lambda_1, \lambda_2) = (0.52, -0.35)$, the price of asset two tends to decrease when this asset is undervalued, but the price of asset one decreases

[5] This and the subsequent figures use the following (annualized) parameter values calibrated to the Chinese bank share data: $\sigma = 0.15$, $b = 0.30$, $\mu_m = 0.05$, $\sigma_m = 0.35$, $r = 0.02$.

by even more, thus ensuring convergence.

The left window of Figure 1 plots the optimal weights under both recurring and non-recurring arbitrage opportunities when the degree of mispricing is varied from zero to 20%, while the investment horizon is kept fixed at $T = 1$ year and $\gamma = 4$. Under these parameter values, and assuming recurring arbitrage opportunities, the optimal weights are of opposite signs and almost identical in magnitude. Moreover, as x gets large, the weights under recurring and non-recurring arbitrage opportunities converge as we would expect, since it becomes unlikely that x will cross zero prior to time T. When x is small, however, the two sets of weights are very different. Whereas the weights under recurring arbitrage opportunities converge to zero, the weights under non-recurring arbitrage opportunities remain bounded away from zero even for small values of x.

The right window of Figure 1 shows that the two sets of weights can differ by even more when λ_1 and λ_2 are of opposite signs. For this case the optimal holdings under recurring arbitrage opportunities are short in both stocks, although the holding in the second stock is quite close to zero. In contrast, the holdings under non-recurring arbitrage opportunities start with a short position in the first stock and a long position in the second stock, although the latter declines towards zero as x gets larger. Once again, as the magnitude of the mispricing grows, the two pairs of weights converge.

4.3 Short-Term Risk-Return Trade-off and Long-Term Arbitrage

Our cointegrated price processes allow for mispricing in the short term but impose that prices revert to their equilibrium (no arbitrage) relation in the long term. These properties are reflected in the portfolio weights. At short horizons, the portfolio weights are dominated by the mean-variance component, which is determined by the instantaneous risk-return trade-off. At long horizons, the portfolio weights reflect equilibrium forces. Conventional long-short arbitrage strategies impose that the two stock portfolio weights should have opposite signs. This is not true for the optimal portfolio strategy. We show below that at short horizons, it is possible that both stock portfolio weights can have the same sign, while at long horizons they can have opposite signs. To illustrate these points, we take advantage of the closed-form solution for the case with continuing cointegrated processes, but the intuition and conclusion applies to the stopped cointegrated processes as well.

From Proposition 1, the optimal stock portfolio weights for the continuing cointegrated process can be written as

$$\begin{pmatrix} \phi_{1t}^* \\ \phi_{2t}^* \end{pmatrix} = \left(\begin{pmatrix} -(\lambda_1+\lambda_2)\frac{\sigma^2}{b^2} - \lambda_1 \\ (\lambda_1+\lambda_2)\frac{\sigma^2}{b^2} + \lambda_2 \end{pmatrix} \frac{1}{2\sigma^2+b^2} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} C(t) \right) \frac{1}{\gamma} \ln\left(\frac{P_{1t}}{P_{2t}}\right). \quad (17)$$

At short horizons the term proportional to $C(t)$ is small. Without loss of generality, we can assume that $\lambda_1 > 0$. If $P_{1t} > P_{2t}$, then (17) shows that it is optimal to short the first stock, i.e., $\phi_{1t}^* < 0$, which is unsurprising since this stock is (relatively) overvalued and has a negative alpha. More surprisingly, however, it is possible that it is optimal to simultaneously short the second stock, i.e., $\phi_{2t}^* < 0$. This follows when $(\lambda_1+\lambda_2)\frac{\sigma^2}{b^2} + \lambda_2 < 0$ and suggests the following corollary:

Corollary 1 *Suppose that* $(\lambda_1+\lambda_2)\frac{\sigma^2}{b^2} + \lambda_2 < 0$. *Then at short horizons the optimal portfolio takes a short position in both stocks.*

The intuition for this result is as follows. At short horizons, intertemporal hedging ceases to be important and so both assets are shorted if they have negative alphas and are overpriced. In this situation, investors can optimally exploit the absolute mispricing by shorting the two assets and are willing to be exposed to idiosyncratic risk. At short horizons, the investor acts myopically and portfolio holdings are dictated by the conventional mean-variance trade-off.

This result is in stark contrast with the delta neutral strategy which consists of symmetric long-short positions in the two stocks.

Turning to the opposite case with a long horizon, using results in the appendix we have,

$$C(t) \to \frac{\lambda_1+\lambda_2 - \sqrt{(\lambda_1+\lambda_2)^2 - 2\frac{(\lambda_1^2+\lambda_2^2)(\sigma^2+b^2)+2\lambda_1\lambda_2\sigma^2}{(b^2+2\sigma^2)}(1-\gamma)}}{2b^2} = \frac{\lambda_1+\lambda_2-\xi}{2b^2}, \quad (18)$$

where

$$\xi = \sqrt{(\lambda_1+\lambda_2)^2 - 2\frac{(\lambda_1^2+\lambda_2^2)(\sigma^2+b^2)+2\lambda_1\lambda_2\sigma^2}{(b^2+2\sigma^2)}(1-\gamma)}.$$

Hence, at long horizons, the optimal stock holdings are given by

$$\begin{aligned} \begin{pmatrix} \phi_{1t}^* \\ \phi_{2t}^* \end{pmatrix} &= \begin{pmatrix} -(\lambda_1+\lambda_2)\sigma^2 - \lambda_1 b^2 + (\lambda_1+\lambda_2-\xi)(2\sigma^2+b^2) \\ (\lambda_1+\lambda_2)\sigma^2 + \lambda_2 b^2 - (\lambda_1+\lambda_2-\xi)(2\sigma^2+b^2) \end{pmatrix} \frac{\ln\left(\frac{P_{1t}}{P_{2t}}\right)}{\gamma b^2(2\sigma^2+b^2)} \\ &= \begin{pmatrix} (\lambda_1+\lambda_2)(\sigma^2+b^2) - \xi(2\sigma^2+b^2) - \lambda_1 b^2 \\ -(\lambda_1+\lambda_2)(\sigma^2+b^2) + \xi(2\sigma^2+b^2) + \lambda_2 b^2 \end{pmatrix} \frac{\ln\left(\frac{P_{1t}}{P_{2t}}\right)}{\gamma b^2(2\sigma^2+b^2)}. \end{aligned} \quad (19)$$

Using this result, we get the following corollary:

Corollary 2 *Suppose that* $-(\lambda_1+\lambda_2)(\sigma^2+b^2)+\xi(2\sigma^2+b^2)+\lambda_2 b^2 > 0$. *Then at long horizons the optimal portfolio takes a short position in one stock and a long position in the other stock.*

At long horizons, the intertemporal hedging component dominates. In our analysis the intertemporal hedging component reflects how prices converge to the equilibrium implied by no arbitrage and so the optimal portfolio holdings of an investor with a long horizon will reflect the no-arbitrage conditions. The condition in Corollary 2 requires that the arbitrage effect dominates the mean-variance risk-return trade-off. When γ is close to 1, ξ is close to $\lambda_1 + \lambda_2$, and the long-run inequality in Corollary 2 reduces to the short-run inequality in Corollary 1, so the two stock positions will be of opposite signs at long horizons if and only if this holds at the short horizon. However, when $\gamma >> 1$, ξ can be arbitrarily large and the positions in the two stocks will always have opposite signs at long horizons.[6]

Figure 2 shows that these results are not only of theoretical interest and apply to the empirical analysis of some of the Chinese banks traded in China and Hong Kong. The left window shows the case where $\lambda_1 \approx \lambda_2 > 0$. For this case the optimal stock weights under recurring and non-recurring arbitrage opportunities are of opposite signs and of nearly identical magnitude at both long and short horizons. In contrast, in the right window we show a case where λ_1 and λ_2 are of opposite signs. For this case it is optimal to hold both shares short, i.e., $\phi_{1t}^* < 0, \phi_{2t}^* < 0$, at short horizons. As the horizon grows, the sign of the holdings in asset two changes from negative to positive, so at long horizons we have $\phi_{1t}^* < 0, \phi_{2t}^* > 0$.

Jurek and Yang (2007) derive the optimal investment strategy under power utility, assuming recurring arbitrage opportunities and $\lambda_1 = \lambda_2$. Like us, they find that the intertemporal hedging demand component can be an important part of the investor's overall position.

4.4 Optimal Convergence Trades with a Single Stock

Another surprising result is that it may be optimal to hold only one stock in a convergence trade. It is often automatically assumed that the optimal convergence trade strategy will hold the two assets simultaneously, i.e., buy the underpriced asset and short the overpriced asset. However, as we shall see, this need not hold, at least not if asset prices are modeled by cointegrated price processes. In fact, it is possible that the optimal position holds only one of the assets.

Specifically, suppose that $\gamma = 1$ so that $C(t) = 0$ and there is no dynamic hedging. Furthermore, assume that one of the assets is mispriced ($\lambda_1 > 0$) while the other is correctly priced

[6] Alternatively, the optimal stock portfolio weights can be written as

$$\begin{pmatrix} \phi_{1t}^* \\ \phi_{2t}^* \end{pmatrix} = \left(\begin{pmatrix} -1 \\ 1 \end{pmatrix} \frac{(\lambda_1 + \lambda_2)\sigma^2}{b^2(2\sigma^2 + b^2)} + \begin{pmatrix} -\lambda_1 \\ \lambda_2 \end{pmatrix} \frac{1}{2\sigma^2 + b^2} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} C \right) \frac{1}{\gamma} \ln \left(\frac{P_{1t}}{P_{2t}} \right).$$

While the first and third terms in the large bracket have opposite signs and identical magnitude, this need not hold for the second term as long as $\lambda_1 \neq \lambda_2$. Thus the two optimal stock positions need not have equal size with opposite signs.

($\lambda_2 = 0$) and unaffected by the error correction term. Finally, assume that $\sigma = 0$ so there is no correlation between idiosyncratic shocks:

$$\begin{aligned} \frac{dP_{1t}}{P_{1t}} &= (r + \beta\mu_m)\,dt + \beta\sigma_m dB_t + bdZ_{1t} - \lambda_1 x_t dt, \\ \frac{dP_{2t}}{P_{2t}} &= (r + \beta\mu_m)\,dt + \beta\sigma_m dB_t + bdZ_{2t}. \end{aligned} \tag{20}$$

The optimal portfolio puts a zero weight on asset two in this case. Thus, the optimal convergence trade strategy holds only one stock, and not the pair. This is surprising, since convergence trades seem to imply that both assets are held.

This result can be understood as follows. Although the two asset prices remain cointegrated, all price adjustment occurs through the first asset. When only asset one is mispriced, asset two will be held to reduce the variance of the optimal strategy due to their common idiosyncratic risks. When $\sigma = 0$, the idiosyncratic risk of asset two is independent of the idiosyncratic risk of asset one, so asset two cannot be used to reduce the variance of asset one and therefore will not be held.

There is a catch to this result, however. Even though the instantaneous correlation between idiosyncratic risks of asset one and asset two is zero, there is a long-term correlation due to their cointegration. In fact, dZ_{2t} is one of the shocks to dx_t, so there is an intertemporal hedging benefit from holding asset two. Therefore, for investors with $\gamma \neq 1$, even though the myopic component of the optimal portfolio weight on asset two is still zero, the intertemporal hedging component will not be zero. The ratio $|\phi^*_{2t}|/|\phi^*_{1t}|$ will then increase from zero as the horizon expands.

5 Optimal Delta Neutral Strategy

Many popular investment strategies assume that the portfolio is delta neutral. For example, Liu and Longstaff (2004) and Liu, Peleg, and Subrahmanyam (2010) directly specify the dynamics of the difference in asset prices and so one can view the strategies studied in these papers as assuming that $\phi_{1t} = -\phi_{2t}$. As pointed out by Gatev et al. (2006), and confirmed empirically by these authors, pairs of stocks are often selected to be market neutral. In our model where the two stocks are assumed to have identical market betas, delta neutrality directly translates into the constraint $\phi_{1t} = -\phi_{2t}$.

This section shows that although the market-neutral strategy is very popular, it can clearly be suboptimal. If the investor really prefers a market-neutral portfolio, this can be better obtained by combining the optimal individual asset holdings with the market index. A utility maximizer could rationally accept some exposure to the market portfolio to earn the market risk premium.

When there is mispricing, the optimal portfolio should have some exposure to the market (and thus not be delta neutral) and some position in the mispriced assets. Our results suggest that the best way to achieve a delta neutral position is to use the market index to hedge away the market exposure in the mispriced assets. Using mispriced assets alone to achieve delta neutrality will necessarily under-exploit opportunities offered by mispricing in the individual stocks.

With the constraint, $\phi_{1t} = -\phi_{2t}$, we have $\tilde{\phi}_{mt} = \phi_{mt} + \beta(\phi_{1t} + \phi_{2t}) = \phi_{mt}$. The HJB equation is then given by

$$\max_{\tilde{\phi}_m,\phi_1} J_t + (-\lambda_x x)J_x + \frac{1}{2}b_x^2 J_{xx} + (r + \tilde{\phi}_m\mu_m - \phi_1(\lambda_1 + \lambda_2)x)WJ_W$$
$$+b^2 2\phi_1 W J_{xW} + \frac{1}{2}(\phi_m^2\sigma_m^2 + 2\phi_1^2 b^2)W^2 J_{WW} = 0. \tag{21}$$

Assume again that the investor has power utility so the value function takes the form

$$J = \frac{W^{1-\gamma}}{1-\gamma}e^{v(t,x)}, \tag{22}$$

where $v(t,x)$ is the delta neutral counterpart to $u(t,x)$ in (10). The HJB equation reduces to

$$\max_{\tilde{\phi}_m,\phi_1} v_t - \lambda_x x v_x + \frac{1}{2}b_x^2(v_x^2 + v_{xx}) + (r + \phi_m\mu_m - \phi_1(\lambda_1 + \lambda_2)x)(1-\gamma)$$
$$+2b^2\phi_1(1-\gamma)v_x + \frac{1}{2}(\phi_m^2\sigma_m^2 + 2\phi_1^2 b^2)(1-\gamma)(-\gamma) = 0.$$

The first-order condition for ϕ_m is

$$\mu_m + \sigma_m^2\phi_m(-\gamma) = 0,$$

which leads to

$$\phi_m^* = \frac{1}{\gamma}\frac{\mu_m}{\sigma_m^2}. \tag{23}$$

Similarly, the first-order condition for ϕ_1 is

$$-(\lambda_1 + \lambda_2)x + 2b^2 v_x + 2b^2\phi_1^* = 0,$$

and so

$$\phi_{1t}^* = \frac{-(\lambda_1 + \lambda_2)x + 2b^2 v_x}{2\gamma b^2}. \tag{24}$$

Substituting the optimal weights back into the HJB equation, we get the following PDE

$$v_t - \lambda_x x v_x + \frac{1}{2}b_x^2(v_{xx} + v_x^2) + (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma)$$
$$+\frac{1}{2\gamma}\frac{(-(\lambda_1 + \lambda_2)x + 2b^2 v_x)^2}{2b^2}(1-\gamma) = 0. \tag{25}$$

The terminal condition is

$$v(T,x) = 0.$$

Once again we separately characterize the optimal portfolio holdings for the cases with recurring and non-recurring arbitrage opportunities.

5.1 Continuing Cointegrated Process

In this case, the PDE specified by equation (25) has a closed-form solution.

Lemma 2 *Suppose asset prices evolve according to equations (1)-(3) and the investor has constant relative risk aversion preferences. Then the function $v(t,x)$ is characterized by*

$$v(t,x) = B(t) + \frac{1}{2}D(t)x^2,$$

where the $B(t)$ and $D(t)$ functions only depend on time, t, and are given in the Appendix.

From equations (23) and (24) and Lemma 2, the optimal portfolio weights under the constraint that $\phi_{1t} = -\phi_{2t}$ can be characterized in closed form as follows:

Proposition 2 *The optimal portfolio weights under the delta neutrality constraint $\phi_{1t} + \phi_{2t} = 0$ are given by*

$$\begin{aligned}
\check{\phi}^*_{mt} &= \frac{\mu_m}{\gamma\sigma_m^2}, \\
\check{\phi}^*_{1t} &= \frac{-(\lambda_1+\lambda_2)\ln\left(\frac{P_{1t}}{P_{2t}}\right) + 2b^2 D(t)\ln\left(\frac{P_{1t}}{P_{2t}}\right)}{2\gamma b^2}.
\end{aligned}$$

To illustrate this result, Figure 3 compares the optimal unconstrained and delta neutral stock weights using the parameter estimates from our empirical analysis. For the first set of parameter values where $\lambda_1 \approx \lambda_2$, the delta neutral and unconstrained optimal weights are essentially identical as we would expect from propositions 1 and 2 since in this case $\phi^*_{1t} \approx \phi^*_{2t}$. Conversely, the delta neutral weights are very different from the unconstrained optimal weights in the right window where λ_1 and λ_2 are of opposite signs. For this case, the delta neutral strategy takes a short position in stock one, which is (relatively) overvalued, and a long position in stock two, which is undervalued. In contrast, the optimal unconstrained strategy takes a large short position in the first stock and a small short position in the second stock.

5.2 Stopped Cointegrated Process

Using similar arguments as for the unconstrained case, the boundary conditions for $v(t,x)$ are

$$v(t,0) = (1-\gamma)(r + \frac{\mu_m^2}{2\gamma\sigma_m^2})(T-\tau). \tag{26}$$

When $x \to \infty$, we have

$$v(t,x) \to B(t) + \frac{1}{2}D(t)x^2. \tag{27}$$

Here, $B(t) + \frac{1}{2}D(t)x^2$ is the $v(t, x)$ function for the case with recurring arbitrage opportunities given in Lemma 2. Again, we can solve the PDE numerically for the case with non-recurring arbitrage opportunities.

Figure 4 illustrates the delta neutral positions for the case with non-recurring arbitrage opportunities. Once again the delta neutral position is virtually identical to the unconstrained position for the first set of parameter estimates for which $\lambda_1 \approx \lambda_2$ (left window). Large differences occur, however, when λ_1 and λ_2 are dissimilar, as they are in the right window. For example, the unconstrained optimal holding of the second stock decreases as a function of x, while conversely the delta neutral holding in the second stock increases as x grows so as to balance out the increasingly short position in the first stock. As a result, the magnitude of the delta neutral and the unconstrained optimal positions can be very different across a wide spectrum of $x-$values.

To better understand the differences between the unconstrained optimal weights and the delta neutral weights, consider again the case with mispricing only in asset one ($\lambda_1 = 1, \lambda_2 = 0$). For this case we expect the magnitude of the myopic demand for asset one in the unconstrained optimal portfolio to exceed that in the delta neutral portfolio. To see why, notice that shocks to asset one have two components: one that is perfectly correlated with shocks to asset two (Z_t) and one that is independent of shocks to this asset (Z_{1t}). The unconstrained allocation ensures that the perfectly correlated shock is completely hedged by taking an appropriate position in asset two. Hence the unconstrained optimal holding is determined by the risk premium and the variance of the independent shock.

The delta neutral portfolio constrained to have suboptimal relative weights (1,-1), earns the same risk premium as asset one because the risk premium of asset two is zero ($\lambda_2 = 0$). Furthermore, the size of the independent shock to asset two is the same as that for asset one, but is not completely hedged. Thus the suboptimal portfolio earns the same risk premium but at a higher risk and so the investor will hold less of asset one under the constrained strategy.

By the same token, because the unconstrained portfolio has the same risk premium as the delta neutral portfolio but also has lower risk, the investor would take a larger position in the unconstrained portfolio for the intertemporal hedging demand.

5.3 Wealth Gains from the Optimal Strategy

Differences between the optimal and delta neutral trading strategies, while interesting in their own right, are not of economic significance unless we can demonstrate that they sometimes lead to sizeable economic losses for sensible choices of parameter values. This subsection therefore explores cases where the expected wealth gain from imposing delta neutrality will be minimal as well as cases where the opposite holds. First, we provide a sufficient condition for the optimal unconstrained strategy to be delta neutral:

Proposition 3 *The optimal strategy is delta neutral if* $\lambda_1 = \lambda_2$.

This follows from equation (11) since $\phi_{1t}^* = -\phi_{2t}^*$ and the optimal strategy is delta neutral when $\lambda_1 = \lambda_2$. When this condition holds, clearly there will be no loss from applying a delta neutral strategy.

We next address more broadly the size of the economic loss associated with adopting the conventional delta neutral strategy. We base our comparison on a simple result that allows us to compute the wealth gain of the optimal investment strategy relative to the suboptimal delta neutral strategy.

The following proposition allows us to compare the wealth under the two investment strategies for the scenario with recurring arbitrage opportunities:

Proposition 4 *The wealth gain of the optimal strategy relative to the delta neutral strategy assuming a mispricing of* x *is*

$$R = e^{\frac{1}{1-\gamma}(u(t,x)-v(t,x))},$$

where $u(t,x)$ *and* $v(t,x)$ *are defined in equations (10) and (22), respectively.*

To see this, note that, given a wealth level W, we need $W \times R$ under the delta neutral strategy to achieve the same level of utility as under the optimal strategy, where

$$\frac{(W \times R)^{1-\gamma}}{1-\gamma} e^{v(t,x)} = \frac{W^{1-\gamma}}{1-\gamma} e^{u(t,x)},$$

from which the result follows. A similar result allows us to compute wealth gains under the stopped cointegrated price process.

Using Proposition 4, it is easy to evaluate the investor's wealth gain. Figure 5 shows the expected wealth gain from adopting the unconstrained optimal trading strategy versus the delta neutral strategy as a function of the initial price difference, x. The graph considers the second set of error correction parameters, $(\lambda_1 = 0.52, \lambda_2 = -0.35)$ since expected gains are close to zero for the first set of parameters because the two sets of weights are nearly identical; see Proposition 3. Assuming recurring arbitrage opportunities, the expected gain rises from close to 3% of initial wealth to 4.5% of initial wealth as x goes from zero to 20%. In the case with non-recurring arbitrage opportunities, the expected gain starts at zero, but increases to a level close to 3.5% when $x = 0.20$.

6 Stochastic Investment Horizon with Poisson Termination Process

So far, we have assumed either that the investor's horizon is deterministic (that is, the horizon is a constant T) or stochastic in a way that is related to the x-process crossing some boundary. Another possibility is that the fund is forced to liquidate its position at a random time for reasons extraneous to the risky arbitrage such as withdrawal of funds or liquidity shocks. For example, Krishnamurthy (2010) argues that small shocks to credit or liquidity conditions can have large balance sheet and/or information amplifiers on financial intermediaries, and these could ultimately force a fund to wind up its positions if it can no longer obtain funding.

To capture this case, suppose that the investor's horizon is given by an exogenous Poisson arrival time, τ. Thus the investor's objective is

$$\mathcal{E}_0\left[\frac{1}{1-\gamma}W_{\bar{\tau}}^{1-\gamma}\right], \tag{28}$$

where τ is a Poisson arrival time, and $\bar{\tau}=\min(\tau,T)$, and T is constant. For example, T could be the lifetime of a limited partnership hedge fund. The expectation operator $\mathcal{E}_0$ denotes the expectation taken with respect to the Poisson process τ, in addition to Brownian motions.

The problem with Poisson exit time when stock prices follow a geometric Brownian motion is studied for the case without transaction costs by Merton (1970) and with transaction costs by Liu and Loewenstein (2002). In our paper, asset prices are not geometric Brownian motions. We obtain the analytical solution for the portfolio weights by combining the dynamic programming approach and the martingale approach.

Following Merton (1970), the objective can be written as

$$\mathrm{E}_0\left[\int_0^T \rho e^{-\rho t}\frac{1}{1-\gamma}W_t^{1-\gamma}dt+\frac{e^{-\rho T}}{1-\gamma}W_T^{1-\gamma}\right], \tag{29}$$

where ρ is the intensity of the Poisson process τ and E_0 denotes the expectation assuming a deterministic horizon, and thus applies to the Brownian motions. This leads to a standard dynamic programming problem.

To obtain an analytical expression for the value function we next consider the martingale approach. Note that the (unique) pricing kernel can be expressed as

$$\pi_t=e^{-rt}e^{-\frac{1}{2}\frac{\mu_m^2}{\sigma_m^2}-\frac{\mu_m}{\sigma_m}dB_{mt}}e^{-\frac{1}{2}(\eta_{1t}^2+\eta_{2t}^2)-(\eta_{1t}dZ_{1t}+\eta_{2t}dZ_{2t})}, \tag{30}$$

where η_t is the market price of risk which is given by

$$\eta_t=\begin{pmatrix}\eta_{1t}\\ \eta_{2t}\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{2\sigma^2+b^2}+b}{2} & \frac{\sqrt{2\sigma^2+b^2}-b}{2}\\ \frac{\sqrt{2\sigma^2+b^2}-b}{2} & \frac{\sqrt{2\sigma^2+b^2}+b}{2}\end{pmatrix}^{-1}\begin{pmatrix}-\lambda_1 x_t\\ \lambda_2 x_t\end{pmatrix}.$$

To solve the investor's problem in this case, note that there are infinitely many constraints, each indexed by t, for all $t > 0$:

$$\mathrm{E}_0[\pi_t W_t] = w_0,$$

where π_t is the pricing kernel. We show in the Appendix that the value function under a stochastic horizon specified by a Poisson arrival process is the weighted average of the value function under deterministic horizons with weights given by the Poisson distribution. Moreover, in this case the optimal portfolio weights are given in the following proposition:

Proposition 5 *Suppose the assumptions of Lemma 1 hold and that the investor's horizon is governed by a Poisson termination process with intensity parameter ρ. Then the optimal weights on the individual assets, (ϕ_{1t}^*, ϕ_{2t}^*), are given by*

$$\phi_1^* = \frac{-\lambda_1 b^2 - (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x + \frac{2\int_0^T C(t) e^{-\rho t} e^{A(t)+C(t)x^2} dt + 2C(t) e^{-\rho(T-t)} e^{A(t)+C(t)x^2}}{\gamma \left(\int_0^T e^{-\rho t} e^{A(t)+C(t)x^2} dt + e^{-\rho(T-t)} e^{A(t)+C(t)x^2} \right)} x,$$

and

$$\phi_2^* = \frac{\lambda_2 b^2 + (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x - \frac{2\int_0^T C(t) e^{-\rho t} e^{A(t)+C(t)x^2} dt + 2C(t) e^{-\rho(T-t)} e^{A(t)+C(t)x^2}}{\gamma \left(\int_0^T e^{-\rho t} e^{A(t)+C(t)x^2} dt + e^{-\rho(T-t)} e^{A(t)+C(t)x^2} \right)} x.$$

where $A(t)$ and $C(t)$ are the same as in Lemma 1.

Figure 6 compares the deterministic horizon (of length T) versus stochastic (Poisson) horizon solutions under recurring arbitrage opportunities, assuming a Poisson intensity parameter of $\rho = 2$ and setting $T = 1$. The solution under the stochastic horizon is quite close to that under the deterministic horizon. The possibility of an early termination of the arbitrage opportunities leads the investor to reduce the positions in the risky assets. This is understandable since the value function under the stochastic horizon is a weighted average of the value function under deterministic horizons with horizon lengths between zero and T. This dampens the sensitivity of the optimal weights in the Poisson case with respect to movements in x, because the sensitivity to x increases with the horizon in the deterministic case.

7 Empirical Example: Chinese Bank Shares

To illustrate the empirical relevance of our theoretical results, we next provide an analysis of pairs of Chinese bank stocks. Some Chinese bank shares are traded simultaneously as A shares on the Shanghai stock exchange and as H Shares in Hong Kong. They represent claims on the same assets and so should not, once converted into the same currency, be priced differently.

Hence this case matches our theoretical setup and we use this as an empirical example to calibrate the parameters of our cointegrated model.

Specifically, we consider seven pairs of Chinese A and Hong Kong H bank shares, namely Agricultural Bank of China (sample period: 7/16/2010 - 2/15/2012) China Merchants Bank (9/22/2006 - 2/15/2012); Bank of China (7/15/2006 - 2/15/2012); China Citic Bank (4/27/2007 - 2/15/2012); China Minsheng Banking (11/26/2009 - 2/15/2012); China Con. Bank (9/25/2007 - 2/15/2012); and Bank of Commerce (5/15/2007 - 2/15/2012). The shortest sample spans 394 days, while the longest sample spans 1,411 days. Chinese A shares are quoted in yuan, while H shares are quoted in Hong Kong Dollars, so we convert both series into US dollar terms to make them comparable.

Figure 7 shows that log-prices of the pairs of A- and H-shares tend to move broadly in synchrony. The price differentials plotted in Figure 8 show that differences at times can be quite substantial, although the figure also suggests that price differences tend to decrease when they get unusually large, consistent with mean reversion towards zero, although the speed of mean reversion can be quite slow. Bearing in mind that cointegration tests can have low power in relatively short samples such as ours, pair-wise tests of cointegration, reported in the first column in Table 1, reject the null of no cointegration for four of the seven series.[7]

To help calibrate the parameters of our model, we estimate pairwise error correction models for the seven sets of A-shares and H-shares. Table 1 reports the estimated values of λ_1 and λ_2 along with their t-statistics. The first set of estimates only include $\log(P_{1t})$ and $\log(P_{2t})$ in the error correction model and so may be subject to omitted variable bias since the effect of the market price is left out.[8] To deal with this issue, we also show results for a two-stage procedure that first orthogonalizes the bank share prices with respect to the market index and then estimates an error correction model for the resulting log-prices.[9] In all but one case, the sum of the estimated values of λ_1 and λ_2 is positive. This is consistent with our modeling assumption of mean reversion in price differentials. The lack of statistical significance of some of the λ_1 and λ_2 estimates can again be attributed to the relatively short data samples and the fairly slow speed of mean reversion.

[7] The test results are based on an ADF unit root test for the log-price difference, $\log(P_{1t}) - \log(P_{2t})$. This has slightly better power than the conventional cointegration test since it does not require estimating the cointegration parameter which is instead assumed to be unity.

[8] The findings are robust to the number of lags included in the analysis and also hold when a Bayesian vector error correction model is used.

[9] Although the two pairs of λ estimates differ, their correlation, at 0.73 (λ_1) and 0.67 (λ_2), is quite high.

7.1 Trading Results

To further illustrate the difference between the delta neutral versus the unconstrained strategies, we undertake a simple trading experiment.[10] For each pair of banking shares we compute the optimal weights using the end-of-day value of $x_t = \log(P_{1t}) - \log(P_{2t})$, the estimates of (λ_1, λ_2) from Table 1 and moment-based estimates of (b^2, σ^2, β). We assume $\mu_m = 5\%, \sigma_m = 35\%, r = 2\%$. We set the terminal date, T, to February 15, 2010 corresponding to the end of our data and focus on the case with recurring arbitrage opportunities which yields more observations than the case with non-recurring arbitrage opportunities and so offers the more informative comparisons for this particular application.[11] Rebalancing is assumed to take place daily and we set $\gamma = 4$. While daily rebalancing does not match the assumption of continuous time price dynamics, it is likely to provide a reasonable approximation, see Bertsimas, Kogan, and Lo (2000).

Table 2 reports the results. In all cases the mean portfolio return associated with the unconstrained strategy is at least as large as that associated with the delta neutral strategy and in some cases it is substantially larger, i.e., 16.7% versus 6.9% per annum for China Merchants Bank and 16.1% versus 8.3% for China Minsheng Banking. Since the volatility of returns on the unconstrained strategy is also higher, we consider the Sharpe ratio and cumulated wealth, starting from \$100. In four of seven cases this is highest for the unconstrained strategy. Similarly, the cumulated wealth of the optimal strategy, reported in the last panel of Table 2, exceeds that of the delta neutral strategy for all but China Con. Bank.

These results ignore transaction costs which are unfortunately unavailable and difficult to assess. Interestingly, however, for four of the seven pairs of banking shares, the turnover of the unconstrained and delta neutral strategies, measured through $(|\Delta\phi^*_{1t}| + |\Delta\phi^*_{2t}|)/2$, are very close, i.e., within 10% of each other, and so transaction costs would not appear to explain the differences in performance of the two strategies, at least for a number of the pairings.

8 Conclusion

Convergence trades form an important part of many relative value investment strategies. It is generally assumed that a delta neutral long-short position should be taken in pairs of under- and over-valued assets so as to ensure market neutrality. This paper argues that such a strategy does not optimally take advantage of the associated risk-return trade-off. Instead, we derive optimal portfolio strategies when pairs of risky asset prices are cointegrated so that their conditional

[10] We thank the editor, Pietro Veronesi, for suggesting to use a trading strategy to illustrate our theoretical results.

[11] For the majority of banking shares, x crosses zero after relatively few observations.

excess return can be characterized through a mean reverting error correction process. When arbitrage opportunities are recurring, the optimal portfolio holdings can be characterized in closed form. We also consider the interesting case where the investor's position is closed out as soon as prices converge. This second case can give rise to very different solutions for the optimal portfolio holdings and expected utility. We compare our optimal solutions to those achieved under conventional trading strategies restricted to be delta neutral, i.e., insensitive to market conditions, and show that considerable gains in expected utility can be achieved by deviating from conventional convergence trades.

Our analysis considers the actions of an unconstrained fund. In reality funds' trades are constrained in important ways, reflecting limits on borrowing, regulatory constraints and other market imperfections. Perhaps the single most important constraint arises from funding risk which arises when a trade has to be closed down early due to lack of funding. Moreover, such funding risk is likely to be greatest in bad states of the world and so could well be correlated with the arbitrage opportunities analyzed here. We view this as a topic of great interest for future research.

Appendix

This appendix derives the results needed for the optimal portfolio weights presented in the paper.

Proof of Lemma 1

Without the boundary condition at $x = 0$, we conjecture

$$u(t,x) = A(t) + \frac{1}{2}C(t)x^2.$$

Substituting this conjecture into equation (12) yields an equation that is affine in x. Setting the terms in the equation that are independent of x and the coefficient of x to zero, we have the following ordinary differential equations (ODE)

$$\begin{aligned} 0 &= A_t + \frac{1}{2}b_x^2 C + (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma), \\ 0 &= C_t - \frac{2\lambda_x}{\gamma}C + \frac{2}{\gamma}b^2C^2 + \frac{\lambda_1^2(\sigma^2+b^2) + 2\lambda_1\lambda_2\sigma^2 + \lambda_2^2(b^2+\sigma^2)}{\gamma b^2(b^2+2\sigma^2)}(1-\gamma). \end{aligned}$$

We first solve for C. The ODE for C can be written as

$$0 = C_t + \frac{2}{\gamma}b^2(C - C_+)(C - C_-),$$

where $C_{\pm}$ are the roots of the following quadratic equation in C

$$0 = C^2 - \frac{\lambda_x}{b^2}C + \frac{(\lambda_1^2 + \lambda_2^2)(\sigma^2 + b^2) + 2\lambda_1\lambda_2\sigma^2}{2b^4(b^2 + 2\sigma^2)}(1 - \gamma).$$

These roots are given by

$$C_{\pm} = \frac{\lambda_x \pm \sqrt{(\lambda_x)^2 - 2\frac{(\lambda_1^2+\lambda_2^2)(\sigma^2+b^2)+2\lambda_1\lambda_2\sigma^2}{(b^2+2\sigma^2)}(1-\gamma)}}{2b^2}.$$

The ODE for C can be solved in the following steps. First,

$$C_t = -\frac{2b^2}{\gamma}(C - C_+)(C - C_-),$$

which can be written as

$$\frac{dC}{(C - C_+)(C - C_-)} = -\frac{2b^2}{\gamma}dt,$$

or, equivalently,

$$(\frac{1}{C - C+} - \frac{1}{C - C_-})dC = -\frac{2b^2}{\gamma}(C_+ - C_-)dt.$$

Integrating on both sides, we have

$$d\ln\left|\frac{C - C_+}{C - C_-}\right| = -\frac{2b^2}{\gamma}(C_+ - C_-)dt.$$

Using the terminal condition $C(T) = 0$, we get

$$\ln\left|\frac{C - C_+}{C - C_-}\right| - \ln\left|\frac{C_+}{C_-}\right| = \frac{2b^2}{\gamma}(C_+ - C_-)(T - t),$$

which is the same as

$$\left|\frac{C/C_+ - 1}{C/C_- - 1}\right| = e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)}.$$

- When $\gamma > 1$, $C_+ > 0$ and $C_- < 0$, and $0 > C > C_-$.
- When $\gamma < 1$, if $C_{\pm}$ are real, then $C \geq 0$, and $C(t)$ decreases from $C < C_-$ at $t = 0$ to 0 at $t = T$. Thus, $C < C_- < C_+$.

In both cases,

$$\left|\frac{1 - C/C_+}{1 - C/C_-}\right| = \frac{1 - C/C_+}{1 - C/C_-}.$$

Thus, we get

$$\frac{1 - C/C_+}{1 - C/C_-} = e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)},$$

which can be solved as

$$C(t) = -\frac{e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)} - 1}{\frac{1}{C_+} - \frac{e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)}}{C_-}} = C_-\frac{e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)} - 1}{e^{\frac{2b^2}{\gamma}(C_+-C_-)(T-t)} - \frac{C_-}{C_+}}.$$

Turning to the solution for $A(t)$, integrating over the ODE for $A(t)$, we get

$$0 = A(t) - \frac{1}{2}b_x^2 \int_t^T C(s)ds + (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma)(t-T),$$

where we have used the terminal condition $A(T) = 0$. Note that

$$\begin{aligned}
& - \int_t^T C(s)ds = C_- \left(t - T - (\frac{C_-}{C_+} - 1)\frac{1}{\frac{C_-}{C_+}\frac{2b^2}{\gamma}(C_+ - C_-)} \ln \left| \frac{1 - \frac{C_-}{C_+}e^{-\frac{2b^2}{\gamma}(C_+ - C_-)(T-t)}}{1 - \frac{C_-}{C_+}} \right| \right) \\
= \quad & C_- \left(t - T + \frac{\gamma}{2C_- b^2} \ln \left(\frac{1 - \frac{C_-}{C_+}e^{-\frac{2b^2}{\gamma}(C_+ - C_-)(T-t)}}{1 - \frac{C_-}{C_+}} \right) \right).
\end{aligned}$$

Hence, we get

$$A(t) = \Big((r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma) + b^2 C_-\Big)(T-t) - \frac{\gamma}{2}\ln \left(\frac{1 - \frac{C_-}{C_+}e^{-\frac{2b^2}{\gamma}(C_+ - C_-)(T-t)}}{1 - \frac{C_-}{C_+}} \right).$$

Note that

$$\begin{aligned}
& -b^2 C_-(T-t) + \frac{\gamma}{2}\ln \left(\frac{C_+ - C_- e^{-\frac{2b^2}{\gamma}(C_+ - C_-)(T-t)}}{C_+ - C_-} \right) \\
= \quad & -\frac{C_+ + C_-}{2}b^2(T-t) \\
& + \frac{\gamma}{2}\ln \left(\frac{C_+ e^{\frac{b^2}{\gamma}(C_+ - C_-)(T-t)} - C_- e^{-\frac{b^2}{\gamma}(C_+ - C_-)(T-t)}}{C_+ - C_-} \right) \\
= \quad & -\frac{C_+ + C_-}{2}b^2(T-t) \\
& + \frac{\gamma}{2}\ln \left(\frac{\frac{C_+ + C_- + (C_+ - C_-)}{2}e^{\frac{b^2}{\gamma}(C_+ - C_-)(T-t)} - \frac{C_+ + C_- - (C_+ - C_-)}{2}e^{-\frac{b^2}{\gamma}(C_+ - C_-)(T-t)}}{C_+ - C_-} \right) \\
= \quad & -\frac{C_+ + C_-}{2}b^2(T-t) + \frac{\gamma}{2} \times \\
& \ln \left(\frac{C_+ + C_-}{2}\frac{e^{\frac{b^2}{\gamma}(C_+ - C_-)(T-t)} - e^{-\frac{b^2}{\gamma}(C_+ - C_-)(T-t)}}{C_+ - C_-} + \frac{1}{2}(e^{\frac{b^2}{\gamma}(C_+ - C_-)(T-t)} + e^{\frac{-b^2}{\gamma}(C_+ - C_-)(T-t)}) \right) \\
= \quad & -\frac{\lambda_1 + \lambda_2}{2}(T-t) \\
& + \frac{\gamma}{2}\ln \left(\frac{\lambda_1 + \lambda_2}{2}\frac{e^{\frac{\xi}{\gamma}(T-t)} - e^{-\frac{\xi}{\gamma}(T-t)}}{\xi} + \frac{1}{2}(e^{\frac{\xi}{\gamma}(T-t)} + e^{\frac{-\xi}{\gamma}(T-t)}) \right),
\end{aligned}$$

with

$$\xi = \sqrt{(\lambda_1 + \lambda_2)^2 - 2\frac{(\lambda_1^2 + \lambda_2^2)(\sigma^2 + b^2) + 2\lambda_1\lambda_2\sigma^2}{(b^2 + 2\sigma^2)}(1-\gamma)}.$$

Hence, we can write $A(t)$ as

$$\begin{aligned} A(t) &= (r + \frac{1}{2\gamma}\frac{\mu_m^2}{\sigma_m^2})(1-\gamma)(T-t) + \frac{\lambda_1+\lambda_2}{2}(T-t) \\ &\quad -\frac{\gamma}{2}\ln\left(\frac{\lambda_1+\lambda_2}{2}\left(\frac{e^{\frac{\xi}{\gamma}(T-t)} - e^{-\frac{\xi}{\gamma}(T-t)}}{\xi}\right) + \frac{1}{2}(e^{\frac{\xi}{\gamma}(T-t)} + e^{\frac{-\xi}{\gamma}(T-t)})\right). \end{aligned}$$

This yields the result for $A(t)$ and $C(t)$ and so proves Lemma 1.

Proof of Lemma 2

Next consider the case with delta neutral portfolio weights where $\phi_{1t} = -\phi_{2t}$. We conjecture that

$$v(t,x) = B(t) + \frac{1}{2}D(t)x^2.$$

Substituting this equation into (25) yields an affine equation in x. Setting the term that is independent of x and the coefficient of x to zero leads to the following ODEs

$$B_t + b^2 D + (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma) = 0,$$

$$D_t - 2(\lambda_1+\lambda_2)D + 2b^2D^2 + \frac{1}{\gamma}\frac{(\lambda_1+\lambda_2)^2 - 4(\lambda_1+\lambda_2)b^2D + 4b^4D^2}{2b^2}(1-\gamma) = 0.$$

These two ODEs can be simplified to

$$\begin{aligned} B_t + b^2D + (r + \frac{1}{2\gamma}(\mu_m^2/\sigma_m^2))(1-\gamma) &= 0, \\ D_t - \frac{2}{\gamma}(\lambda_1+\lambda_2)D + \frac{2}{\gamma}b^2D^2 + \frac{1}{\gamma}\frac{(\lambda_1+\lambda_2)^2}{2b^2}(1-\gamma) &= 0. \end{aligned}$$

Let

$$D_\pm = \frac{\lambda_1+\lambda_2}{2b^2}(1 \pm \sqrt{\gamma}).$$

The solution to the ODE is then

$$D(t) = \frac{1 - e^{\frac{2(\lambda_1+\lambda_2)}{\sqrt{\gamma}}(T-t)}}{1/D_+ - e^{\frac{2(\lambda_1+\lambda_2)}{\sqrt{\gamma}}(T-t)}/D_-},$$

and

$$\begin{aligned} B(t) &= (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma)(T-t) + \frac{\lambda_1+\lambda_2}{2}(T-t) \\ &\quad -\frac{\gamma}{2}\ln\left((\lambda_1+\lambda_2)\frac{e^{\frac{\eta}{\gamma}(T-t)} - e^{-\frac{\eta}{\gamma}(T-t)}}{2\eta} + \frac{1}{2}(e^{\frac{\eta}{\gamma}(T-t)} + e^{-\frac{\eta}{\gamma}(T-t)})\right), \end{aligned}$$

where

$$\eta = (\lambda_1+\lambda_2)\sqrt{\gamma}.$$

Proof of Proposition 5

The wealth dynamics of a self-financing trading strategy satisfies

$$dW_t = W_t\Big(rdt + (\phi_{mt} + \beta(\phi_{1t} + \phi_{2t}))(\mu_m dt + \sigma_m dB_t) + \phi_{1t}(\sigma dZ_t + bdZ_{1t} - \lambda_1 x_t dt) + \phi_{2t}(\sigma dZ_t + bdZ_{2t} + \lambda_2 x_t dt)\Big).$$

Let $\tilde{\phi}_{mt} = \phi_{mt} + \beta(\beta_{1t} + \beta_{2t})$. Note that maximizing over (ϕ_m, ϕ_1, ϕ_2) is equivalent to maximizing over $(\tilde{\phi}_m, \phi_1, \phi_2)$.

The HJB equation is

$$\begin{aligned}0 = \max\ & J_t + \ (\mu_x - \lambda_x x)' J_x + \frac{1}{2}\mathrm{Tr}\Big((\beta_x \sigma_m \sigma_m' \beta_x' + \sigma_x \sigma_x' + b_x b_x')J_{xx'}\Big)\\ & + \Big(r + \tilde{\phi}_m' \mu_m - \phi' \lambda x\Big) W J_W\\ & + \Big(\beta_x \sigma_m \sigma_m' \tilde{\phi}_m + \sigma_x \sigma' \phi + \alpha b b' \phi\Big)' W J_{xW}\\ & + \frac{1}{2}\Big(\tilde{\phi}_m' \sigma_m \sigma_m' \tilde{\phi}_m + \phi' \sigma \sigma' \phi + \phi' b b' \phi\Big) W^2 J_{WW} + \rho\left(\frac{W^{1-\gamma}}{1-\gamma} - J\right).\end{aligned}$$

We conjecture that the value function takes the form

$$J(x, W) = \frac{W^{1-\gamma}}{1-\gamma} f(x).$$

In terms of f, the HJB equation is

$$\begin{aligned}\max_{\tilde{\phi}_m, \phi_1, \phi_2}\ & 0 = f_t - \lambda_x x f_x + \frac{1}{2} b_x^2 f_{xx} + (r + \tilde{\phi}_m \mu_m + (-\lambda_1 \phi_1 + \lambda_2 \phi_2)x)(1-\gamma) f\\ & + b^2(\phi_1 - \phi_2)(1-\gamma) f_x + \frac{1}{2}(\tilde{\phi}_m^2 \sigma_m^2 + (\phi_1 + \phi_2)^2 \sigma^2 + (\phi_1^2 + \phi_2^2) b^2)(1-\gamma)(-\gamma) f + \rho(1 - f).\end{aligned}$$

The first order condition for $\tilde{\phi}_m$ is

$$\mu_m + \sigma_m^2 \tilde{\phi}_m(-\gamma) = 0,$$

which leads to

$$\tilde{\phi}_m^* = \frac{1}{\gamma}\frac{\mu_m}{\sigma_m^2}.$$

The first-order conditions for ϕ_1 and ϕ_2 are

$$\begin{aligned}-\lambda_1 x + b^2 f_x/f + ((\phi_1 + \phi_2)\sigma^2 + \phi_1 b^2)(-\gamma) &= 0,\\ \lambda_2 x - b^2 f_x/f + ((\phi_1 + \phi_2)\sigma^2 + \phi_2 b^2)(-\gamma) &= 0.\end{aligned}$$

These equations lead to

$$\begin{aligned}\phi_1^* &= \frac{-\lambda_1 b^2 - (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x + \frac{1}{\gamma} f_x/f,\\ \phi_2^* &= \frac{\lambda_2 b^2 + (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x - \frac{1}{\gamma} f_x/f.\end{aligned}$$

Substituting the optimal weights back into the HJB equation, we get

$$
\begin{aligned}
&0 = f_t + (-\lambda_x x) f_x + \frac{1}{2} b_x^2 f_{xx} + (r + \frac{1}{2\gamma}\mu_m^2/\sigma_m^2)(1-\gamma) f \\
&+\frac{1}{2\gamma}\left(2(f_x/f)^2 b^2 - 2(\lambda_1+\lambda_2) x f_x/f + \frac{\lambda_1^2(\sigma^2+b^2) + 2\lambda_1\lambda_2^2\sigma^2 + \lambda_2^2(b^2+\sigma^2)}{b^2(b^2+2\sigma^2)} x^2\right)(1-\gamma) f \\
&+\rho(1-f).
\end{aligned}
$$

This equation be can simplified to

$$
\begin{aligned}
&0 = f_t - \frac{\lambda_x}{\gamma} x f_x + \frac{1}{2} b_x^2 f_{xx} + \frac{1}{\gamma} b^2 (1-\gamma) f_x^2/f \\
&+\left(r + \frac{1}{2\gamma}\left(\mu_m^2/\sigma_m^2 + \frac{\lambda_1^2(\sigma^2+b^2) + 2\lambda_1\lambda_2\sigma^2 + \lambda_2^2(b^2+\sigma^2)}{b^2(b^2+2\sigma^2)} x^2\right)\right)(1-\gamma) f \\
&+\rho(1-f)
\end{aligned}
$$

which is the same as the equation for a deterministic horizon with f_t replaced by $\rho(1-f)$. The boundary condition is $f(x) \to 0$ as $x \to \pm\infty$ for $\gamma > 1$.

Note that when $\gamma = 1$, the PDE for f becomes

$$
0 = f_t - \frac{\lambda_x}{\gamma} x f_x + \frac{1}{2} b_x^2 f_{xx} + \rho(1-f),
$$

and the solution is $f = 1$. Next, we use the martingale approach to obtain a closed form expression for $f(t,x)$. It is straightforward to show that the (unique) pricing kernel, π_t, is

$$
\pi_t = e^{-rt} e^{-\frac{1}{2}\frac{\mu_m^2}{\sigma_m^2} - \frac{\mu_m}{\sigma_m} dB_{mt}} e^{-\frac{1}{2}(\eta_{1t}^2+\eta_{2t}^2) - (\eta_{1t} dZ_{1t} + \eta_{2t} dZ_{2t})}.
$$

Note that

$$
\eta_{1t}^2 + \eta_{2t}^2 = \frac{(\sigma^2+b^2)(\lambda_1^2+\lambda_2^2) + 2\lambda_1\lambda_2\sigma^2}{b^2(2\sigma^2+b^2)} x_t^2 = H x_t^2,
$$

where

$$
H = \frac{(\sigma^2+b^2)(\lambda_1^2+\lambda_2^2) + 2\lambda_1\lambda_2\sigma^2}{b^2(2\sigma^2+b^2)}.
$$

The agent's objective is to solve the optimization problem

$$
\mathrm{E}_0\left[\int_0^\infty \rho e^{-\rho t} \frac{1}{1-\gamma} W_t^{1-\gamma} dt + e^{-\rho T} \frac{1}{1-\gamma} W_T^{1-\gamma}\right],
$$

subject to infinitely many constraints, each indexed by t, for $t > 0$:

$$
\mathrm{E}_0[\pi_t W_t] = w_0.
$$

The Lagrangian is

$$
\mathrm{E}_0\left[\int_0^\infty \rho e^{-\rho t} \frac{1}{1-\gamma} W_t^{1-\gamma} dt - \kappa_t(\mathrm{E}_0[\pi_t W_t] - w_0) dt\right]
$$

where κ_t is the Lagrangian multiplier corresponding to the 't' constraint. The first order condition is

$$\rho e^{-\rho t} W_t^{-\gamma} - \kappa_t \pi_t = 0.$$

Thus, solving for W_t^*,

$$W_t^* = (e^{\rho t} \kappa_t \pi_t / \rho)^{-1/\gamma}.$$

From the constraint, we can derive the Lagrangian multiplier κ_t

$$\kappa_t^{-1/\gamma} = (e^{\rho t}/\rho)^{1/\gamma} (\mathrm{E}_0[\pi_t^{1-1/\gamma}])^{-1} w_0.$$

Inserting this in the wealth expression, we have

$$W_t^* = \pi_t^{-1/\gamma} (\mathrm{E}_0[\pi_t^{1-1/\gamma}])^{-1} w_0.$$

The value function is

$$\begin{aligned} J(w_0, x) &= \int_0^\infty \rho e^{-\rho t} \mathrm{E}_0 \left[\frac{W_t^{*1-\gamma}}{1-\gamma} \right] dt = \int_0^\infty \rho e^{-\rho t} \mathrm{E}_0 \left[\frac{\left(\pi_t^{-1/\gamma} (\mathrm{E}_0[\pi_t^{1-1/\gamma}])^{-1} w_0 \right)^{1-\gamma}}{1-\gamma} \right] dt \\ &= \frac{w_0^{1-\gamma}}{1-\gamma} \int_0^\infty \rho e^{-\rho t} \mathrm{E}_0 \left[\pi_t^{1-1/\gamma} \left(\mathrm{E}_0[\pi_t^{1-1/\gamma}] \right)^{\gamma-1} \right] dt = \frac{w_0^{1-\gamma}}{1-\gamma} \int_0^\infty \rho e^{-\rho t} \left(\mathrm{E}_0[\pi_t^{1-1/\gamma}] \right)^{\gamma} dt. \end{aligned}$$

Using the equation for the pricing kernel, we have

$$\mathrm{E}_0[\pi_t^{1-1/\gamma}] = e^{-(1-1/\gamma)rt - \frac{1}{2\gamma} \frac{(1-1/\gamma)\mu_m^2}{\sigma_m^2} t} \mathrm{E}_0[e^{-\frac{1}{2}(1-1/\gamma) \int_0^t (\eta_{1u}^2 + \eta_{2u}^2) du - (1-1/\gamma)(\eta_{1u} dZ_{1u} + \eta_{2t} dZ_{2u})}].$$

Note that

$$\begin{aligned} & \mathrm{E}_0 \left[e^{-\frac{1}{2}(1-1/\gamma) \int_0^t (\eta_{1u}^2 + \eta_{2u}^2) du - (1-1/\gamma)(\eta_{1u} dZ_{1u} + \eta_{2t} dZ_{2u})} \right] \\ = \; & \mathrm{E}_0^Q \left[e^{-\frac{1}{2\gamma}(1-1/\gamma) \int_0^t (\eta_{1u}^2 + \eta_{2u}^2) du} \right] = \mathrm{E}_0^Q \left[e^{-\frac{H}{2\gamma}(1-1/\gamma) \int_0^T x_u^2 du} \right], \end{aligned}$$

where E_0^Q denotes the expectation under the equivalent measure Q defined by the following Radon-Nikodym derivative

$$\frac{dQ}{dP} = e^{-\frac{1}{2}(1-1/\gamma)^2 \int_0^(t\eta_{1u}^2 + \eta_{2u}^2) du - (1-1/\gamma)(\eta_{1u} dZ_{1u} + \eta_{2t} dZ_{2u})}.$$

Under Q, the dynamics of x_t is

$$dx = -\frac{\lambda_x}{\gamma} x_t dt + b dZ_{1t}^Q - b dZ_{2t}^Q.$$

From Feynman-Kac, we know that

$$l(x,t) = \mathrm{E}_t^Q \left[e^{-\frac{H}{2\gamma}(1-1/\gamma) \int_t^T x_u^2 du} \right]$$

satisfies the following PDE

$$l_t - \frac{\lambda_x}{\gamma} x l_x + \frac{1}{2} b_x^2 l_{xx} - \frac{H}{2\gamma}(1 - 1/\gamma) x^2 l = 0,$$

with the terminal condition

$$l(x, T) = 1.$$

We conjecture that

$$l(x, t) = e^{h_0(t) + \frac{1}{2} h_1(t) x^2}.$$

The equation for $l(t, x)$ then becomes

$$\frac{dh_0}{dt} + \frac{1}{2}\frac{dh_1}{dt} x^2 - \lambda_x h_1(t) x^2 + \frac{1}{2} b_x^2 (h_1 + h_1^2 x^2) - \frac{H}{2\gamma}(1 - 1/\gamma) x^2 = 0.$$

Hence,

$$\begin{aligned} \frac{dh_0}{dt} + \frac{1}{2} b_x^2 h_1 &= 0, \\ \frac{dh_1}{dt} - 2\frac{\lambda_x}{\gamma} h_1(t) + b_x^2 h_1^2 - \frac{H}{\gamma}(1 - 1/\gamma) &= 0, \end{aligned}$$

subject to the terminal conditions

$$\begin{aligned} h_0(T) &= 0, \\ h_1(T) &= 0. \end{aligned}$$

The solution to $h_0(t)$ and $h_1(t)$ is given by $h_0(t) = \frac{A(t)}{\gamma}$ and $h_2(t) = \frac{C(t)}{\gamma}$, where $A(t)$ and $C(t)$ are given in Lemma 1. Therefore, we have

$$f(x, t) = \int_t^T \rho e^{-\rho(T-u)} e^{A(u) + \frac{1}{2} C(u) x^2} du + e^{-\rho(T-t)} e^{A(t) + \frac{1}{2} C(t) x^2}$$

where $A(t)$ and $C(t)$ are given in Lemma 1. The value function under a stochastic horizon specified by a Poisson arrival process is the weighted average of the value function under deterministic horizons with weights given by the Poisson distribution. The optimal portfolio weights are given by

$$\begin{aligned} \phi_1^* &= \frac{-\lambda_1 b^2 - (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x + \frac{1}{\gamma} f_x / f \\ &= \frac{-\lambda_1 b^2 - (\lambda_1 + \lambda_2)\sigma^2}{\gamma b^2 (b^2 + 2\sigma^2)} x \\ &+ \frac{\int_t^T C(u) e^{-\rho(T-u)} e^{A(u) + \frac{1}{2} C(u) x^2} du + C(T-t) e^{-\rho(T-t)} e^{A(t) + \frac{1}{2} C(t) x^2}}{\gamma \int_t^T e^{-\rho(T-u)} e^{A(u) + \frac{1}{2} C(u) x^2} du + e^{-\rho(T-t)} e^{A(t) + \frac{1}{2} C(t) x^2}} x \end{aligned}$$

and

$$
\begin{aligned}
\phi_2^* &= \frac{\lambda_2 b^2 + (\lambda_1+\lambda_2)\sigma^2}{\gamma b^2(b^2+2\sigma^2)} x - \frac{1}{\gamma} f_x / f \\
&= \frac{\lambda_2 b^2 + (\lambda_1+\lambda_2)\sigma^2}{\gamma b^2(b^2+2\sigma^2)} x \\
&- \frac{\int_t^T C(u) e^{-\rho(T-u)} e^{A(u)+\frac{1}{2}C(u)x^2} du + C(T-t) e^{-\rho(T-t)} e^{A(t)+\frac{1}{2}C(t)x^2}}{\gamma \int_t^T e^{-\rho(T-u)} e^{A(u)+\frac{1}{2}C(u)x^2} du + e^{-\rho(T-t)} e^{A(t)+\frac{1}{2}C(t)x^2}} x.
\end{aligned}
$$

When $\gamma = 1$, $A(t) = 0$ and $C(t) = 0$, and so

$$
f(x) = \int_0^T \rho e^{-\rho t} dt + e^{-\rho T} = 1,
$$

which is consistent with the solution using the dynamic programming approach.

References

[1] Alexander, Carol, 1999, Optimal Hedging Using Cointegration, Philosophical Transactions of the Royal Society, Series A, 357, 2039-2058.

[2] Bertsimas, D., L. Kogan, and A. Lo, 2000. When Is Time Continuous? Journal of Financial Economics 55, 173-204.

[3] Bondarenko, Oleg, 2003, Statistical Arbitrage and Securities Prices. Review of Financial Studies 16, 875-919.

[4] Bossaerts, Peter and Richard Green, 1989, A General Equilibrium Model of Changing Risk Premia: Theory and Evidence. Review of Financial Studies 2, 467-493.

[5] Brennan, Michael J. and Ashley Wang, 2006, Asset Pricing and Mispricing. UCLA working paper.

[6] Brenner, Robin J. and Kenneth F. Kroner, 1995, Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets. Journal of Financial and Quantitative Analysis 30, 23-42

[7] Chen, Zhiwu and Peter J. Knez, 1995, Measurement of Market Integration and Arbitrage. Review of Financial Studies 8, 287-325.

[8] Dumas, Bernard, Alexander Kurshev and Raman Uppal, 2009, Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility. Journal of Finance 64, 579-629.

[9] Engle, Robert F., and Clive W.J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica 55, 251-276..

[10] Gatev, Evan, William N. Goetzmann and K. Geert Rouwenhorst, 2006, Pairs Trading: Performance of a Relative-Value Arbitrage Rule. Review of Financial Studies 19, 797-827.

[11] Hasbrouck, Joel, 2003, Intraday Price Formation in U.S. Equity Index Markets. Journal of Finance 58, 2375-2399.

[12] Hogan, Steve, Robert Jarrow, Melvyn Teo and Mitch Warachka, 2004, Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. Journal of Financial Economics 73, 525-565.

[13] Jurek, Jakub and Halla Yang, 2007, Dynamic Portfolio Selection in Arbitrage. Harvard University working paper.

[14] Kondor, Peter, 2009, Risk in Dynamic Arbitrage: Price Effects of Convergence Trading. Journal of Finance 64(2), 638-658.

[15] Krishnamurthy, A., 2010, Amplification Mechanisms in Liquidity Crises. American Economic Journal - Macroeconomics 2(3), 1-30.

[16] Lehmann, Bruce N., 2002, Some Desiderata for the Measurement of Price Discovery Across Markets. Journal of Financial Markets 5, 259-276.

[17] Liu, Hong and Mark Loewenstein, 2002, Optimal Portfolio Selection with Transaction Costs and Finite Horizons. Review of Financial Studies 15(3), 805-835.

[18] Liu, Jun, 2007, Portfolio Selection in Stochastic Environments, Review of Financial Studies 20, 1-39.

[19] Liu, Jun and Francis Longstaff, 2004, Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities, Review of Financial Studies 17, 611-641.

[20] Liu, Jun, Ehud Peleg, and Avanidhar Subrahmanyam, 2010, Information, Expected Utility, and Portfolio Choice. Journal of Financial and Quantitative Analysis 45, 1221-1251.

[21] Merton, Robert C, 1970, Optimum Consumption and Portfolio Rules in a Continuous-Time Model, Journal of Economic Theory 3, 373-413.

[22] Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of Risk and Return in Risk Arbitrage. Journal of Finance 56, 2135-2175.

[23] Shleifer, Andrei and Robert W. Vishny, 1997, The Limits of Arbitrage. Journal of Finance 52, 35-55.

[24] Xiong, W., 2001, Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets. Journal of Financial Economics 62, 247-292.

Stock	Regular Prices					Orthogonalized Prices			
	Coint. test	λ_1	t-stat (λ_1)	λ_2	t-stat (λ_2)	λ_1	t-stat (λ_1)	λ_2	t-stat (λ_2)
Agricultural Bank of China	-2.261	0.288	2.100	0.309	1.009	0.192	1.564	0.473	1.613
China Merchants Bank	-4.281***	-0.370	-1.981	0.653	3.053	0.023	0.163	0.622	3.177
Bank of China	-2.879	0.086	0.660	0.291	1.773	0.308	2.968	-0.481	-3.100
China Citic Bank	-4.623***	0.314	1.733	0.269	1.285	0.526	3.962	-0.346	-1.754
China Minsheng Banking	-2.333	-0.284	-1.727	0.606	2.890	-0.235	-1.715	0.540	2.679
China Con. Bank	-3.189*	0.135	0.868	0.377	1.745	-0.029	-0.247	0.679	3.435
Bank of Commerce	-3.603**	-0.137	-0.780	0.612	2.843	-0.141	-1.061	0.593	2.976

Table 1: **Cointegration estimates for pairs of Chinese Bank A and H-shares.** This table reports parameter estimates from a cointegration model fitted to the log-prices of pairs of Chinese banks traded as A-shares in China and as H-shares in Hong Kong. All prices have been converted into a common currency (US dollars). The first column reports a test for cointegration between the two log-prices. The second and fourth column reports estimates of the loadings on the error-correction terms for the price process in China (λ_1) and Hong Kong (λ_2). We consider both the regular stock prices (columns 2-5) as well as prices that have been orthogonalized with respect to a common China market index (columns 6-9). For the cointegration test, * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

	Regular Prices							
	Mean		Std		Sharpe		Wealth	
Stock	unc	con	unc	con	unc	con	unc	con
Agricultural Bank of China	8.185	8.048	9.862	9.755	0.828	0.823	112.757	112.534
China Merchants Bank	16.677	6.902	14.300	6.539	1.165	1.052	231.925	143.235
Bank of China	9.878	8.437	19.475	16.103	0.506	0.523	156.519	149.204
China Citic Bank	18.568	18.299	26.538	26.769	0.699	0.683	205.692	202.500
China Minsheng Banking	16.110	8.340	15.388	6.907	1.046	1.205	138.702	119.419
China Con. Bank	6.305	6.304	15.398	13.529	0.408	0.464	125.077	126.549
Bank of Commerce	21.690	11.469	23.239	14.708	0.932	0.778	246.202	163.590

	Orthogonalized Prices							
	Mean		Std		Sharpe		Wealth	
Stock	unc	con	unc	con	unc	con	unc	con
Agricultural Bank of China	11.082	9.160	13.136	11.290	0.842	0.810	117.280	114.217
China Merchants Bank	19.284	16.195	14.846	12.091	1.298	1.338	265.633	229.510
Bank of China	10.460	3.292	19.227	11.306	0.543	0.289	161.831	115.979
China Citic Bank	23.323	6.293	24.825	13.981	0.939	0.449	263.167	128.917
China Minsheng Banking	14.078	7.595	13.567	6.384	1.036	1.187	133.422	117.575
China Con. Bank	10.647	9.823	27.611	21.235	0.385	0.462	134.949	139.276
Bank of Commerce	20.804	10.728	22.386	13.952	0.928	0.768	238.218	158.755

Table 2: **Results from daily trading of pairs of Chinese banking shares.** This table reports the performance of unconstrained and constrained delta-neutral trading strategies based on daily rebalancing of the holdings in the Chinese A share, Hong-Kong H share and the market portfolio. All results assume that trading ceases on February 15, 2012. The results assume $\mu_m = 0.05$, $\sigma_m = 0.35$, $r = 0.02$, $\gamma = 4$, and use empirical estimates of the remaining stock-specific parameters. All numbers are reported in percentage terms and ignore transaction costs.

Figure 1: **Comparison of optimal portfolio holdings under recurring versus non-recurring arbitrage opportunities.** This figure plots the optimal holdings under recurring (ϕ_1^{rec} and ϕ_2^{rec}) and the optimal holdings under non-recurring arbitrage opportunities ($\phi_1^{non-rec}$ and $\phi_2^{non-rec}$) as a function of the log-price differential, x. The results are illustrated for two different combinations of the loadings on the error correction term, λ_1 and λ_2, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong.

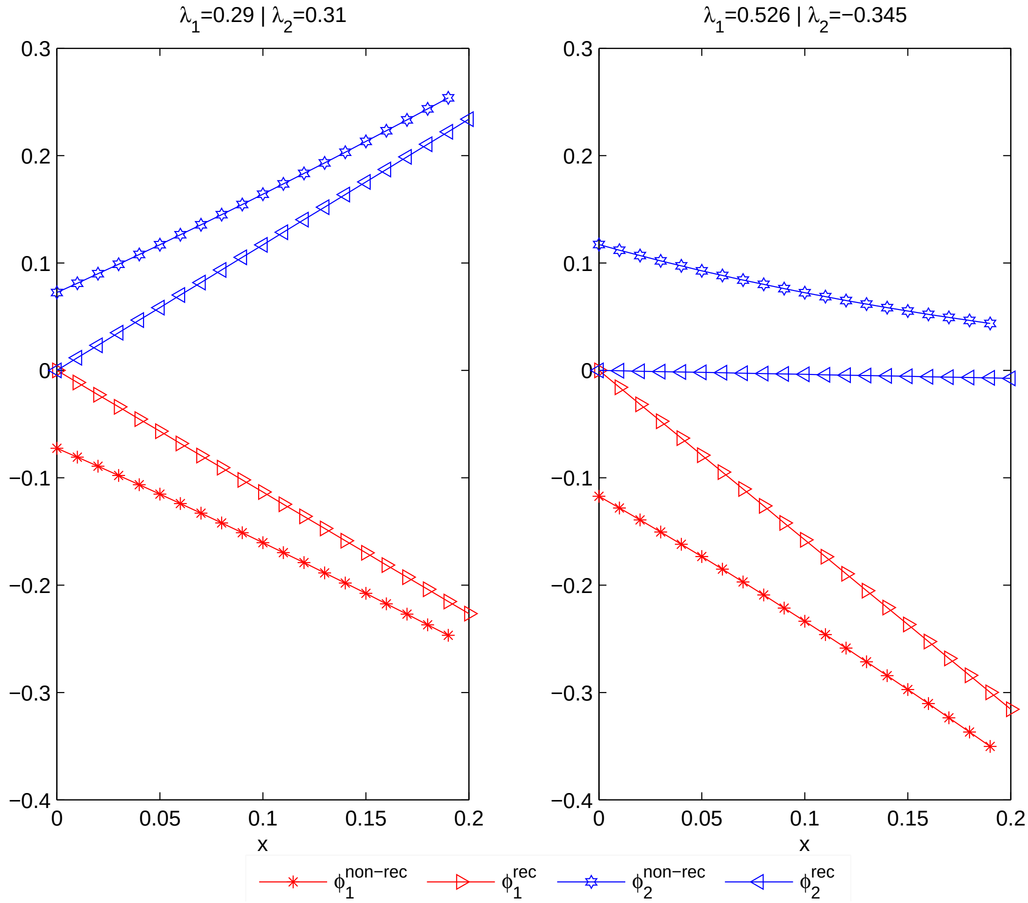

Figure 2: **Optimal positions in assets 1 and 2 as a function of the horizon, T-t.** This figure plots the optimal holdings under both recurring and non-recurring arbitrage opportunities, holding the price difference between the two assets fixed at 20%. The results are illustrated for two different combinations of the loadings on the error correction term, λ_1 and λ_2, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong.

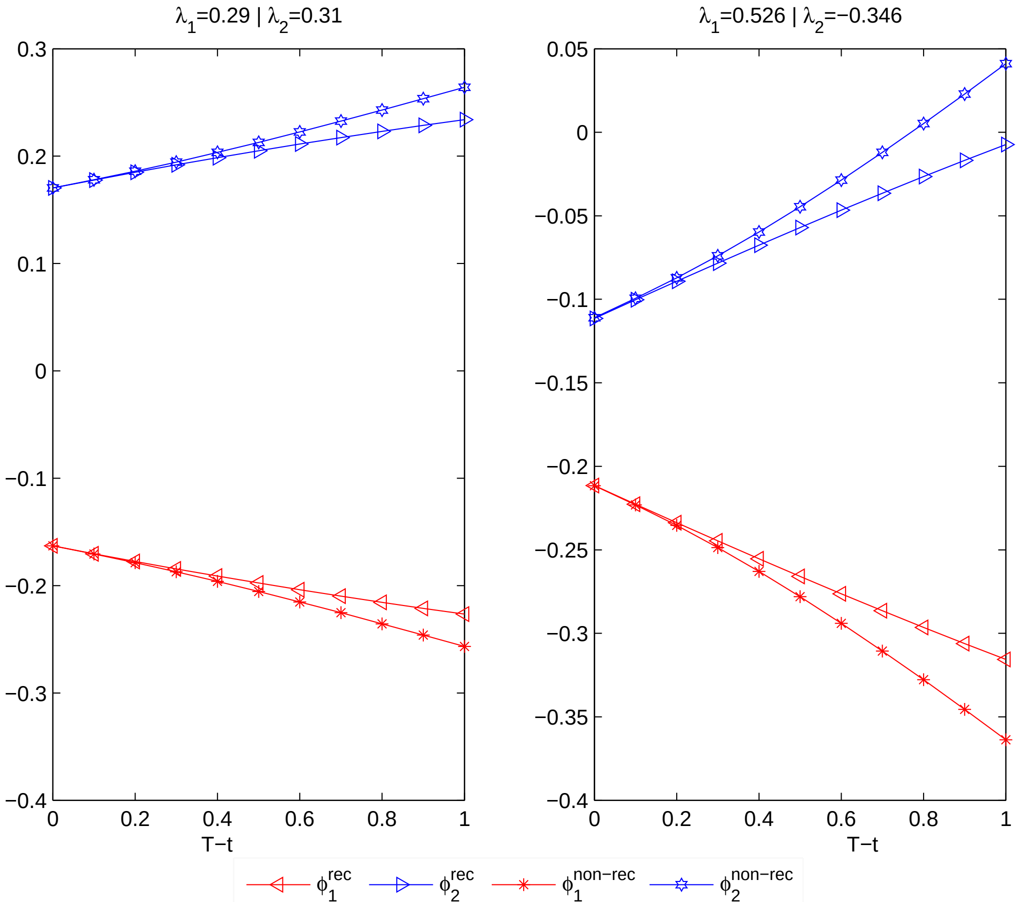

Figure 3: **Optimal holdings under recurring arbitrage opportunities: constrained versus unconstrained solutions.** This figure plots the optimal unconstrained holdings (ϕ_1^* and ϕ_2^*) and the optimal constrained holdings (ϕ_1 and ϕ_2) as a function of the log-price differential, x. The results are illustrated for two different combinations of the loadings on the error correction term, λ_1 and λ_2, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong. The figure assumes recurring arbitrage opportunities.

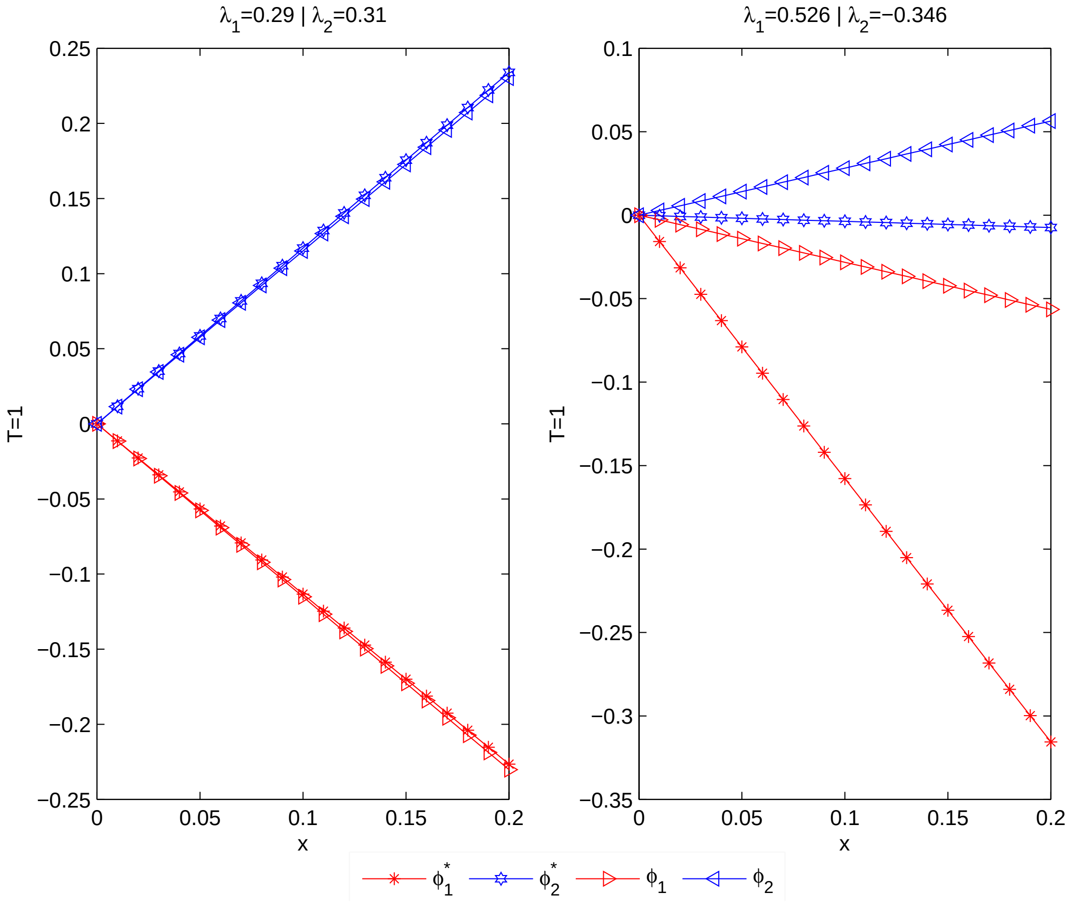

Figure 4: **Optimal holdings under non-recurring arbitrage opportunities.** This figure plots the optimal unconstrained holdings (ϕ_1^* and ϕ_2^*) and the optimal constrained holdings (ϕ_1 and ϕ_2) as a function of the log-price differential, x. The results are illustrated for two different combinations of the loadings on the error correction term, λ_1 and λ_2, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong. The figure assumes non-recurring arbitrage opportunities, as positions are closed down when the price differential crosses zero.

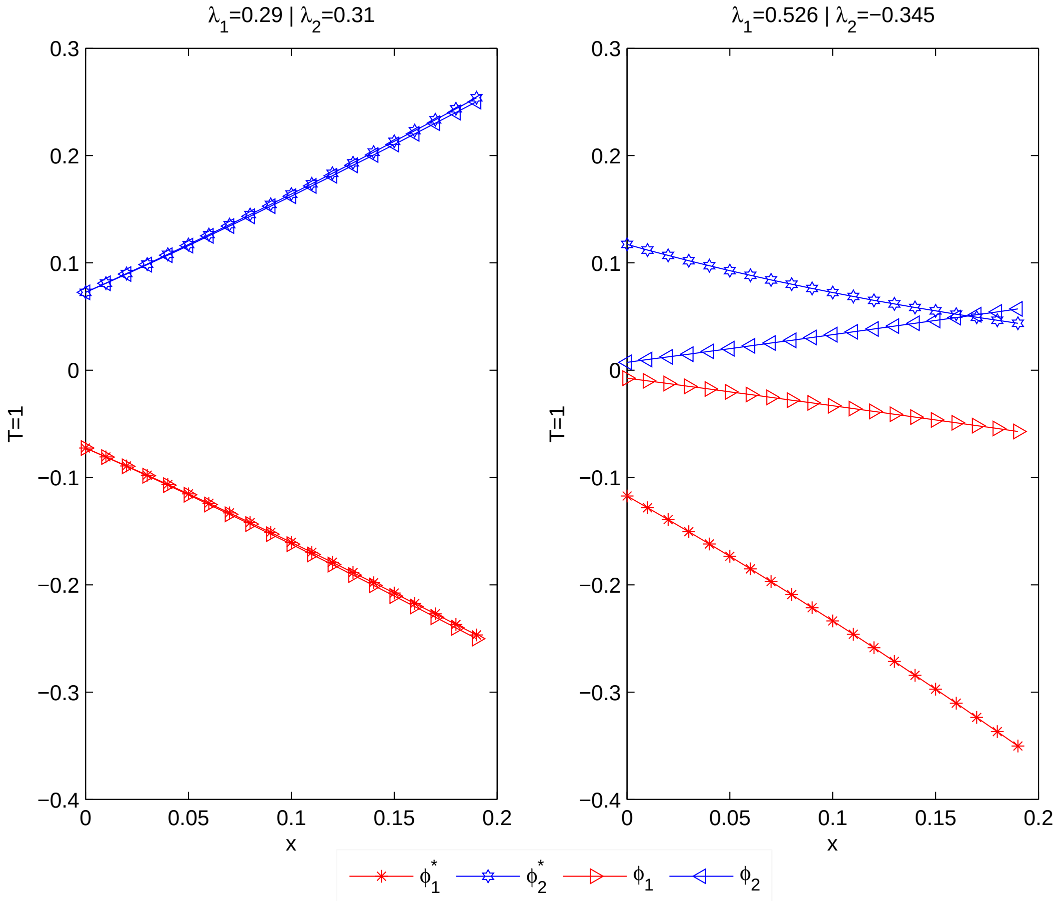

Figure 5: **Wealth gain under recurring and non-recurring arbitrage opportunities.** The plot shows the percentage wealth gain (in percentage of initial wealth) from not imposing the constraint that the position be delta neutral, as a function of the log-price differential, x. The results are illustrated for $\lambda_1 = 0.526$ and $\lambda_2 = -0.346$, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong.

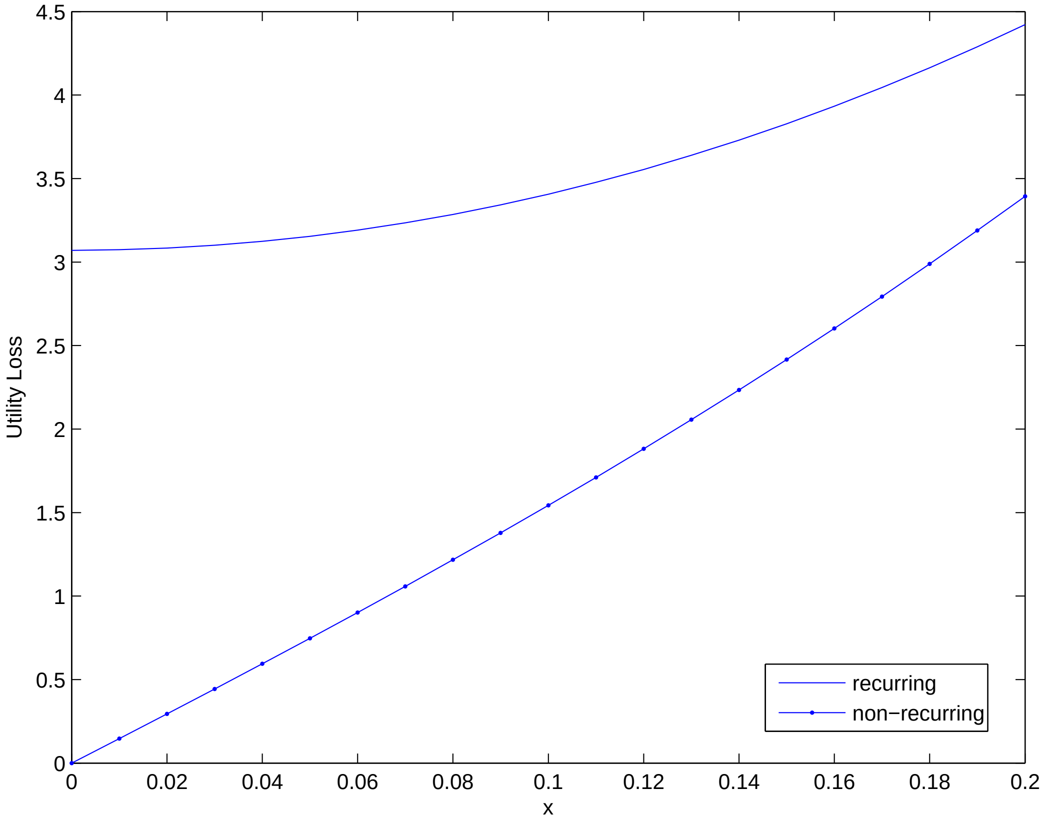

Figure 6: **Optimal portfolio holdings under recurring arbitrage opportunities with stochastic (Poisson) horizon versus fixed horizon.** This figure plots the optimal unconstrained holdings under a stochastic (Poisson) horizon ($\phi_1^{Poisson}$ and $\phi_2^{Poisson}$) versus the optimal unconstrained holdings under a fixed horizon (ϕ_1^{rec} and ϕ_2^{rec}) as a function of the log-price differential, x. The results are illustrated for two different combinations of the loadings on the error correction term, λ_1 and λ_2, with values based on the estimates fitted to pairs of Chinese banks traded in China and Hong Kong.

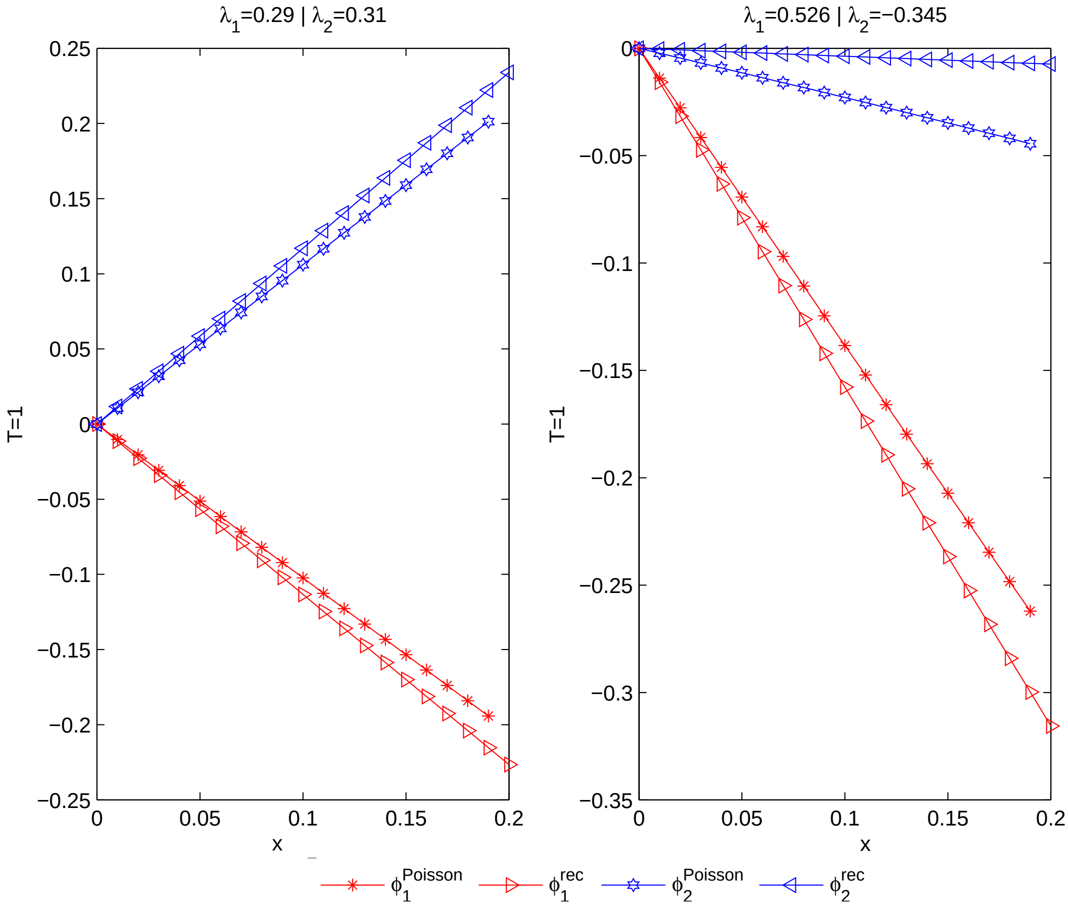

Figure 7: **Time series plots of pairs of Chinese banking shares.** This figure plots stock prices for pairs of Chinese banks traded as A shares in China and as H shares in Hong Kong.

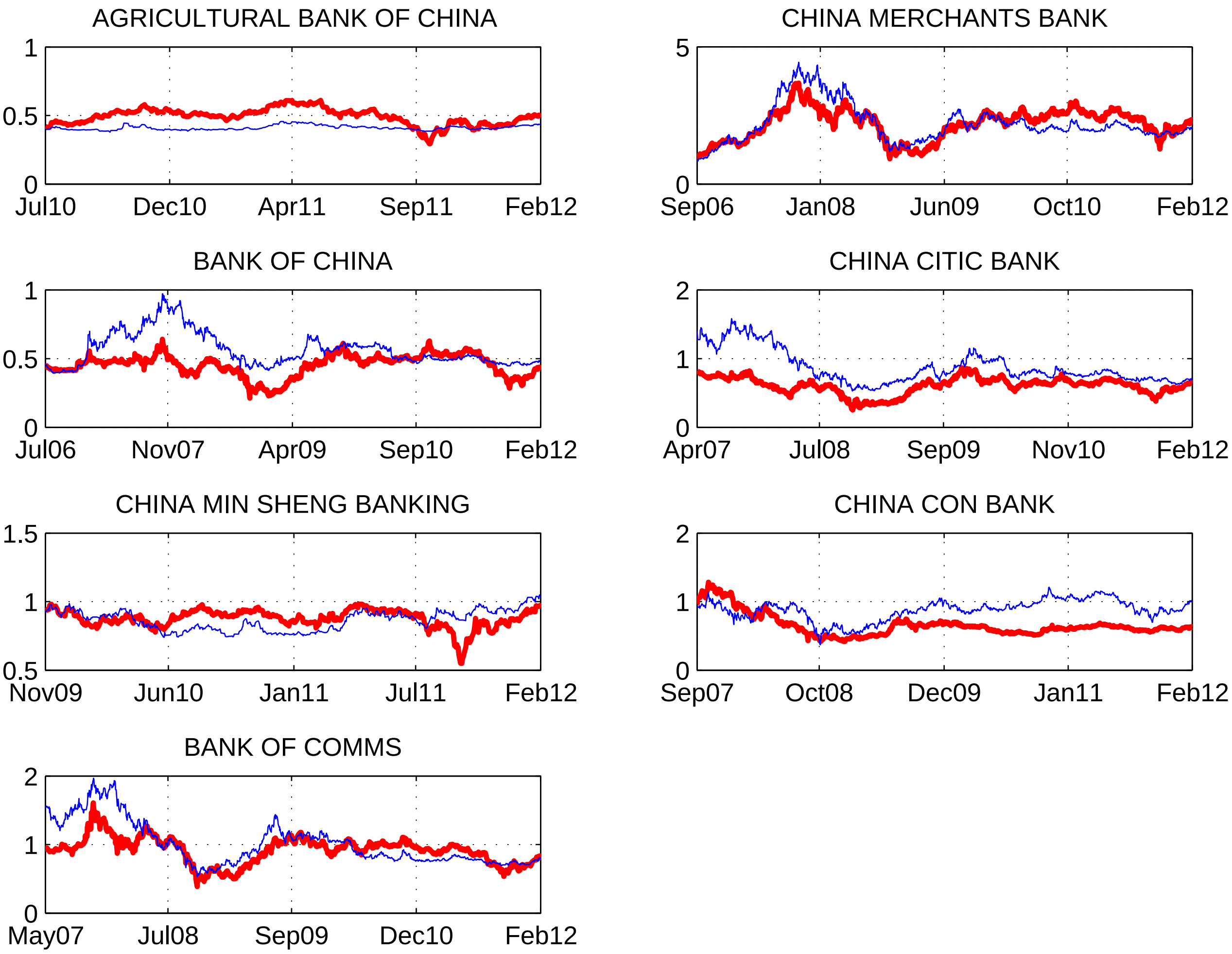

Figure 8: **Time series plots of price difference between pairs of Chinese banking shares.** This figure plots the price difference between pairs of Chinese banks traded as A shares in China and as H shares in Hong Kong.