

# LOSS PROTECTION IN PAIRS TRADING THROUGH MINIMUM PROFIT BOUNDS: A COINTEGRATION APPROACH

YAN-XIA LIN, MICHAEL McCRAE, AND CHANDRA GULATI

*Received 4 September 2005; Revised 10 May 2006; Accepted 15 May 2006*

Pairs trading is a comparative-value form of statistical arbitrage designed to exploit temporary random departures from equilibrium pricing between two shares. However, the strategy is not riskless. Market events as well as poor statistical modeling and parameter estimation may all erode potential profits. Since conventional loss limiting trading strategies are costly, a preferable situation is to integrate loss limitation within the statistical modeling itself. This paper uses cointegration principles to develop a procedure that embeds a minimum profit condition within a pairs trading strategy. We derive the necessary conditions for such a procedure and then use them to define and implement a five-step procedure for identifying eligible trades. The statistical validity of the procedure is verified through simulation data. Practicality is tested through actual data. The results show that, at reasonable minimum profit levels, the protocol does not greatly reduce trade numbers or absolute profits relative to an unprotected trading strategy.

Copyright © 2006 Yan-Xia Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

Pairs trading is a statistical arbitrage strategy with a long history of modest but persistent profits on Wall St (Peskin and Boudreau [10]; Gatev et al. [3]). The strategy identifies pairs of shares whose prices are driven by the same economic forces, then trades on any temporary deviations of those two-share prices from their long-run average relationship (Gillespie and Ulph [4]). The arbitrage or risk-free nature of the strategy arises from the opening of opposing positions for each trade—shorting the over-valued share and buying the under-valued share.

The simple statistical techniques used for share pairs selection and trading decisions make pairs trading an appealing arbitrage strategy. But simplicity comes at a cost. Correlation, covariance, and regression analysis of share price associations provide an imprecise, simplistic statistical definition of a long-run equilibrium relationship between share

## 2 Pairs trading based on cointegration approach

prices. Moreover, they do not necessarily imply mean reversion to a long-run equilibrium price spread.

This paper assumes that such deficiencies of the statistical techniques are best dealt with by systematic improvement within the underlying statistical modeling itself, rather than left to costly hedging and conditional order techniques. We use cointegration theory to provide a statistically precise foundation for the decisions involved in pairs trading and then use these principles to derive a loss limiting rule that ensures that each eligible trade will return some preset minimum profit, subject, of course, to the previously mentioned endemic market risks which are always present.

In this paper we use the principle of cointegrated series to derive a precise, dynamic definition of long-run equilibrium price spread that inherently implies mean reversion in component series. We then use cointegration principles to establish a protocol for ensuring that any selected trade will satisfy preset minimum profit conditions.

The paper extends the work on cointegration in pairs trading by Gatev et al. [3], Gillespie and Ulph [4], and Alexander and Dimitriu [1] to integrate a minimum nominal profit requirement into other trading strategy decisions such as the choices about share pairs, dollar weighting of long/short positions, trade opening and closing criteria and total dollar investment. We first use cointegration coefficient weighting (CCW) principles to derive the necessary conditions that will ensure that a trade delivers a preset minimum nominal profit per trade (MNPPT). These conditions are then incorporated into a practical, five step procedure for achieving any given MNPPT.

The analysis proceeds in six sections. Section 2 summarizes pairs trading fundamentals and identifies the main parameter estimates required for a pairs trading strategy. Section 3 introduces the concept of cointegration-based dollar weighting of long/short positions as the theoretical foundation for deriving the necessary conditions of a preset minimum profit per trade. A five step procedure for putting the necessary conditions into practice is presented at the end of Section 3. Section 4 uses the simulated data series to investigate the procedure's sensitivity to alternative opening trade hurdle values under two trading conditions: un-constrained and constrained total investment dollars. Actual daily share price data is used in Section 5 to examine the effect of investment dollar constraints on the number of valid trades for six preset minimum profit levels, with decreasing open condition values at each level. Section 6 discusses the risk minimization implications of our results in the context of arbitrage trading strategies.

The data simulation exercise in Section 4 indicates that while all trades can be immunized in a theoretical sense, the crucial factor that determines the number of eligible trades is the allowable investment dollar maximum, since some trades require large outlays. However, real data are needed to test the practical limitations of the capital requirement.

The practicality of imposing minimum profit conditions is tested on daily closing price data for two Australian Stock Exchange quoted bank shares—the Australia New Zealand Bank (ANZ) and the Adelaide Bank (ADB) over the period January 2, 2001 to August 30, 2002. The results show that trading strategy profit potential is not unduly constrained by adding a reasonable minimum profit condition to protect against losses.

## 2. Pairs trading and statistical arbitrage

Pairs trading relies on the principle of equilibrium pricing for near-equivalent shares. In efficient markets, capital asset pricing model-based valuation theory and the law of one price require price equality for equivalent financial assets over time (Reilly and Brown [11]; Sharpe et al. [12]). The price spreads of near-equivalent assets should also conform to a long-term stable equilibrium over time. Hendry and Juselius [6] use this principle to show that short-term deviations from these equivalent pricing conditions may create opportunities for arbitrage profits depending upon the size and duration of the price shock.

When a sufficiently large deviation of price spread from the long-run norm is identified, a trade is opened by simultaneously buying (go long) the under-valued share and selling (short) the over-valued share. The trade is closed out when prices return to their equilibrium price spread levels by selling the long position and off-setting the short position. Net trading profit sums the profits from the long and short positions, calculated as the difference between the opening and closing prices (net of trading costs less interest on short sale receipts). See Gillespie and Ulph [4] and L'Habitant [8].

The “risk free” characteristic of pairs trading arises from the simultaneous long-short (buy-sell) opening market positions. The opposing positions ideally immunize trading outcomes against systematic market-wide movements in prices that may work against uncovered positions (see Jacobs and Levy [7]).

But arbitrage trading of the “convergence trade” type is rarely risk-less. Market events, persistent pricing inefficiencies or structural price changes may invalidate statistical pricing models, confound future price expectations or require parameter reestimation. Price spreads after position opening may escalate rather than revert, or the equilibrium position may shift. The inherent nature of losses were dramatically demonstrated by the unraveling of long-term capital management’s highly leveraged long/short sovereign bond positions in the late 90s (Lowenstein [9]).

Pairs trading is also exposed to risk from the inherent limitations in the statistical techniques used to identify and extract profit potential. Traditional techniques may appeal in their simplicity but suffer severe limitations as a foundation for trading decision choices that determine arbitrage profit potential and extraction.

The profit reduction consequences of these risks may be offset by loss limitation strategies including stop loss and time limit orders and derivatives hedging. But these strategies are costly and only limit rather than prevent loss. With regard to statistical inefficiency, a preferable situation is integrating loss protection into the statistical modeling itself. This paper develops and tests such a procedure by using cointegration theory to define the necessary conditions for ensuring a minimum nominal profit before a trade is opened. The next section describes the foundation for this analysis.

## 3. Cointegration-based strategies

Alexander et al. [2] demonstrate that the arbitrage profit potential between two shares depends critically on the presence of a long-term equilibrium spread between share prices, the existence of short-run departures (price shocks) from that equilibrium and

#### 4 Pairs trading based on cointegration approach

re-convergence to equilibrium. In this situation, the statistical technique used for pairs trading must be able to provide an effective model of share price time behavior; detect equilibrium value relationships, and provide a measure of the extent and size of short-term variations from that equilibrium relationship. Gatev et al. [3], Gillespie and Ulp [4] and Alexander and Dimitriu [1] all suggest that cointegration theory offers a more integrative framework for statistical arbitrage strategies than current techniques.

*Definition 3.1.* A time series  $X_t$  is called an  $I(1)$  series if the first difference of the time series forms a stationary series, denoted by  $I(0)$ .

Many share price series are  $I(1)$  series. Therefore, the following cointegration definition is given based on  $I(1)$  series.

*Definition 3.2.* Let  $X_{1t}, X_{2t}, \dots, X_{nt}$  be a sequence of  $I(1)$  time series. If there are nonzero real numbers  $\beta_1, \beta_2, \dots, \beta_n$  such that

$$\beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_n X_{nt} \tag{3.1}$$

becomes an  $I(0)$  series, then  $X_{1t}, X_{2t}, \dots, X_{nt}$  are said to be cointegrated.

Cointegrated price series possess a stationary long-run stable equilibrium relationship with the associated property of mean reversion. By the definition, the linear combination of cointegrated price series is stationary and will always revert back to the mean of the stationary series. This is an important fact, which will ensure that the pair trading technique developed in this paper becomes practicable. Further details on cointegration analysis can be found in Harris [5].

We now use the concept of cointegrated share price series to derive the necessary conditions for ensuring a given minimum profit per trade over a selected trading horizon. The selection of share pairs and all subsequent trading decisions is made on cointegration principles. Long and short positions are weighted by their cointegration coefficients rather than being equally weighted. Trade opening boundaries are defined in terms of deviations from the linear combination of cointegrated series rather than deviations from some absolute mean-spread value.

The following assumptions simplify the analysis:

- (A1) the two-share price series are always cointegrated over the relevant time horizon;
- (A2) long and short positions always apply to the same shares in the share pair. For any trade,  $S_1$  always represents the short position while  $S_2$  represents the long position;
- (A3) at the opening of any trade, the price for the shorted share  $S_1$  is always higher than the price of the share in long position  $S_2$ .

*Remarks 3.3.* (1) Since divergence from equilibrium pricing is random, any one share in a pair is just as likely to be over-priced as under-priced. However, since we are only concerned with profit per trade and since any one's trade must be concluded before the next trade is started, assumption (A2) does not affect either the validity of the simulation or empirical tests in relation to the ability of the necessary conditions to generate greater than minimum required profits per trade.

(2) Assumption (A3) is quite general since a linear transformation of a price series has no effect on its cointegrating properties.

(3) The above simplifying assumptions do not affect the validity or generality of the “necessary conditions” tests, except that assumption (A3) under-estimates the number of possible trades in any period, since, on standardized prices, the shorted share,  $S_1$ , is just as likely to be below  $S_2$  as above.

**3.1. Profit produced by a completed pair trade.** The first step is to derive a profit formula for a pair of shares  $S_1$  and  $S_2$ .

Let  $t_o$  and  $t_c$  be the times of opening and closing out a trade position, respectively. A trade is opened at  $t_o$  when a preset open trade condition (OTC) is met. The trade is closed out at  $t_c$  when a preset close trade condition (CTC) is met.

The following notations are used in the analysis. Denote by

- $N_{S_1}(t_o)$  the number of shares in short position at  $t_o$ ;
- $N_{S_2}(t_o)$  the number of shares in long position at  $t_o$ ;
- $P_{S_1}(t_o)$  the price of share  $S_1$  at  $t_o$ ;
- $P_{S_1}(t_c)$  the price of share  $S_1$  at  $t_c$ ;
- $P_{S_2}(t_o)$  the price of share  $S_2$  at  $t_o$ ;
- $P_{S_2}(t_c)$  the price of share  $S_2$  at  $t_c$ .

A trade is opened when the OTC is met at time  $t_o$ . The over-valued share,  $S_1$ , is shorted (sold), so that  $N_{S_1}(t_o)$  shares are sold for the receipt of  $N_{S_1}(t_o)P_{S_1}(t_o)$  dollars. A long position on  $N_{S_2}(t_o)$  shares is taken in the under-valued share  $S_2$  at a cost of  $N_{S_1}(t_o)P_{S_2}(t_o)$  dollars.

The trade is then closed out when the CTC is met at time  $t_c$ , by simultaneously selling the long position shares for the receipt of  $P_{S_2}(t_c)N_{S_2}(t_o)$  dollars and buying back the  $N_{S_1}(t_o)$  of  $S_1$  shares at a cost of  $N_{S_1}(t_o)P_{S_1}(t_c)$  dollars.

Thus, the total profit from the trade is

$$TP_{t_c} = N_{S_2}(t_o)[P_{S_2}(t_c) - P_{S_2}(t_o)] + N_{S_1}(t_o)[P_{S_1}(t_o) - P_{S_1}(t_c)]. \quad (3.2)$$

A trade is profitable if and only if  $TP_{t_c} > 0$ . So a loss prevention strategy equates to ensuring that  $TP_{t_c} > 0$  or, more generally, that  $TP_{t_c} > K > 0$  for any preset value  $K$ .

**3.2. The conditions of minimum profit per trade under the CCW rule.** We now establish a cointegration coefficient weighting (CCW) rule and derive the conditions necessary to ensure a minimum profit per trade (MPPT).

Under (A1), the prices of shares  $S_1$  and  $S_2$  are cointegrated; say

$$P_{S_1}(t) + \beta P_{S_2}(t) = \varepsilon_t, \quad t \geq 1 \quad (3.3)$$

where  $\varepsilon_t$  is an  $I(0)$  series.

The following study is restricted to the situation where the cointegration coefficient  $\beta < 0$ . This condition is not restrictive since previous studies show that most cointegrated share price series conform to this condition.

## 6 Pairs trading based on cointegration approach

To ensure that the money gained from  $S_1$  at  $t_o$  will cover the outlay to buy  $S_2$  at  $t_o$ , we need the following condition for opening a trade:

$$N_{S_1}(t_o)P_{S_1}(t_o) \geq N_{S_2}(t_o)P_{S_2}(t_o). \quad (C1)$$

In general, a trade can be opened at any time as long as (C1) is satisfied. Here we introduce an open trade criterion by the following.

*Open trade condition (OTC(a)).* Let  $a$  be a positive real number. A time  $t_o$  can be considered as an open trading time if  $t_o$  satisfies the following condition:

$$P_{S_1}(t_o) + \beta P_{S_2}(t_o) = \varepsilon_{t_o} > a > 0. \quad (3.4)$$

To ensure that both conditions OTC(a) and (C1) are true, a condition on  $N_{S_1}(t_o)$  and  $N_{S_2}(t_o)$  needs to be imposed. If a trader decides to buy  $n$  shares, that is,  $N_{S_2}(t_o) = n$ , then, the trader should sell at least  $n/|\beta|$  shares in the short position. For simplicity, fractional share holdings are permitted. In this situation we will have  $P_{S_1}(t_o)N_{S_1}(t_o) > P_{S_2}(t_o)N_{S_2}(t_o)$ . So (C1) holds. After manipulation, under OTC(a) with  $N_{S_2}(t_o) = n$  and  $N_{S_1}(t_o) = n/|\beta|$ , the total profit made at time  $t_c$  can be calculated below:

$$\begin{aligned} TP_{t_c} &= N_{S_2}(t_o)[P_{S_2}(t_c) - P_{S_2}(t_o)] + N_{S_1}(t_o)[P_{S_1}(t_o) - P_{S_1}(t_c)] \\ &= \frac{n}{\beta} \{ [\varepsilon_{t_c} - P_{S_1}(t_c)] - [\varepsilon_{t_o} - P_{S_1}(t_o)] \} + \frac{n}{|\beta|} [P_{S_1}(t_o) - P_{S_1}(t_c)] \\ &= -\frac{n}{|\beta|} [P_{S_1}(t_o) - P_{S_1}(t_c)] + \frac{n}{|\beta|} [P_{S_1}(t_o) - P_{S_1}(t_c)] \\ &\quad + \frac{n}{|\beta|} (\varepsilon_{t_o} - \varepsilon_{t_c}) = \frac{n(\varepsilon_{t_o} - \varepsilon_{t_c})}{|\beta|}. \end{aligned} \quad (3.5)$$

This derivation shows that for any pair of cointegrated shares, if at open time  $t_o$  the number of shares in the long and short positions are  $N_{S_2}(t_o) = n$  and  $N_{S_1}(t_o) = n/|\beta|$ , respectively, the total profit from long/short trading can be expressed solely in terms of  $\beta$ ,  $\varepsilon_{t_o}$ ,  $\varepsilon_{t_c}$ , and  $n$ .

We now need to define an appropriate close time  $t_c$  such that a trader, who opened a trade under OTC(a) with  $N_{S_2}(t_o) = n$  and  $N_{S_1}(t_o) = n/|\beta|$ , will be able to gain a minimum of  $\$K$  when the trader closes the trade.

From (3.5), to ensure the minimum gain requirement,  $\varepsilon_{t_c}$  has to satisfy the following inequality:

$$\frac{n(\varepsilon_{t_o} - \varepsilon_{t_c})}{|\beta|} > K. \quad (3.6)$$

In other words, the value of  $\varepsilon_{t_c}$  has to be lower than  $\varepsilon_{t_o}$  and the difference between  $\varepsilon_{t_o}$  and  $\varepsilon_{t_c}$  has to be greater than  $|\beta|K/n$ . Thus, to ensure a minimum profit of  $\$K$ , we use the following closing condition.

*Close trade condition (CTC(a), (b)).* If a trade is opened at OTC(a) with  $N_{S_2}(t_o) > K|\beta|/(a - b)$  and  $N_{S_1}(t_o) = N_{S_2}(t_o)/|\beta|$ , where  $a > b$ , then the trade needs to be closed at  $t_c$  when  $\varepsilon_{t_c} < b$ .

In practice,  $N_{S_1}(t_o)$  can take value  $[N_{S_2}(t_o)/|\beta|] + 1$ , in case  $N_{S_2}(t_o)/|\beta|$  is not an integer, where “[ $d$ ]” denotes the maximum integer less than  $d$ .

**3.3. Five-step trading strategy.** We now use the above necessary conditions to build a five-step procedure for obtaining the required minimum profit  $\$K$  on any completed trade.

*Step 1.* Choose an opening condition  $a$  and closing condition  $b$  such that  $a > b$ . Usually  $b$  is assigned as the mean of  $\varepsilon_1$  and  $a$  is assigned as  $k\sigma$  where  $k$  is a positive real number and  $\sigma$  is the standard deviation of  $\varepsilon_1$  (recall that  $\varepsilon_t$  is a stationary time series).

*Step 2.* Choose an integer  $n > K|\beta|/(a - b)$ .

*Step 3.* Open a trade at  $t_o$  when  $P_{S_1}(t_o) > P_{S_2}(t_o)$  and condition OTC(a) is satisfied.

*Step 4.* Buy  $n$  shares of  $S_2$  and sell  $[n/|\beta|] + 1$  shares of  $S_1$  at time  $t_o$ .

*Step 5.* Close out the trade at  $t_c$  when  $\varepsilon_{t_c} < b$ .

Following the above steps, we have

$$\frac{n(\varepsilon_{t_o} - \varepsilon_{t_c})}{|\beta|} > \frac{n(a - b)}{|\beta|} > K, \quad (3.7)$$

which will ensure a given MPPT of  $\$K$  for the trade.

In the above strategy, the proportion of shares assigned to the long and short positions is determined by the cointegration coefficient  $\beta$  rather than by the more traditional equal weighting strategy. We label this the cointegration coefficient weighting strategy (CCW).

*Remark 3.4.* In practice, the CCW weighting strategy will always work if the total dollar investment is permitted by the broker. This is because the open and close conditions are based on the movement of the stationary time series  $\varepsilon_t$ . To ensure an appropriate frequency of trades, the opening condition (a) and the closing condition (b) should be chosen such that they are regularly crossed by the process  $\varepsilon_t$ , thus ensuring the frequent opening and closing of trades.

#### 4. The application of minimum profit conditions

The preceding analysis derived the theoretical conditions for achieving a given MPPT and formulated a five step procedure to implement the trading strategy. We now examine practical application issues of constraints imposed by the procedure on numbers of trades and sensitivity to maximum investment levels.

The theoretical derivation of the necessary conditions for achieving a given level of minimum profit may be enhanced if the procedure is a practical one in terms of its impact on trade numbers and trading profitability. The current analysis concentrates on profit per trade, trade numbers, and dollar investment implications.

## 8 Pairs trading based on cointegration approach

Table 4.1. Total profit and trades for varying MPPT levels under CCW strategy: simulated data.

$K$	Open condition * (a)	Close condition * (b)	Average total profit	Average total trades
10	$m + \sigma$	$m$	1350	51.8
50	$m + \sigma$	$m$	6006	51.8
100	$m + \sigma$	$m$	11751	51.8
10	$m + (\sigma/2)$	$m$	3248	81.7
50	$m + (\sigma/2)$	$m$	15180	81.7
100	$m + (\sigma/2)$	$m$	30032	81.7

\* $m$  and  $\sigma$  denote the mean and standard deviations, respectively, of the  $I(0)$  series  $\varepsilon_t$

We investigate the trade number and profit constraint issues using data generated from a known cointegration model. The simulation study has two purposes. First, to demonstrate how the CCW strategy works for the simulated data and whether altering the values for (a) and (b) affects the number of trades for any given MPPT level. Second, to demonstrate the effect on trade numbers of introducing a constraint on the total dollar investment allowed in any trade.

As an investigative technique, simulation enhances control over the data generation process by ensuring that sample data conform to a given cointegration model with the prescribed parameters. This filters out data “noise” that may complicate results on the effects of the treatment variable on the target variable/s.

The sample price data are simulated from a cointegration model:

$$\begin{aligned} P_{S_1}(t) + \beta P_{S_2}(t) &= \varepsilon_t, \\ P_{S_2}(t) - P_{S_2}(t-1) &= e_t, \end{aligned} \tag{4.1}$$

where  $e_t - 0.1e_{t-1} = 13 + \delta_{1,t}$  and  $\delta_{1,t}$  are iid normally distributed,  $N(0, 0.5)$ ;  $\beta = -0.2$ ;  $\varepsilon_t$  follows model  $\varepsilon_t - 0.2\varepsilon_{t-1} = 13 + \delta_t$  and  $\{\delta_t\}$  are iid with standard normal distribution  $N(0, 1)$ . 100 independent samples are simulated from this model and each sample has 500 data points equally spaced over the trading horizon to permit calculation of profit per time period over a horizon of 500 periods. Simulations are run with  $\$K$  equal to  $\$10$ ,  $\$50$ , and  $\$100$ , respectively.

**4.1. Application to the CCW strategy: unconstrained investment.** Following the five step trading strategy, the simulation output in Table 4.1 is given by setting  $N_{S_2}(t_o) = [K|\beta|/(a-b)] + 1$  and  $N_{S_1}(t_o) = N_{S_2}(t_o)/|\beta|$  for each trade.

Table 4.1 shows that the average total trade numbers per sample is just over 51. When the value (a) is closer to the mean of  $\varepsilon_t$ , the average total trade numbers increase to over 81 trades per sample.

Under the CCW rule, the number of trades in a trading horizon is largely determined by the open and close criterion values. Since both criteria now relate to the stationary time series  $\varepsilon_t$ , reversion to the long-run equilibrium value  $m$  is more frequent.

Table 4.2. Total profit and trades for varying MPPT under CCW strategy with constrained investment dollars: simulated data.

$K$	$W$	Open condition (a)	Close condition (b)	Average total profit	Average total trades
10	90000	$m + \sigma$	$m$	523	15.25
10	100000	$m + \sigma$	$m$	875	25.37
10	250000	$m + \sigma$	$m$	1932	51.78
10	100000	$m + (\sigma/2)$	$m$	0	0
10	250000	$m + (\sigma/2)$	$m$	3999	81.75
50	250000	$m + \sigma$	$m$	0	0
50	400000	$m + \sigma$	$m$	5864	46.50
50	250000	$m + 1.5\sigma$	$m$	1523	14.16
50	100000	$m + 1.5\sigma$	$m$	0	0
100	400000	$m + 1.5\sigma$	$m$	1114	5.69

**4.2. Application to the CCW strategy: constrained investment dollars.** In the previous unconstrained CCW simulation, the total dollar investment in long/short positions cannot be preset. They depend on the price of shares at each open trade position. So while minimum profit  $\$K$  requirement is met, the total dollars investment required to produce this result may be large. In this section, another simulation study is considered. This simulation constrains the total dollar investment permitted per trade. Trades that require  $\$W$  investment above the indicated values are now deleted. Table 4.2 presents the results.

The results indicate that the size of the average dollar commitment per trade necessary to meet the MPPT condition can make the rate of return on investment very small at the given entry hurdle—even when set below the prevailing risk free rate. However, recall that we are not deriving a profit maximizing strategy, but a strategy ensuring a given minimum profit. The simulations emphasize the sensitivity of required capital outlay to the other decision parameter values. Keeping outlays feasible implies the selection of realistic parameter values and reasonably priced shares relative to intended outlay. Expensive shares require more capital.

The low rate of return on investment may reflect the lack of price shocks in the simulated model of this section. A more realistic test of returns requires actual data. The next section details an empirical investigation of these results.

## 5. Application of the CCW strategy to empirical data

We now use empirical data to examine how alternative levels of maximum investment affect trade numbers for a given MPPT value  $K$ . Maximum investment limits ( $W$ ) are set at  $\$5,000$  through to  $\$100,000$  for given alternative MPPT levels  $K$  of  $\$10$  through to  $\$2000$ . Within each investment level, the opening condition is varied from  $m + 1.5\sigma$  to  $m + \sigma/5$  and the closing condition is always set at  $m$ , where  $m$  and  $\sigma$  are the mean and standard deviations of  $\varepsilon_t$  in (5.1), respectively.

The data are daily closing prices from January 2, 2001 to August 30, 2002 for two Australian Stock Exchange quoted bank shares—the Australia New Zealand Bank (ANZ) and the Adelaide Bank (ADB). The cointegration parameters are estimated on the first year's data, that is, from January 2, 2001 to January 1, 2002. Both price series are  $I(1)$  processes with a stationary cointegrated spread of the form:

$$P_{\text{ADB}}(t) + \beta P_{\text{ANZ}}(t) = \varepsilon_t, \quad (5.1)$$

where  $\varepsilon_t$  is an  $I(0)$  series. The estimate of  $\beta$  is  $-1/2.0237 = -0.4941$ . The model is then applied to the data from January 2, 2002 to August 30, 2002, which are 167 trade days. The outputs are presented by Tables 5.1 and 5.2.

Several patterns emerge from the tables. First, at least one valid trade is generated at all MPPT levels, except where the open condition becomes too low to allow potential trades to develop at the given investment levels. Predictably, the number of valid trades yielding a given MPPT increases with increased investment dollars. Second, a reduction in open trade boundary values increases the number of valid trades and then falls to zero trades as the spread becomes too small to generate trades within the given investment levels. This pattern reflects the functional relationship between the open condition level and the level of MPPT.

Third, the number of valid trades may appear low for all MPPT levels. But the restrictive nature of the second analytical assumption makes the results conservative. The restriction of valid trades to those situations where  $S_2$  is the shorted share will eliminate number of potential trades. So the actual number of trades in an unrestricted trading situation is probably higher than those reported here at all MPPT levels.

The purpose of the analysis did not include an examination of the effects of the MPPT procedure on the total profit levels of pairs trading. However, the total profit figures for the trading during the 167 days are included in Tables 5.1 and 5.2. At all MPPT levels the rate of return on investment increases as open condition boundaries are lowered, until they become too low to generate eligible trades at the MPPT level within the given investment levels. The pattern and level of increases in the rate of return on investment appear consistent across increasing levels of MPPT levels and invariant to that level.

## 6. Discussion

In this paper we derived a cointegration-based procedure that would always return at least a given minimum profit level. We then tested the feasibility of the procedure in terms of the number of possible trades that could be immunized at different MPPT levels for several combinations of open trade values, and investment dollars. The results of the empirical analysis suggest that the five-step strategy is feasible for commonly used parameter values.

Pairs trading strategies involve several decision choices. Taken together, these choices determine how much arbitrage profit potential is actually extracted from each pairs trade. Our cointegration-based analysis provides exploitable information on the long-run time

Table 5.1. Total profit and trades under varying MPPT for three levels of investment: ANZ and ADB share pairs.

$K$	$W$	Open condition (a)	Close condition (b)	Number of trades	Total profit $P$	$TP/W$
10	5000	$m + 1.5\sigma$	$m$	1	12.76	0.00255
		$m + \sigma$		1	14.08	0.00282
		$m + 0.75\sigma$		2	26.26	0.00525
		$m + (\sigma/2)$		2	37.26	0.00745
		$m + (\sigma/3)$		3	74.38	0.01487
50	5000	$m + 1.5\sigma$	$m$	1	63.83	0.01276
		$m + \sigma$		1	68.42	0.01368
		$m + 0.75\sigma$		2	131.30	0.02626
		$m + (\sigma/2)$		2	186.31	0.03726
		$m + (\sigma/3)$		0	0	0
50	10000	$m + 1.5\sigma$	$m$	1	63.83	0.00638
		$m + \sigma$		1	68.42	0.00684
		$m + 0.75\sigma$		2	131.30	0.01313
		$m + (\sigma/2)$		2	186.31	0.01863
		$m + (\sigma/3)$		3	368.26	0.03683
100	5000	$m + 1.5\sigma$	$m$	1	127.65	0.02553
		$m + \sigma$		1	135.84	0.02716
		$m + 0.75\sigma$		0	0	0
		$m + (\sigma/2)$		0	0	0
		$m + (\sigma/3)$		0	0	0
100	10000	$m + 1.5\sigma$	$m$	1	127.65	0.01277
		$m + \sigma$		1	135.84	0.01358
		$m + 0.75\sigma$		2	262.60	0.02626
		$m + (\sigma/2)$		2	372.61	0.03726
		$m + (\sigma/3)$		0	0	0
100	50000	$m + 1.5\sigma$	$m$	1	127.65	0.00255
		$m + \sigma$		1	135.84	0.00272
		$m + 0.75\sigma$		2	262.60	0.00525
		$m + (\sigma/2)$		2	372.61	0.00745
		$m + (\sigma/3)$		3	734.70	0.01469

## 12 Pairs trading based on cointegration approach

Table 5.2. Total profits and trades under varying MPPT with constant constrained investment: ANZ and ADB share pairs.

$K$	$W$	Open condition (a)	Close condition (b)	Number of trades	Total profit $P$	$TP/W$
500	100000	$m + 1.5\sigma$	$m$	1	636.84	0.00637
		$m + \sigma$		1	678.18	0.00678
		$m + 0.75\sigma$		2	1310.06	0.01310
		$m + (\sigma/2)$		2	1858.92	0.01859
		$m + (\sigma/3)$		3	3666.25	0.03666
		$m + (\sigma/4)$		4	5998.04	0.05998
		$m + (\sigma/5)$		0	0	0
1000	100000	$m + 1.5\sigma$	$m$	1	1273.67	0.01274
		$m + \sigma$		1	1355.36	0.01355
		$m + 0.75\sigma$		2	2620.11	0.02620
		$m + (\sigma/2)$		2	3717.84	0.03718
		$m + (\sigma/3)$		0	0	0
		$m + (\sigma/4)$		0	0	0
		$m + (\sigma/5)$		0	0	0
2000	100000	$m + 1.5\sigma$	$m$	1	2547.34	0.02547
		$m + \sigma$		1	2710.72	0.02710
		$m + 0.75\sigma$		0	0	0
		$m + (\sigma/2)$		0	0	0
		$m + (\sigma/3)$		0	0	0
		$m + (\sigma/4)$		0	0	0
		$m + (\sigma/5)$		0	0	0

series behavior of share pairs that is not available through currently used statistical methods. Unlike these current techniques, cointegration also offers a technique for systematically analyzing the interdependence of strategic choices. Our analysis shows that the profitability of a pairs trading strategy depends upon using weighting rules, minimum profit hurdles, and open/close criterion that reflect traders' preferences and are appropriate to the short and long-run price behavior of the component shares. Unrealistic values imply low trading rates, excessive trade durations, and low profits per share trade. Through cointegration the trader has a tool for investigating the statistical relationship between parameters.

Our analysis also emphasizes a range of other fundamental issues in statistical arbitrage strategies that require further study.

- (1) The contribution to arbitrage profit of each share depends upon relative price volatilities and mean-reversion characteristics of the component shares.
- (2) “Success” in pairs trading is a compromise between arbitrage levels and profit levels. Alternative weighting rules may optimize one objective but not both. For example, at the extreme, the most profitable strategy is to weight investment in the more volatile share at 100 percent and zero weight the other share, but this strategy offers minimal systematic risk protection.
- (3) Any trading strategy is a compromise between trading frequency, duration, and per trade profitability. Arbitrage profit levels depend on achieving a suitable mix relevant to the price series behavior of a given pair of financial assets.
- (4) For cointegrated share pairs, the latent profit potential relates directly to both the size and the frequency of short-term shocks characterizing each price series. Exploiting that potential depends on strategic choices.

Pairs trading, although limited to the simplest long/short case of two shares, is directly congruent with the much wider case of  $n$ -share long/short portfolios. Moreover, since there is no reason why pairs trading should not use put and call options rather than the underlying shares, our statistical analysis also translates to the derivatives portfolio context. It also reflects the statistical equivalent of the economic maxim that there are no “free lunches.”

## Acknowledgments

The paper has benefited from comments of Professor S. Gupta, Professor B. D. Sharma, and participants of the SCRA Conference, University of Maine, 2003, and Dr. W. Friesling, Director of Research, Commonwealth Bank, Australia, and participants at the 2003 IMMACS Workshop on Mathematical Finance, University of Wollongong. The authors are thankful to the anonymous referees for useful comments and remarks that led to a better exposition of the results.

## References

- [1] C. Alexander and A. Dimitriu, *The cointegration alpha: enhanced index tracking and long-short equity market neutral strategies*, 2002, Discussion Papers in Finance ISMA Center 2002-08, University of Reading.
- [2] C. Alexander, I. Giblin, and W. Weddington III, *Cointegration and asset allocation: a new active hedge fund strategy*, 2001, Discussion Papers in Finance ISMA Center 2001-03, University of Reading.
- [3] E. Gatev, W. Goetzmann, and G. Rouwenhorst, *Pairs trading: performance of a relative value arbitrage rule*, Working Paper 7032, National Bureau of Economic Research, Washington DC, 1999.
- [4] T. Gillespie and C. Ulph, *Pair trades methodology: a question of mean reversion*, Proceedings of International Conference on Statistics, Combinatorics and Related Areas and the 8th International Conference of Forum for Interdisciplinary Mathematics, NSW, December 2001, unpublished paper.
- [5] R. Harris, *Using Cointegration Analysis in Econometric Modelling*, Prentice Hall, London, 1995.

## 14 Pairs trading based on cointegration approach

- [6] D. Hendry and K. Juselius, *Explaining cointegration analysis: part II*, Energy Journal **22** (2001), no. 1, 75–120.
- [7] B. Jacobs and K. Levy, *Long/short equity investing*, Journal of Portfolio Management **20** (1993), no. 1, 52–63.
- [8] F.-S. L'Habitant, *Hedge Funds: Myths and Limits*, John Wiley & Sons, Chichester, 2002.
- [9] R. Lowenstein, *When Genius Failed: The Rise and Fall of Long-Term Capital Management*, Random House, New York, 2000.
- [10] M. Peskin and B. Boudreau, *Why hedge funds make sense*, 2000, <http://www.thehfa.org/articles/1.pdf>.
- [11] F. Reilly and K. Brown, *Investment Analysis and Portfolio Management*, 6th ed., Harcourt College, New York, 2000.
- [12] W. F. Sharpe, G. J. Alexander, and J. V. Bailey, *Investments*, 6th ed., Prentice-Hall, New Jersey, 1999.

Yan-Xia Lin: School of Mathematics and Applied Statistics, University of Wollongong,  
Northfields Avenue, Wollongong, NSW 2500, Australia  
*E-mail address:* yanxia@uow.edu.au

Michael McCrae: School of Finance and Accounting, University of Wollongong,  
Northfields Avenue, Wollongong, NSW 2500, Australia  
*E-mail address:* mccrae@uow.edu.au

Chandra Gulati: School of Mathematics and Applied Statistics, University of Wollongong,  
Northfields Avenue, Wollongong, NSW 2500, Australia  
*E-mail address:* cmg@uow.edu.au