Cointegrated Portfolio Management*

Zijiang YANG and Laiwan CHAN
Department of Computer Science and Engineering
CUHK

September 19, 2016

Technical Report : CUHK-CSE-2016-09
Cointegrated Portfolio Management

Abstract

This paper proposes a novel approach which uses cointegration for portfolio management. This approach is motivated by two of our findings. First, we observed that nonstationary stock data, comparing with the stationary ones, often result in larger discrepancy between the in-sample and out-of-sample statistics, including their return mean and standard deviation. Second, portfolios using traditional methods such as mean variance portfolio have better performance in the out-of-sample period if they are constructed based on stationary stock data than on the nonstationary ones due to the smaller discrepancy between the in-sample and out-of-sample data. Unfortunately, real stock data are mostly nonstationary and only few are stationary. Thus, in this paper, we propose a portfolio management approach in which we make use of the property of cointegration for addressing the nonstationary nature of the data. Our approach first detects pairs of cointegrated nonstationary stock data and then combines the pairs to form their stationary equilibrium series in a unique way. By replacing the original stock prices time series which are nonstationary with these stationary time series, we can construct portfolios using them as the base units. Various traditional portfolio management methods could be applied at this step. We present three versions of the portfolios using this approach as examples including equal weight, minimum variance and optimizing the trend and variance. Particularly, the one that maximizes the upward trend and minimizes the fluctuation of the portfolio around its trend shows the best performance. Compared with the corresponding portfolio management methods based on the original price series, the cointegrated approach achieves higher return and lower drawdown on real datasets. One key advantage of this portfolio is that the risk is low compared to other portfolios, even under adverse situations. Further, we use simulated data with one or two common factors to model the shared market effects and investigate their effects to the performance of our approach. We find that the portfolio has its best performance when there is one dominating market factor among the data.
1 Introduction

Portfolio management refers to the decision of allocating the investment in order to achieve certain goals. A well-known approach is the modern portfolio theory which is built upon Markowitz's [12] work. It gives the optimal rule for allocating the portfolio weights to the assets based on the mean and covariance of the stocks' returns. Subsequently, numerous works have devoted to the development of this approach. As the portfolio is constructed using the mean and covariance of the data, it is assumed that we have good estimation of these statistical information and they do not change until we have the updated estimation. When the estimation of these statistics deviate from the true values, the performance of the portfolio would be degraded. This is shown by a number of empirical results [13, 18, 19] and as discussed by many authors, the poor performance is mainly due to the estimation error of the stocks' returns in the out-of-sample period.

In an early work, Merton [20] showed that it is more difficult to estimate the mean than the covariance of the stocks' returns in the out-of-sample period. Jagannathan and Ma [18] recently reported that the estimation error in the stocks' returns is so large that nothing much is lost by ignoring the mean when constructing the portfolio. DeMiguel [26] conducted a series of empirical experiments on the real US equity data to investigate the performance of Markowitz's optimal portfolio in the out-of-sample period. He discovered that the stocks' returns have much larger influence onto the Markowitz's optimal portfolio than the covariance does. The key to the success of Markowitz's portfolio is to improve the estimation accuracy of the stocks' out-of-sample return. While there are numerous ways to estimate the statistical information, without the loss of generality, this paper simply uses the in-sample data for the estimation and the out-of-sample data for measuring the performance. The focus of this paper is not on how to make the estimations more accurate, but to construct a portfolio that has satisfactory performance without relying on an accurate estimation of the statistics.

As a preliminary study for our proposed approach described later, section 2 presents the investigation of the relationship between the stationarity property of the stock price time series and the accuracy of the estimation of the stock return in the out-of-sample period. We discover that the estimation of out-of-sample return is more accurate for stationary stocks than the nonstationary ones. Also, we compare the out-of-sample performance between portfolios constructed by stationary
and nonstationary stocks.

As indicated by the real data, very few stationary stocks exist in the market. The portion of the stationary stocks in the whole market is only around 5%. The small number of stationary stocks is too limited for us to construct a meaningful portfolio. To fully utilize the other nonstationary stocks in portfolio construction, we propose the cointegrated portfolio. It uses cointegration to create the stationary time series bases from the nonstationary stock price time series via the specific cointegration relationship between pairs of stocks. This process is vital as it transforms the nonstationary stocks into stationary stock bases. Thus, the stock bases are more accurate in terms of the estimation of the out-of-sample return. Then we build the portfolio using these stationary stock bases. The experiment results on the real equity data show that this portfolio obtains superior performance over the traditional ones.

The contribution of this paper is as follows:

(1) We show that the estimation of stationary stocks’ out-of-sample returns based on their in-sample returns is more accurate than the nonstationary stocks’ ones. The portfolio constructed by the stationary stocks performs better than the one constructed by the nonstationary stocks.

(2) We propose a novel portfolio management method to assign the weightings of assets via cointegration. Based on the experimental results using the real data, the proposed method is shown to be superior to the existing portfolio management methods in multiple stock markets. The superiority is justified by three measures: (a) Sharpe ratio, (b) final return, and (c) downside risk.

(3) We study the characteristics of the proposed cointegrated portfolio. We discover that the proposed portfolio’s performance improves with increasing the number of the stocks while the traditional portfolios do not. Also, we simulate the market condition and test the proposed portfolio performance under the simulated market. We discover that the proposed portfolio performs best when there is a dominating I(1) factor in the market.

The structure of this paper is arranged as follows: Section 2 investigates the effect of stationarity onto the return estimation and the portfolio performance. Section 3 establishes the theory and the method for managing the cointegrated portfolio. The experiment results run on the real data are also presented. Section 4 studies various characteristics of cointegrated portfolio. Section 5 draws
Table 1. Basic description of datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Starting Date</th>
<th>Ending Date</th>
<th>Time Length</th>
<th>No. of stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P500</td>
<td>2-Jan-01</td>
<td>30-Dec-14</td>
<td>3520</td>
<td>50</td>
</tr>
<tr>
<td>NYSE</td>
<td>3-Jul-62</td>
<td>31-Dec-84</td>
<td>5651</td>
<td>36</td>
</tr>
<tr>
<td>TSE</td>
<td>4-Jan-94</td>
<td>31-Dec-98</td>
<td>1258</td>
<td>88</td>
</tr>
<tr>
<td>HKSE</td>
<td>2-Jan-10</td>
<td>30-Dec-14</td>
<td>1293</td>
<td>48</td>
</tr>
</tbody>
</table>

Note: S&P500 stock dataset contains adjusted daily stock close price ranges from 2001/01/02 to 2014/12/30. This dataset contains 50 largest market capital companies of S&P500 component companies as of 2014/01/02. The NYSE and TSE stock datasets are obtained at http://www.cs.technion.ac.il/~rani/portfolios/ portfolios.htm. The HKSE stock dataset is obtained from the Yahoo finance. For the HKSE stock dataset, we obtain the daily stock prices for 48 component stocks of the Hang Seng Index(HSI) and the adjusted prices are taken as the daily stock price.

the conclusions and gives the future directions on this portfolio management method.

2 Stationary versus Nonstationary Data Series

In this section, we study the statistical discrepancy between the in-sample and the out-of-sample return return for stationary stock price time series and compare it with that for the nonstationary ones using real datasets. Then we conduct empirical studies for comparing the performance between portfolios constructed using stationary stocks and nonstationary stocks.

2.1 Estimation of the Out-of-sample Return

Traditionally, we use the in-sample statistics of the return for the estimation of the out-of-sample return. In the following context, we denote the price of i-th stock at time t as S_{i,t}. We use the log-return as the return computation, (i.e. r_{i,t} = \ln S_{i,t} - \ln S_{i,t-1}) and in vector form \( r_t = [r_{1,t}, r_{2,t}, \cdots, r_{N,t}]' \), where the prime symbol denotes vector transpose. We have a total of N stocks and let \( P \) be the set of the time period in concern. The sample mean return for i-th stock of the sample period is \( \mu_i = \frac{1}{|P|} \sum_{k \in P} r_{i,k} \) and the sample return sd is \( \sigma_i = \sqrt{\frac{1}{|P|-1} \sum_{k \in P} (r_{i,k} - \mu_i)^2} \), where \(|P|\) denotes the cardinality of set \( P \). The return vector for these N stocks is \( \mu = [\mu_1, \mu_2, \cdots, \mu_N]' \). The sample covariance of the return is computed as \( \Sigma = E(r - \mu)(r - \mu)' = \frac{1}{|P|-1} \sum_{k \in P} (r_k - \mu)(r_k - \mu)' \).

Our experiments rely on the "rolling window" approach and we label the segment of sample data concern as the period, \( p \), indexing at time \( T_p \). At any time \( T \), we use the previous one year
daily stock prices for training and the consecutive half a year prices for testing. To be more specific, at any sample period given by time $T$, daily prices of stock $i$ between $T - 251$ to $T$ are included in the in-sample (IS) data set and we refer the sample mean return as $r_{i,T}^{IS}$. The daily stock prices between $T + 1$ to $T + 126$ are included in the out-of-sample (OOS) data set and we refer this sample mean return as $r_{i,T}^{OOS}$. After we have computed the statistics for time $T$, we slide half a year forward and repeat the computation for $T + 126$. So, in the first period $T_1 = 252$ and in the second period $T_2 = 378$, so and so forth. Figure 1 illustrates this process.

**Figure 1.** Rolling window approach illustration

2.1.1 **Experimental Setup.** We first conduct the Augmented Dickey Fuller (ADF)\cite{7, 8, 9} tests for detecting stationarity property on the stocks price time series using the in-sample data. The confidence level is set to be 95%. Let set $S$ be the set that contains all stationary stocks and set $NS$ be the set that contains all nonstationary stocks. For each stock at each period $p$, we compute the absolute difference in the return mean ($r$) between the in-sample (IS) and out-of-sample (OOS) data. Similarly, we also compute the absolute difference in the return sd ($\sigma$) between the in-sample (IS) and out-of-sample (OOS) data. Then, we compute their average values among the stocks in the set $S$ at each period $p$ as belows.

\[
\text{meanDiff}^{(S)}_{T_p} = \frac{1}{|S|} \sum_{i \in S} |r_{i,T_p}^{IS} - r_{i,T_p}^{OOS}|
\]

\[
\text{SDDiff}^{(S)}_{T_p} = \frac{1}{|S|} \sum_{i \in S} |\sigma_{i,T_p}^{IS} - \sigma_{i,T_p}^{OOS}|
\]

We repeat the same for the set $NS$ and obtain the followings.

\[
\text{meanDiff}^{(NS)}_{T_p} = \frac{1}{|NS|} \sum_{i \in NS} |r_{i,T_p}^{IS} - r_{i,T_p}^{OOS}|
\]

\[
\text{SDDiff}^{(NS)}_{T_p} = \frac{1}{|NS|} \sum_{i \in NS} |\sigma_{i,T_p}^{IS} - \sigma_{i,T_p}^{OOS}|
\]
Table 2. No. of stationary stocks for the data sets (Median value of all periods)

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P500</th>
<th>NYSE</th>
<th>TSE</th>
<th>HKSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median of no. of stationary stocks</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>3.5</td>
</tr>
<tr>
<td>Total no. of stocks</td>
<td>50</td>
<td>36</td>
<td>88</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 3. The number of periods in which the discrepancy between in-sample and out-of-sample data for stationary and nonstationary stocks is smaller than the counterpart.

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P500</th>
<th>NYSE</th>
<th>TSE</th>
<th>HKSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Return Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>19</td>
<td>7</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>NS</td>
<td>19</td>
<td>41</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>Total Mean</td>
<td>26</td>
<td>43</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Return SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>17</td>
<td>9</td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td>NS</td>
<td>21</td>
<td>7</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Total SD</td>
<td>26</td>
<td>43</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: S stands for stationary, NS stands for nonstationary.

Finally, we compare the performance of stationary and nonstationary stocks by the followings:

(1) the number of periods in which the absolute difference between the mean return of the in-sample and out-of-sample data is smaller than the counterpart. For example, for stationary stocks, it is the number of periods that satisfies $meanDiff^{(S)}_{T_p} < meanDiff^{(NS)}_{T_p}$.

(2) the number of periods in which the absolute difference in the standard deviation (sd) of the in-sample and out-of-sample data is smaller than the counterpart. For example, for stationary stocks, it is the number of periods that satisfies $SDDiff^{(S)}_{T_p} < SDDiff^{(NS)}_{T_p}$

2.1.2 Experimental Result. We conduct the experiments on the datasets listed in Table 1. In each sample period, we conduct the ADF test for stationarity to each stock and Table 2 shows the median numbers of stationary stocks among all sample periods. Compared with the total number of stocks, only a couple of the stocks are stationary and the majority of them are nonstationary. Given this, Table 3 shows there are more periods that the stationary stocks has smaller difference in the return statistics between the in-sample and out-of-sample statistics than the nonstationary stocks (Table 3). For example, among 19 periods out of 26 periods in S&P500, the stationary stocks have lower estimation error for the out-of-sample return mean, whereas only 7 periods in which the nonstationary stocks have lower estimation error. The discrepancy on the mean return is more significant than the discrepancy on the standard derivation.
2.2 Performance of the Portfolios

We have shown that the statistical information of the out-of-sample returns for stationary stocks is closer to that of in-sample returns than nonstationary stocks. In this subsection, we first give a brief review on a few traditional methods in managing the portfolio [12, 26] which would be used and compared in Section 3. Then, we apply the mean variance portfolio on the stationary and nonstationary stocks in the data sets separately for comparison. Their difference is used to determine the effects of the stationarity of the stock prices onto the portfolio performance.

2.2.1 Existing Portfolio Management Methods.

2.2.2 Equally Weighted Portfolio. Equally Weighted Portfolio (a.k.a. "1/N portfolio") involves holding a portfolio with equal weight 1/N for each asset. Mathematically, the weight vector can be expressed as \( \mathbf{w}_t = \frac{1}{N} \mathbf{1} \), where \( \mathbf{1} \) represents the \( N \times 1 \) vector with all entries equal to 1. The weight vector is independent of the data and their statistics. It is the simplest and most naive strategy in portfolio management.

2.2.3 Minimum Variance Portfolio. The Minimum Variance Portfolio aims at minimizing the variance of the portfolio. The weight vector \( \mathbf{w}_t \) is the one that the resulting portfolio has the smallest variance as follows.

\[
\min_{\mathbf{w}_t} \mathbf{w}_t' \Sigma_t \mathbf{w}_t, \quad s.t. \quad \mathbf{w}_t' \mathbf{1} = 1
\]

This management method requires estimating the covariance matrix \( \Sigma_t \) of the return of each asset only whereas the expected return is not considered.

2.2.4 Mean Variance Portfolio. The Mean Variance Portfolio (Markowitz's Portfolio[12]) considers both the return and variance of the portfolio. The weight \( \mathbf{w}_t \) is the one that the resulting portfolio maximizes the following quantity which is a trade-off between the return and variance of the portfolio as follows.

\[
\max_{\mathbf{w}_t} \mathbf{w}_t' \mu_t - \kappa \mathbf{w}_t' \Sigma_t \mathbf{w}_t, \quad s.t. \quad \mathbf{w}_t' \mathbf{1} = 1
\]
Figure 2. Illustration of the portfolio formation process for mean variance portfolio (a) and trend variance cointegrated portfolio (b)

The coefficient $\kappa (\kappa > 0)$ specifies the risk aversion coefficient of the variance of the return. This management method needs to estimate both the mean return of each asset and the covariance matrix of the return of the assets. Figure 2(a) illustrates the process for the formation of mean variance portfolio. The analytical solution for mean variance portfolio can be obtained via Lagrangian multiplier method.

2.3 Experimental Settings

Again, we use the rolling window approach. We use the data from $T - 251$ to $T$ (inclusive), a total of 252 trading days daily stock prices, for computing the weight $w_t$. This weight is used for building the portfolio at time $T$ and it is kept unchanged for the testing period from $T+1$ to $T+126$ days. Transactions take place only at the beginning and end of the testing periods, and no buy and sell activities would take place in between. Thus, it is reasonable to ignore the transaction cost which is just a fixed percentage of the final portfolio return. Similar to the previous experiments, $T$ advances by 126 days for each period.
2.4 Evaluating the Performance

2.4.1 Evaluation Metrics. To evaluate the out-of-sample performance of the portfolios, we use three quantities for comparison. The first one is the out-of-sample Sharpe ratio (SR) of each strategy. It is defined as the sample mean of excess return divided by the sample standard deviation of the return in each out-of-sample period. To simplify the computation, we assume that the risk-free rate \( r_f \) is 0.

\[
SR = \frac{\mu_{\text{portfolio}} - r_f}{\sigma_{\text{portfolio}}}
\]

Secondly, we evaluate the profitability (FinalReturn) of each strategy using the percentage gain of the strategy in the out-of-sample period. \( V_{\text{start}} \) and \( V_{\text{end}} \) are the values of the portfolio at the start and end of the out-of-sample period respectively.

\[
\text{FinalReturn} = \frac{V_{\text{end}} - V_{\text{start}}}{V_{\text{start}}}
\]

Lastly, we measure the downside risk of each strategy by computing the maximum drawdown (maxDD), which is the maximum percentage of the wealth that one can lose during the out-of-sample period. \( s \) and \( t \) are any time within the out-of-sample period. \( V_s \) and \( V_t \) denote the value of the portfolio at time \( s \) and \( t \) respectively.

\[
\text{maxDD} = \max \frac{V_s - V_t}{V_s} \quad \text{for any } s < t, \text{ and } V_s > V_t.
\]

In each period, we conduct the ADF test for stationarity to categorize the stocks into stationary ones and nonstationary ones based on the in-sample data. We discard those periods that the number of stationary stocks is less than 3 as the portfolios constructed using few stocks are not very representative. So the numbers of valid periods are smaller than those shown in Table 3. Two mean variance portfolios with \( \kappa = 0.025 \) are then constructed; one is based on the stationary stocks and the other one on nonstationary stocks. As reviewed in Table 2, only around 5% of the stocks in each period of the datasets are stationary. For fair comparison, we take a random sub-sample of the nonstationary stocks from the nonstationary stock pool in each period such that the number of nonstationary stocks used for portfolio construction is the same as that of the stationary stocks. This process is repeated 100 times. We take the average of these 100 out-of-sample nonstationary stock portfolios as the final performance for the nonstationary stock portfolio.
Similar to the methodology for the return estimation, we use the followings to compare the performance of the portfolio constructed by the stationary and nonstationary stocks. The superscripts of $S$ and $NS$ are appended to the three evaluation quantities in order to distinguish for the type of portfolios.

(1) The number of periods in which the portfolio yields higher final return in the out-of-sample data than the counter portfolio. For example, for the portfolio based on stationary stocks, it is the number of periods that satisfies $FinalReturn_{T_T}^{(S)} > FinalReturn_{T_T}^{(NS)}$.

(2) The number of periods in which the portfolio yields lower standard deviation in portfolio return among the out-of-sample data than the counter portfolio. For example, for the portfolio based on stationary stocks, it is the number of periods that satisfies $returnSD_{T_T}^{(S)} < returnSD_{T_T}^{(NS)}$.

(3) The number of periods in which the portfolio Sharpe ratio in the out-of-sample data is higher than the counterpart. For example, for the portfolio based on stationary stocks, it is the number of periods that satisfies $SR_{T_T}^{(S)} > SR_{T_T}^{(NS)}$.

(4) The number of periods in which the portfolio yields lower maximum drawdown in the out-of-sample data than the counterpart. For example, for the portfolio based on stationary stocks, it is the number of periods that satisfies $maxDD_{T_T}^{(S)} < maxDD_{T_T}^{(NS)}$.

2.4.2 Experimental Result. Again, we conduct experiments on the datasets listed in Table 1, and the result is shown in Table 4. As all but three entries shown under the S columns are higher than the counterparts under the NS columns, this indicates that the portfolio constructed by the stationary stocks generally performs better than that constructed by the nonstationary ones. In particular, the Sharpe ratios are higher for the portfolios with stationary stocks for all four datasets.

3 Cointegrated Portfolio Management

We conclude from the previous section that the stationarity of the stocks would affect the portfolio’s performance. It is therefore preferable to construct the portfolio using stationary stocks. However, due to the limited number of the stationary stocks, it is difficult for us to diversify the risk just across very few stocks. Besides, we should not exclude the potential profits in the investment of the nonstationary stocks. In this paper, we propose a new portfolio management system which
Table 4. The number of periods that the out-of-sample portfolio based on either stationary or nonstationary stocks performs better than the counterpart.

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P500</th>
<th>NYSE</th>
<th>TSE</th>
<th>HKSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>NS</td>
<td>S</td>
<td>NS</td>
</tr>
<tr>
<td>Final Return</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>return SD</td>
<td>6</td>
<td>3</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>SR</td>
<td>6</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>maxDD</td>
<td>6</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>17</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: S stands for stationary, NS stands for nonstationary.

uses cointegration to alleviate the problem caused by the nonstationary stock price time series. When two nonstationary data series are said to be cointegrated, it is possible to combine them in a specify way such that the resultant data series is stationary. We use this property to transform nonstationary stock price time series into sets of stationary time series and then the portfolio is constructed based on them. Following this idea, we propose a framework of cointegrated portfolio and multiple portfolios are tested and compared. The following context gives a brief overview of cointegration. Then we show how to construct the stationary time series from the nonstationary ones using cointegration. Finally, we propose the framework of cointegrated portfolio.

3.1 Definition of Cointegration

Cointegration was introduced by the Nobel laureates Engle and Granger in the 1980s [11, 21, 22]. Since then, cointegration has widely been used in the analysis of macroeconomic time series. The followings present the definition of integration and cointegration.

Definition (Integration): Let $X_t$ be a $n$-dimensional vector time series. If $X_t$ achieves stationary after differencing $d$ times, then $X_t$ is said to be integrated of order $d$, denoted as $X_t \sim I(d)$.

In real applications, the parameter $d$ typically takes the value 0 or 1. If $X \sim I(0)$, $X_t$ itself is stationary. If $X \sim I(1)$, the first difference $X_{t+1} - X_t$ is stationary. A typical example for $I(1)$ is random walk: $X_{t+1} - X_t = e_{t+1}$, where $e_{t+1}$ is a vector random innovation. It is generally true that the linear combination of two vector $I(1)$ time series will also be $I(1)$. However, there is a possibility that this linear combination will be $I(0)$. For example, if the first two components of
$X_t$ are following the following process: $X_{1,t+1} - X_{1,t} = u_{t+1}$ and $X_{2,t+1} - X_{2,t} = u_{t+1}$. The linear combination $Z_t = X_{1,t} - X_{2,t}$ is a constant and so $Z_t$ is $I(0)$. When this occurs, it means that an equilibrium exists between $X_{1,t}$ and $X_{2,t}$. Though $X_{1,t}$ and $X_{2,t}$ themselves are unpredictable, their differences are perfectly predictable. [11, 21] formalized this idea as cointegration.

**Definition (Cointegration):** If $X_t \sim I(d)$ and there exists a vector $\beta (\beta \neq 0)$ so that $Z_t = \beta'X_t \sim I(d - b)$ for $b > 0$, then the components of $X_t$ are said to be cointegrated of order $d$, $b$. The vector $\beta$ is called the cointegrating vector.

It can be easily seen that cointegration characterizes the equilibrium relationship among the components of $X_t$ by a simple linear combination $\beta'X_t$. Granger’s representation theorem states that this cointegration relationship can be expressed in an Error Correction Model.

$$\Delta X_t = \mu + \Phi D_t + \alpha \beta'X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \epsilon_t \quad (1)$$

Where $\Delta X_t \equiv X_t - X_{t-1}$, $\mu (n \times 1)$ is the constant term, $\Phi D_t$ is the deterministic trend term. $\Gamma_i$’s are the $n \times n$ coefficient matrices and $\epsilon_t(n \times 1)$ is the random innovation. $\beta$ is an $n \times r$ matrix whose column vectors are cointegrating vectors. The number of the columns $r$ means that there are $r$ cointegration relationships among the time series. $\alpha$ is an $n \times r$ mixing matrix that mix all these $r$ stationary time series $\beta'X_{t-1}$ together.

To estimate the cointegration relationship, Johansen’s trace or eigenvalue test [23, 24, 25] based on ECM is usually performed to test the existence of cointegration and to estimate the deterministic trend $\Phi D_t$ and the cointegrating vector $\beta$.

### 3.2 Constructing the Pairs

Since cointegration aims at capturing the stationary equilibrium among the time series, we can exploit this property among the stocks to help us to construct the portfolio. As there may exist complicated multivariate cointegration relationships among the time series, we only consider the bivariate cointegration relationship for simplicity. We first perform the bivariate cointegration test on any stock pair $i$ and $j$, using the log price time series $\{\ln S_{i,t}, \ln S_{j,t}\}$ to detect for cointegration and extract the stationary equilibrium if it exists. Mathematically, we extract the cointegrated

13
pairs that exhibit the following relationship.

\[
\ln S_{i,t} - \beta_{i,j} \ln S_{j,t} = c_{i,j} t + I(0)_{i,j}
\]  

(2)

Under cointegration, the pair of \(S_{i,t}\) and \(S_{j,t}\) forms a bivariate relationship \(\ln S_{i,t} - \beta_{i,j} \ln S_{j,t}\) where \(\beta_{i,j}\) is the cointegrating coefficient. The pair is composed of a linear trend component \(c_{i,j} t\) where \(c_{i,j}\) is the trend coefficient, and a stationary series \(I(0)_{i,j}\). We assume that the time series data do not deviate significantly from the starting position, i.e., \(\left| \frac{S_{i,t} - S_{i,0}}{S_{i,0}} \right| \leq 1\) and \(\left| \frac{S_{j,t} - S_{j,0}}{S_{j,0}} \right| \leq 1\).

Using the Taylor series for approximation, we have

\[
\ln S_{i,t} - \beta_{i,j} \ln S_{j,t} \approx \ln S_{i,0} + \frac{S_{i,t} - S_{i,0}}{S_{i,0}} - \beta_{i,j} (\ln S_{j,0} + \frac{S_{j,t} - S_{j,0}}{S_{j,0}}) \\
= \frac{1}{S_{i,0}} S_{i,t} + \frac{\beta_{i,j}}{S_{j,0}} S_{j,t} + (\ln S_{i,0} - \beta_{i,j} \ln S_{j,0}) - 1 + \beta_{i,j}
\]

If stocks \(i\) and \(j\) are cointegrated, the left hand side of the equation reduces to the trend component plus the stationary component. On the right hand side of the equation, only the first two terms are time varying and the rest are not. The first two terms are indeed equivalent to the value of a mini-portfolio composing of \(1/S_{i,0}\) shares of stock \(i\) and \(\beta_{i,j}/S_{j,0}\) shares of stock \(j\). Without loss of generality, we merge the remaining non-time varying terms with the stationary component \(I(0)_{i,j}\) on the right hand side. Thus, the mini-portfolio results as the stationary equilibrium of the cointegration process.

Among the pool of \(N\) stocks, suppose that there are \(K\) cointegrated pairs from all \(N(N-1)/2\) possible pairs from the original log stock price \(\{\ln S_{i,t}\}_{i=1}^{N}\). We define \(P_{k,t}\) as the value of the \(k\)th mini-portfolio at time \(t\) which composes of the cointegrated pair of stocks \(k_i\) and \(k_j\).

\[
P_{k,t} = \left( \frac{1}{1 + \beta_{k_i,k_j}} \right) \left( \frac{1}{S_{k_i,t}} S_{k_i,t} + \frac{\beta_{k_i,k_j}}{S_{k_j,t}} S_{k_j,t} \right)
\]  

(3)

The fraction \(\frac{1}{1 + \beta_{k_i,k_j}}\) is used to standardize each pair such that the value of each pair at time \(t = 0\) is 1. Thus, each pair will hold \(\frac{1}{1 + \beta_{k_i,k_j} S_{k_i,0}}\) shares of stock \(k_i\) and \(\frac{\beta_{k_i,k_j}}{(1 + \beta_{k_i,k_j}) S_{k_j,0}}\) shares of stock \(k_j\).

As we have explained that the \(k\)th mini-portfolio reduces to the stationary equilibrium, we have
\[ P_{k,t} = \hat{c}_k t + I(0)_k \]

where \( P_{k,0} = 1, \hat{c}_k = c_{k_i,k_j}/(1 + \beta_{k_i,k_j}) \) and \( I(0)_k \) is a stationary time series. For each \( P_{k,t} \), the trend component \( \hat{c}_k t \) specifies the trend of the pair and the stationary component \( I(0)_k \) specifies how the value of the pair varies around the trend.

### 3.3 Constructing the Portfolio Based on the Pairs

We have shown the approach that we construct mini-portfolios from pairs of cointegrated stocks and they result in stationary time series, as denoted by \( P_{k,t} \) for \( (k = 1, 2, \ldots K) \). Our cointegrated portfolio framework is to use these \( P_{k,t} \), replacing the original \( N \) stock data series, as the bases for portfolio construction. Under this framework, we propose the following schemes.

#### 3.3.1 Equally Weighted Cointegrated Portfolio (EWCP)

In analog to the Equally Weighted Portfolio [12, 26], we assign equal weight to each of the base units which arose from the standardized pairs. Those standardized pairs have negative long-run trend \( (\hat{c}_k < 0) \) are excluded as they deteriorate the performance of the final portfolio. By screening out the standardized pairs with negative long-run trend, we have \( \hat{K} \) base pairs where \( \hat{K} \leq K \). We assign equal weight \( \hat{w}_t = \frac{1}{\hat{K}} \) to these non-negative trend base unit \( P_{k,t} \). The expected return of EWCP is \( \frac{1}{\hat{K}} \sum_{k=1}^{\hat{K}} \hat{c}_k t \), which increases linearly with time. This method does not require the estimation of the trend (except the sign) and variance of the standardized pairs \( \{(k_i, k_j)\}_{k=1}^{\hat{K}} \).

#### 3.3.2 Minimum Variance Cointegrated Portfolio (MVCP)

Similar to the minimum variance portfolio management method mentioned above, we minimize the variance of the final portfolio, which corresponds to the stationary component of the final portfolio. Similar to the treatment in Equally Weighted Cointegrated Portfolio (EWCP), we screen out the base pairs with negative trend. The weight \( \hat{w}_t \) is obtained by solving the optimization problem for \( \{P_{k,t}\}_{k=1}^{\hat{K}} \)

\[
\min_{\hat{w}_t} \hat{w}_t' \hat{\Sigma}_t \hat{w}_t, \quad \text{s.t.} \quad \hat{w}_t' \mathbf{1} = 1
\]

where the variance \( \hat{\Sigma}_t \) is the sample covariance of the stationary component \( [I(0)_1, I(0)_2, \ldots, I(0)_{\hat{K}}]' \)

This portfolio management method requires the estimation of the sample covariance of the stationary component \( [I(0)_1, I(0)_2, \ldots, I(0)_{\hat{K}}]' \). The analytic solution for \( \hat{w}_t \) can be obtained by Lagrange multiplier method.
3.3.3 Trend Variance Cointegrated Portfolio (TVCP). Using the same idea of return-risk trade-off as in the mean variance portfolio management, we seek for a balance between maximizing the trend of the final portfolio and minimizing the fluctuation of the final portfolio around the trend.

\[
\max_{\mathbf{\hat{w}}_t} \mathbf{\hat{w}}'_t \hat{c} - \tau \mathbf{\hat{w}}'_t \hat{\Sigma}_t \mathbf{\hat{w}}_t, \quad s.t. \quad \mathbf{\hat{w}}'_t \mathbf{1} = 1
\]

where the trend vector \( \hat{c} \) is \( [\hat{c}_1, \hat{c}_2, \cdots, \hat{c}_K]' \) and the variance \( \hat{\Sigma}_t \) is the sample covariance of the stationary component \( [I(0)_1, I(0)_2, \cdots, I(0)_K]' \). The coefficient \( \tau (\tau > 0) \) measures the risk-aversion level of the fluctuation of the final portfolio around the trend. Figure 2(b) illustrates the process for the portfolio formation.
Table 5. Performance of the strategies on real stock data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Final SR Ret. (%)</td>
<td>Drawdown (%)</td>
<td>Final SR Ret. (%)</td>
<td>Drawdown (%)</td>
</tr>
<tr>
<td>1/N</td>
<td>0.57</td>
<td>229.58</td>
<td>51.60</td>
<td>1.09</td>
</tr>
<tr>
<td>Min variance</td>
<td>0.87</td>
<td>290.44</td>
<td>40.98</td>
<td>1.10</td>
</tr>
<tr>
<td>Mean variance(κ = 0.0025)</td>
<td>0.66</td>
<td>698.50</td>
<td>64.30</td>
<td>0.74</td>
</tr>
<tr>
<td>Mean variance(κ = 0.005)</td>
<td>0.66</td>
<td>583.20</td>
<td>62.65</td>
<td>0.78</td>
</tr>
<tr>
<td>Mean variance(κ = 0.01)</td>
<td>0.67</td>
<td>503.30</td>
<td>59.34</td>
<td>0.83</td>
</tr>
<tr>
<td>EWCP</td>
<td>0.58</td>
<td>298.14</td>
<td>37.31</td>
<td>0.66</td>
</tr>
<tr>
<td>MVCP</td>
<td>0.74</td>
<td>365.36</td>
<td>59.45</td>
<td>0.81</td>
</tr>
<tr>
<td>TVCP(τ = 5)</td>
<td>0.88</td>
<td>736.70</td>
<td>48.16</td>
<td>1.00</td>
</tr>
<tr>
<td>TVCP(τ = 10)</td>
<td>0.91</td>
<td>724.20</td>
<td>38.00</td>
<td>0.92</td>
</tr>
<tr>
<td>TVCP(τ = 20)</td>
<td>1.02</td>
<td>700.70</td>
<td>33.81</td>
<td>0.80</td>
</tr>
</tbody>
</table>
3.3.4 Interpreting the Performance.

3.3.4.1 Comparison of the performance among the cointegrated portfolio management methods. We compare the performance among the cointegrated portfolio management methods (EWCP, MVCP and TVCP) as shown in Table 5. It is found that among these three methods, TVCP has the best performance across all data sets. Both EWCP and MVCP are not as superior as TVCP probably because they do not optimize the upward trends in the portfolio. This also suggests that the trends obtained by the cointegration equilibrium states provide useful information in gaining profits in the portfolios. Therefore, the profits of EWCP and MVCP are not comparable to that earned by TVCP. For TVCP, as \( \tau \) increases, we have less risky performance which is indicated by the decrease in the max drawdown.

3.3.4.2 Comparison of the performance between the cointegrated portfolio management methods and the traditional methods. We then compare the performance of strategies between the cointegrated portfolio management methods and the corresponding traditional methods on the most recent dataset S&P500 which covers the period from 2001 to 2014. From the three measurements, the performance of the cointegration approaches are generally better than the traditional ones. It is observed that the TVCP strategy has the best out-of-sample performance in terms of maximizing the final return and minimizing the maximum drawdown.

The final returns for the TVCP are 736\%, 724\% and 700\% for \( \tau = 5, 10 \) and 20 respectively. They are much larger than the final return of others.

An encouraging observation of this TVCP strategy is its small downside risk. Let's compare the method mean variance(\( \kappa = 0.005 \)) and the method TVCP(\( \tau = 10 \)). These two methods generate the final return, 583\% and 724\% for mean variance and TVCP respectively. TVCP has generated more than 24\% (724\%/583\% − 1 = 0.24) of the return of the mean variance approach. However, the maximum drawdown of TVCP(\( \tau = 10 \)) strategy is only about 38\% and the maximum drawdown of mean variance(\( \kappa = 0.005 \)) is about 63\%. It is observed that the maximum drawdown of TVCP is about 40\% less than that of the mean variance approach. Notice that this dataset includes the period of the 2008 financial crisis, where the S&P500 index dropped by more than 65\% from 2007 to
Figure 3. Hang Seng Index from Jan 04, 2010 to Dec 31, 2014

2009. This approach appears to be very promising in controlling the downside risk of the portfolio.

Similar results are also found on the datasets of TSE. TVCP achieves a high Sharpe ratio, high final return and low maximum drawdown. It is found that the risk-aversion coefficient has an impact on controlling the final return and the drawdown. The higher the risk-aversion level, the lower the drawdown and the lower the return. This reflects the parity between the return and the risk. The balance of return generation and risk management is subject to the preference of the practitioners.

HKSE is found different to the above. The final return surprisingly increases with $\tau$ while the drawdown decreases as expected. This is regarded as abnormal as an increasing $\tau$ would sacrifice more of the return for a less risky performance. The maximum drawdowns are still much lower than those of mean variance portfolios which demonstrates the superiority of the cointegration portfolios. The final returns are however lower than the mean variance ones but it shows an increasing trend with $\tau$. By carefully analyzing the data, we find that there is one period which the equilibrium pairs indicate promising trends but they turn out to be the other way round. During the period from Jan 03, 2011 to Dec 30, 2011 (shown as the highlighted area in Figure 3), the Hang Seng Index drops by 21.34% from 23436.05 to 18434.39. It then recovers quickly to the level of 21000
Table 6. Performance of the strategies on HKSE and HKSE(reduced) datasets

<table>
<thead>
<tr>
<th></th>
<th>HKSE (Jan 2011 - Dec 2014)</th>
<th>HKSE(reduced) (Jan 2011 - Dec 2011, Jul 2012 - Dec 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR</td>
<td>Final Return (%)</td>
</tr>
<tr>
<td>1/N</td>
<td>0.39</td>
<td>24.60</td>
</tr>
<tr>
<td>minVar</td>
<td>1.04</td>
<td>56.96</td>
</tr>
<tr>
<td>meanVar(κ = 0.0025)</td>
<td>0.78</td>
<td>135.56</td>
</tr>
<tr>
<td>meanVar(κ = 0.005)</td>
<td>0.76</td>
<td>116.56</td>
</tr>
<tr>
<td>meanVar(κ = 0.01)</td>
<td>0.75</td>
<td>98.43</td>
</tr>
<tr>
<td>EWCP</td>
<td>0.32</td>
<td>21.15</td>
</tr>
<tr>
<td>MVC</td>
<td>0.47</td>
<td>28.69</td>
</tr>
<tr>
<td>TVCP(τ = 5)</td>
<td>0.48</td>
<td>39.31</td>
</tr>
<tr>
<td>TVCP(τ = 10)</td>
<td>0.74</td>
<td>59.90</td>
</tr>
<tr>
<td>TVCP(τ = 20)</td>
<td>0.95</td>
<td>75.73</td>
</tr>
</tbody>
</table>

in Feb 2012. Figure 3 shows the Hang Seng Index and the highlighted areas are the training and testing time frames of this special period. We would observe that all three quantities improve with a larger τ, which is against the trade-off between the return and the risk. We conduct another experiment on the same HKSE dataset but excluding this special period and compare it with the original HKSE dataset. The results are shown in the Table 6. We can see clearly that the TVCP returns have been improved by excluding this period and more importantly, the trend of the return and risk reflects the trade-off between them. Actually, we find that by excluding this period, all three performance measures have been improved. The Sharpe ratio and maximum drawdown are shown superior than the mean variance portfolios. In contrast, the final return of the mean variance portfolios are reduced by omitting this special period.

As for the NYSE dataset, the cointegration approach gains improvement in the downside risk but not on the final returns. Table 3 shows that for NYSE, it has 44% (19/43) of chance that the nonstationary stocks have less discrepancy than the stationary ones in terms of mean return. However, in portfolio performance, Table 4 shows that a much higher ratio, 59% (10/17) is found in the final return that the portfolio based on nonstationary gives a better performance. This indicates that the statistical inconsistency of the nonstationary stocks tends to lead towards a favorable out-of-sample portfolio performance in the final returns. Another reason to explain the portfolio performance on NYSE is the number of cointegrated pairs. Table 7 shows the number of cointegrated pairs found on each data set. In terms of percentage, all datasets have about 2% of the
Table 7. Median number of cointegrated pairs found on each data set

<table>
<thead>
<tr>
<th>Data set</th>
<th>S&amp;P500</th>
<th>NYSE</th>
<th>TSE</th>
<th>HKSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Stocks</td>
<td>50</td>
<td>36</td>
<td>88</td>
<td>48</td>
</tr>
<tr>
<td>Median no. of pairs</td>
<td>23.5</td>
<td>18</td>
<td>80</td>
<td>23.5</td>
</tr>
</tbody>
</table>

pairs that show cointegration. Since NYSE has the smallest number of stocks among all datasets, it is also the one showing the least number of pairs among all. A careful check on the weight assigned on the stocks shows that this portfolio concentrates only on a few stocks when there are insufficient pairs. Therefore, in such case, the risk of the portfolio cannot be well-diversified. A more thorough examination on the number of stocks and pairs are shown in section 4.2.

4 Characteristics of the Cointegrated Portfolio

While we have compared the performance of the portfolios in the previous section, we intend to give a more in-depth analysis on the characteristics of the cointegrated portfolio. Three studies are conducted in this section. We were motivated by the fact that there are large discrepancy between in-sample and out-of-sample statistics of the raw stock prices. It is natural for us to examine the statistical discrepancy of the new bases and how it affects the return of the portfolios. Then we analyze the effects of the number of stocks on the performance of cointegrated portfolio. Finally, the last subsection uses the factor analysis approach to simulate the existence of common factors under different market conditions, and the performance of the portfolio is also presented.

4.1 Effects of the New Bases

In our work, portfolios are constructed using the in-sample data and they are tested and compared using the out-of-sample data. In an earlier section, we have illustrated stationary time series results in less discrepancy between the in-sample and out-of-sample data, and the impact of stationarity on the performance of the mean variance method. By replacing the highly nonstationary raw data used in the traditional mean variance method with the new base units obtained by cointegration, our approach is expected to construct portfolios based on these data series. This leads to two issues that we want to explore. The first one is that if these new bases have less discrepancy between
the in-sample and out-of-sample data than the raw data. The second question is that if the return of the portfolio in the out-of-sample period is predictable from the in-sample period. This is an important question in portfolio management as we often need to anticipate the return and risk.

To tackle the first issue about the new bases, we have to study the cointegrated pairs. Pairs are detected using the in-sample data and it is questionable if the property of cointegrated would still persist during the out-of-sample period. Besides, any pair that is detected in one time period does not guarantee that it would be detected in other time period. The numbers of pairs detected vary from one period to another one. Another characteristics of the new bases is that their ranges (i.e. the values of $P_{k,i}$) vary substantially from one pair to another one. This is due to the fact that they are obtained from the weighted differences of pairs of stocks, and it is not appropriate to measure their absolute difference. The method we used in the section 2.1.1 is thus not appropriate. We then propose to use the follow regression approach for our analysis. The "rolling window" approach is adopted again; at period indexing at time $T$, we use the past 252 trading day information (day $T-251$ to day $T$) to compute two in-sample quantities (IS); the net return for each raw stock data $i$, $r_{i,T}^{IS}$, and the standard deviation of the return (sd) $\sigma_{r,T}^{IS}$. Similarly we find the out-of-sample (OOS) return and sd values; $r_{i,T}^{OOS}$, and $\sigma_{r,T}^{OOS}$ for the corresponding testing period. For each period, linear regressions are performed on the stock data to obtain two linear relations. One is the linear relationship between the out-of-sample stock return $r_{T}^{OOS}$ and the in-sample stock return $r_{T}^{IS}$. The other one is the linear relationship between the out-of-sample stock return standard deviation $\sigma_{T}^{OOS}$ and the in-sample stock return sd $\sigma_{T}^{IS}$. The following two equations capture the relationships.

\[ r_{T}^{OOS} = c_{r,T} + \beta_{r,T} r_{T}^{IS} + \epsilon_{T} \quad (4) \]

\[ \sigma_{T}^{OOS} = c_{\sigma,T} + \beta_{\sigma,T} \sigma_{T}^{IS} + \xi_{T} \quad (5) \]

If the in-sample data has the same distribution as the out-of-sample data, we expect that both $\beta_{r,T}$ and $\beta_{\sigma,T}$ equal to one, and both $c_{r,T}$ and $c_{\sigma,T}$ equal to zero.

These regressions are repeated for each valid $T$ in the entire period of the datasets. The raw stock prices (i.e. $S_{i,T}$) are used to obtain the regression coefficients for the traditional mean variance portfolio. Similarly, the same methodology is applied to analyze the equilibrium states of the detected cointegrated pairs. For the cointegrated pairs, we use $P_{k,T}$ instead of $S_{i,T}$. The
dimensionality of the vectors changes from $N$ for raw stock prices to $K$ for the cointegrated pairs.

Regression for the returns

We first compare the agreement on the in-sample returns and out-of-sample returns using regression as shown in Equation 4. We use the following measures for evaluation and the results are tabulated in Table 8.

(1) The number of periods that have positive $\beta_{r,T}$. We observed that the returns for in-sample data and out-of-sample data vary substantially in most of the periods. We use a lenient measurement that detects the sign only. Positive $\beta_{r,T}$ indicates that the in-sample and out-of-sample returns are at least of the same sign. This corresponds to the consistency in having profit or loss.

(2) The number of periods that have a higher $R^2$-squared value ($R^2$) of the linear regression than the counter approach. $R^2$ gives the indication of the percentage that the variability between two variables (in-sample and out-of-sample) have been accounted for.

Regression for the standard deviations

Similarly, we compare the prediction performance of the out-of-sample return sd using linear regression as shown in Equation 5. The following measurements are used and the results are tabulated in Table 8.

(3) The number of periods where the coefficient $\beta_{\sigma,T}$ falls into the region $1 - 0.2 < \beta_{\sigma,T} < 1 + 0.2$. In other words, we count the number of periods where the in-sample standard deviation of the return can be an accurate proxy for the out-of-sample return sd.

(4) The number of periods that has a higher $R^2$ than the counter approach, i.e. for the mean variance method, the raw dataset has a higher $R^2$ than the new bases obtained by the cointegrated pairs.

Finally we take into consideration of the combined effects on both return and standard deviation. The Sharpe ratios of each stock during the in-sample and out-of-sample periods are measured. The average Sharpe ratios of all stocks for the two periods are then compared separately. Similarly, for the new base obtained by the cointegrated pairs during the in-sample and out-of-sample periods, we also compute the average Sharpe ratios of all bases. The performance is measured by the following:
Table 8. The number of periods that the bases of the portfolio using raw stock prices or cointegration pairs provide good prediction performance.

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P500</th>
<th>NYSE</th>
<th>TSE</th>
<th>HKSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>raw price</td>
<td>pair</td>
<td>raw price</td>
<td>pair</td>
</tr>
<tr>
<td>$\beta_{t,T} &gt; 0$</td>
<td>13</td>
<td>17</td>
<td>28</td>
<td>23</td>
</tr>
<tr>
<td>Higher return $R^2$</td>
<td>11</td>
<td>15</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>$|\beta_{t,T} - 1| &lt; 0.2$</td>
<td>11</td>
<td>9</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>Higher SD $R^2$</td>
<td>13</td>
<td>13</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td>Lower Sharpe ratio diff</td>
<td>11</td>
<td>15</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>Total no. of periods</td>
<td>26</td>
<td>26</td>
<td>43</td>
<td>43</td>
</tr>
</tbody>
</table>

(5) The number of periods that has a lower absolute difference between the in-sample and out-of-sample average Sharpe ratios than the counter approach.

Based on the results shown in Table 8, we can see that the cointegrated pairs give more accurate estimate of the out-of-sample return for S&P500. As shown in the first two rows of the table, out of 26 periods, the cointegrated pairs have positive $\beta_{t}$ for 17 periods whereas there are only 13 periods for the raw stock prices. There are also more periods in which the cointegrated pairs show better $R^2$ in return than that for raw stock prices, i.e. 15 out of 26 periods verses 11. This finding generalizes to other data sets, except NYSE. For NYSE, it has a surprising consistent estimation of $\beta_{t}$. However, the reverse is observed for prediction of standard deviation, the raw stock prices give better prediction than the cointegrated pairs in all of the data sets. When both return and standard deviation are taken into account, it is obvious that the cointegrated pairs have a better estimation of the out-of-sample Sharpe ratios than that of the raw stock prices in all data sets. Therefore, based on the above findings, we conclude that the use of the cointegrated pairs gives good prediction on the sign of the returns. The excess profit of the portfolio based on the cointegrated pairs over that of the one based on raw stock prices probably comes from the accurate estimation of the return in the out-of-sample periods using the in-sample information.

4.2 The Impact of the Number of Stocks on the Cointegrated Portfolios

4.2.1 Enlarge the Stock Pool. This experiment involves a larger set of the stock price data in the S&P500 component companies. The time span of the data is from 01 Jan, 2001 to 30 Dec, 2014. We extract the stock price time series of the largest 152 companies and rank them in descending order of their market capital, so the company with the largest capital ranks No.1 and the one
with least capital ranks No.152. We perform experiments using the top \( n \) companies and "rolling window" approach as before is adopted. The number of pairs found for each \( n \) is tabulated in Table 9.

**Table 9.** The number of pairs found from the top \( n \) stocks of the S&P500 data set

<table>
<thead>
<tr>
<th>( n )</th>
<th>35</th>
<th>50</th>
<th>75</th>
<th>100</th>
<th>125</th>
<th>152</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median no. of pairs</td>
<td>15</td>
<td>23.5</td>
<td>74</td>
<td>128</td>
<td>196.5</td>
<td>280.5</td>
</tr>
</tbody>
</table>

We apply TVCP with various risk tolerance \( \tau \) and mean variance method with various \( \kappa \) to the dataset and study their performance when we vary \( n \). The results are presented in Table 10.

We can see that the returns significantly increase from around 340% with 25 stocks to about 1700% with 152 stocks. This corresponds to about 5 times increase in return. The maximum drawdown, however, remains more or less the same. The Sharpe ratio of course increases when the return increases. On the other hand, the mean variance portfolio enjoys about 2.5 times gain in the return with the size of the stocks. The gain in the Sharpe ratio is also not as significant as that in the cointegrated portfolio. One may argue that there are 18.5 times increase in the number of bases for the cointegrated portfolio while the increase in those of the mean variance one is 4.3 times. But, the reality is that both portfolios are built using the same pool of stocks which changes from 25 to 152. Also, the cointegration one has fewer active stocks as some stocks may never be included in any cointegrated pairs.

Based on the experimental result, we conclude that (1) when more stocks are included, the performance of TVCP improves significantly whereas the performance of the mean variance approach improves only slightly; (2) the outperformance of the TVCP approach over the mean variance method is more substantial in the data set with more stocks. Thus, the larger the stock pool size, the more pairs we are able to find in general, and hence the better diversification of the risk and better allocation of the weights on the pairs.
Table 10. Comparison of the performance of TVCP and mean variance approach with various numbers of stocks

<table>
<thead>
<tr>
<th>No. of stocks</th>
<th>TVCP</th>
<th>Mean Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\tau)</td>
<td>(\kappa)</td>
</tr>
<tr>
<td>25</td>
<td>SR</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>371.34</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>23.27</td>
</tr>
<tr>
<td>35</td>
<td>SR</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>346.47</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>30.32</td>
</tr>
<tr>
<td>50</td>
<td>SR</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>700.70</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>33.81</td>
</tr>
<tr>
<td>75</td>
<td>SR</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>425.73</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>43.23</td>
</tr>
<tr>
<td>100</td>
<td>SR</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>691.12</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>42.16</td>
</tr>
<tr>
<td>125</td>
<td>SR</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>1114.07</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>31.81</td>
</tr>
<tr>
<td>152</td>
<td>SR</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>Return(%)</td>
<td>1912.55</td>
</tr>
<tr>
<td></td>
<td>Max DD(%)</td>
<td>33.51</td>
</tr>
</tbody>
</table>

Note: The strategy that gives the highest Sharpe ratio is bolded and underlined.
4.3 Further Analysis Using Synthetic Data

To further explore the performance of the cointegrated portfolio under the existence of market effect, we generate several sets of artificial data with common factors for comparing the performance between the cointegrated portfolio and the traditional mean variance method.

4.3.1 One Common I(1) Factor. We consider the case that a common I(1) factor $F_t$ exists for all stocks. We generate four time series using the following equation.

$$\ln S_{i,t} = \mu_i t + \beta_i F_t + e_{i,t}$$

The factor $\{F_t\}_{T=1}^T$ has ARIMA(2,1,0) structure where the AR coefficients are calibrated according to those obtained from the S&P500 index. The error terms are mutually independent Gaussian noise ($e_{i,t} \sim \mathcal{N}(0, \sigma_e^2)$, $\sigma_e = 3 \times 10^{-3}$) in the setting. We simulate four time series with the parameter $\mu = [10 \times 10^{-4}, 5 \times 10^{-4}, 2 \times 10^{-4}, 1 \times 10^{-4}]'$. Three sets of $\beta$ are used and they are $\beta = \hat{\beta}_0 \times [0.7, -0.2, 1.1, 1.5]'$ for $\hat{\beta}_0 = 0.5$, 1 and 2. We plot the distribution of the out-of-sample performance of the mean variance method and TVCP under various beta loading condition in Figure 4 for 5000 runs of each simulated setting. First we observe that the TVCP approach outperforms the mean variance in terms of achieving higher Sharpe ratio, higher return and lower maximum drawdown. By comparing the rows in the figure, we notice that mean variance strategy is performing worse and worse with the increase of the beta loading. We can see an obvious increase in the maximum drawdown in the $\hat{\beta}_0 = 2$ case (bottom row) compared with the $\hat{\beta}_0 = 0.5$ case (top row). However, the performance of TVCP is not much affected by the beta loading.

4.3.2 Two Common I(1) Factors. We then consider the case with two common I(1) factors $F_{1,t}$ and $F_{2,t}$. We generate four time series using the following equation.

$$\ln S_{i,t} = \mu_i t + \beta_{1,i} F_{1,t} + \beta_{2,i} F_{2,t} + e_{i,t}$$

The factors $\{F_{1,t}\}_{T=1}^T$ and $\{F_{2,t}\}_{T=1}^T$ are two independent factors that each has ARIMA(2,1,0) structure where the AR coefficients are also calibrated according to those obtained from the S&P500 index. The error terms are mutually independent Gaussian noise ($e_{i,t} \sim \mathcal{N}(0, \sigma_e^2)$, $\sigma_e = 3 \times 10^{-3}$) in the setting. Again, we simulate four time series with the same $\mu_i$ as in the experiment for
one common factor. The loading $\beta_1$ is set to $[0.7, -0.2, 1.1, 1.5]'$ and the loading $\beta_2$ is set to $\hat{\beta}_2 \times [1, 1, 1, 1]'$ where $\hat{\beta}_2 = 0.25$, 0.5 or 0.75. This means that we keep the first beta loading fixed and change the magnitude of the second beta. The out-of-sample performance of mean variance method and TVCP under various $\hat{\beta}_2$ values is shown in Figure 5. Based on the results, it is obvious that with the increasing influence of the second factor, the performance of the mean variance method and TVCP are both declining, especially the maximum drawdown. Moreover, when these two factors have similar magnitudes, i.e. in the case of $\hat{\beta}_2 = 0.75$, TVCP does not perform well. It has a high chance of giving maximum drawdown over 20%. However, in most of the two factor cases, TVCP still outperforms the mean variance method slightly as it captures the nonstationarity property of the factors and utilizes it to reduce the risk.
Figure 5. Out-of-sample Sharpe ratio, return and max drawdown distribution of cointegrated portfolio TVCP ($\tau = 20$) and mean variance portfolio ($\kappa = 0.005$) for 5,000 two factors experiments; Note: the top, middle and bottom rows are for $\beta_2$ with values 0.25, 0.5, and 0.75 respectively.

Comparing the performance of the cointegrated portfolios on the one factor cases with the two factor cases, we can see that the cointegration provides hedging for the dominating factor. When there are two similar factors, their effects may not be hedged sufficiently.

4.3.3 Connection Between the Experiments on Synthetic Data and Real Data. Finally, we analyze the variations on the performance of the mean variance method and TVCP on real data along with time which include different financial situation in history. The observations are then connected with those we observed using synthetic data. Here we focus on the mean variance method with $\kappa = 0.005$ and TVCP with $\tau = 20$ on S&P500 data set as both give similar median Sharpe ratio across the testing periods. Their performance along with time at each testing period is plotted in Figure 6. It is observed that the returns of TVCP are very stable while the returns of the mean variance method have more fluctuation (top left figure). The fluctuations are most
Figure 6. Comparison of the performance of mean variance ($\kappa = 0.005$) and TVCP ($\tau = 20$) on S&P500 dataset

severe before and after the financial crisis in 2008. Moreover, based on the annualized return SD plot (bottom left figure), TVCP achieves much lower return SD than the mean variance approach. The difference of the maximum drawdown between the TVCP approach and the mean variance approach is even more obvious (bottom right figure). TVCP consistently achieves lower drawdown than the mean variance approach does along the testing period. The mean variance approach as a maximum drawdown of 50% during the 2008 financial crisis.

Now, we would like to apply the factor analysis approach to the S&P500 data set for exploring if the difference in the performance of the two portfolios has any link to the common market factors. We adopt the method described in the paper by Bai and Ng [14] which the stock price time series is fitted to the following structure.

$$\ln S_{i,t} = c_i + b_i t + \sum_{j=1}^{N} \lambda_{i,j} F_{j,t} + e_{i,t}$$

Principal component analysis is applied to the detrended time series $\{\ln S_{i,t} - c_i - b_i t\}$ and we extract the nonstationary common factors $F_1, F_2, \ldots, F_N$. The parameters $\lambda_{i,j}$ are scaled to unit
length, i.e. $\sum \lambda_{i,j}^2 = 1$. It is observed that the first factor contributes most to the variance of the stock price movement whereas the other factors contribute almost equally. An example is shown in Figure 7. We repeat the process for each testing period and plot the temporal changes of the variance explained by the first factor, which is given by $\text{variance Explained} = \text{var}(F_1)/\sum_j \text{var}(F_j)$.

**Figure 7.** Factor variance distribution (S&P500)

![Factor variance distribution (S&P500)](image)

**Figure 8.** Comparison between the return spread and the variance explained by the first factor

![Comparison between the return spread and the variance explained by the first factor](image)

The corresponding difference in the return(return spread) between the mean variance portfolio and TVCP are also plotted alongside in Figure 8. As we have discovered in the experiment using synthetic experiment, the TVCP approach performs well when there exists a strong I(1) factor. This
corresponds to the bearish market around 2007 to 2009, that the effect of the first factor raises to accounts for about 50% of the variance and TVCP achieves much better performance than the mean variance method, not only in generating more return but also in controlling the downside risk. Although it suffers some loss around that period, the loss is only around 10% comparing to the 50% drop in the market (Figure 9) and 70% loss by the mean variance method (Figure 6). In 2011, we have another period that shows a dominating first factor but we do not observe noticeable difference in the performance between both portfolios. This is probably due to the fact that the index plunged but picked up again within half a year (Figure 9). Both portfolios performed well at that time. The above results agree with what we observed from the one-factor experiments.

Around the period between 2002 and 2006, the market is less volatile and the market is mostly bullish. We notice that the first factor only accounts for 30% percent of the variance. TVCP has a steady return of no more than 20%. The mean variance portfolio performs well most of the time. However, mean variance portfolio also has a moment of extremely poor performance in 2004 which leads to a loss of 20%.
5 Discussion and Conclusion

This paper discovers that stationary stocks achieve less discrepancy between the estimation of the in-sample and out-of-sample stock returns than that of the nonstationary stocks. Due to the advantage of making better return estimation, portfolios constructed by the stationary stocks have better performance than that constructed by nonstationary ones. Real financial data, however, have limited numbers of stationary stocks in the market for us to construct a diversified portfolio. The mean variance portfolio has a strong mathematical framework by optimizing the return and risk, but in practice it does not perform well because it uses the incorrect estimation of return and risk especially when we use nonstationary stocks. Unless we have frequent re-estimation and re-balancing of the portfolio weightings, the difference would cause poor performance.

We propose a framework of portfolio management approach that makes use of pairs of nonstationary stocks that show cointegration. We extract their stationary equilibria and construct the portfolio based on these equilibria. The out-of-sample performance of this portfolio management method is compared with other traditional methods. In summary, we have the following observations which show the advantages over the existing portfolio management methods.

(a) High Sharpe ratio. The Sharpe ratio is about 25% higher than the mean variance approach.
(b) High return. The return is comparable to the mean variance approach.
(c) Low downside risk. The maximum drawdown is about 40% less than the mean variance approach.
(d) Performance improves with the number of the stocks.

The first advantage is that its excellent ability to reduce the downside risk. Because of the use of the cointegrated pairs, hedging has already been built into the system. This effect is observed in our experimental results.

The second advantage is that using the equilibria gives a more consistent statistical properties between the in-sample and out-of-sample data. Therefore, the portfolio performance is more reliable, particularly in the estimation of the expected returns. The raw stock prices are known to have difficulties in estimating the returns actually. Our proposed approach rectifies this issue of inconsistent estimation of returns by using the cointegration properties and changes the stationary
time series into nonstationary ones.

The third advantage of our approach is that it can easily be integrated with many different portfolio management approaches. Rather than taking the original stock prices, the bases of the portfolio are simply replaced by the cointegration equilibria. For the purpose of demonstrating our novel approach, we use only the cointegrated equilibria to form the bases and nothing else in our experimental settings. One possible extension is that we can adopt a hybrid approach which the bases compose of both the cointegrated equilibria and other stock prices. The stationary stocks can be included or we can also include some selected stationary ones but with careful control on their portfolio weightings.

We note that there are other issues when we replace the bases. In equation 3, \( P_{k,t} \) appears comparable to the stock prices \( S_{k,t} \). The cointegration coefficient, \( \beta \), could be positive or negative. The implication of a negative \( \beta \) is that the system holds one stock while it shortsells the other one. If shortselling is not allowed, we could restrict our system to search and include any pairs with positive \( \beta \) only. This would further reduce the number of valid pairs in our portfolio. Also, when \( \beta \) is negative, it reduces the numerical values of \( P_{k,t} \). In some extreme cases, \( P_{k,t} \) could be negative. Comparing with the raw stock prices, \( S_{k,t} \) can never go below zero. When the values of \( P_{k,t} \) are negative or just bary positive, one has to be careful in the calculation of the percentage return and also the Sharpe ratio. Further, if both \( P_{k,t} \) and \( S_{k,t} \) are included in the same portfolio, the ranges of return and standard deviations are different. Normalization or adjustment is needed if both are included.

It is observed that the performance of TVCP is also sensitive to the risk-aversion level \( \tau \), in much the same way as \( \kappa \) to the mean variance portfolio. The scales of \( \tau \) and \( \kappa \) differ significantly. This is a direct result of the difference in the variance of the raw prices verses the variance of the equilibria.

Regarding to the numbers of the bases, we found that we have about 1 to 2% of pairs showing cointegration. This corresponds to the case that if we have about 100 stocks in our pool, the number of cointegrated pairs is roughly the same as the number of stocks in the pool. It has to be noted that the number of pairs does not correspond to the number of active stocks in the portfolio. The cointegrated pairs may often involve some stocks in common. Thus, the cointegrated portfolios are less diversified. If the number of stocks in the pool is small, we have very few valid pairs. The
portfolio can only invest in a limited number of stocks and we are unable to have a diverse portfolio.

When the number of the pairs is a concern, one simple and easy solution is to slightly relax the percentage level of the null hypothesis in the ADF test. By doing this, we allow for more pairs with mild cointegration properties in our system. This makes our portfolio more diversified. We have performed tests on this. We find that the portfolio gains improvement by including more pairs even some pairs have mild cointegration properties.

To the best of our knowledge, this is the first portfolio management method that is built on the trends of the cointegration relationships. The experimental results on the real data has shown that it is a promising portfolio management method. Also, due to its simple structure, it serves as a framework for constructing the portfolio. More work could be done to further enhance the performance of this method.

References


36


