ARTICLE IN PRESS

North American Journal of Economics and Finance xxx (xxxx) xxx

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

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

Multi-asset pair-trading strategy: A statistical learning approach

Tsai-Yu Lin [a], Cathy W.S. Chen [b,*], Fong-Yi Syu [a]

[a] Department of Applied Mathematics, Feng Chia University, Taichung 40724, Taiwan
[b] Department of Statistics, Feng Chia University, Taichung 40724, Taiwan

ARTICLE INFO

Keywords:
Exponentially weighted moving average (EWMA) model
GARCH model
Multiple pair assets
Predictive function
Tolerance interval
Volatility forecast

ABSTRACT

Pair trading is a widely used market neutral strategy and also a statistical arbitrage method that allows investors to take a position in two assets with similar trends in their historical data in order to gain low-risk profits. Combining both "diversification" and "pair trading", this study proposes a statistical learning method to explore the most promising pair among multiple pair assets for each trading time. We incorporate estimated volatility into tolerance limits as a predictive function for finding buying and selling signals in order to capitalize on market inefficiencies. One-step-ahead volatility prediction follows either the exponentially weighted moving average (EWMA) method or the GARCH model. The study selects five artificial intelligence (AI) stocks in the U.S. equities market to target profitability through the proposed strategy with a rolling window training approach over two annual testing periods from April 2017 to March 2019. We recommend that conservative investors use *p*-content at 95%, which is less adventurous and can generate positive excess profits. The idea behind this strategy is to help investments be more diversified and also more profitable.

1. Introduction

Diversification across asset classes is essential for the long-term success of a stock portfolio, but it is not a new concept in investment. The idea behind diversification is that a wider variety of investments will help yield a higher overall return; see Gallmeyer, Kaniel, and Tompaidis (2006); Gibson (2004). No matter how many assets are targeted, the main goal is to reduce risk and increase the diversification of an investment portfolio. The well-known strategy called pair trading is a common market-neutral trading tactic, as it allows investors the ability to gain low-risk profits by selecting two assets with similar trends in their historical data. This method is an important statistical arbitrage technique used by hedge funds after it was developed by the quantitative group under Nunzio Tartaglia at Morgan Stanley in the 1980 s. In fact, Do and Faff (2010) confirm that the pair trading strategy performs powerfully in profitability during periods of prolonged turbulence, including the 2008 global financial crisis. Krauss (2017) classifies pair-trading methods into five categories: distance methods, cointegration methods, time series methods, stochastic control methods, and other methods. Gatev, Goetzmann, and Rouwenhorst (2006) first promote using standardized historical prices to find potential pairs based on the distance method, followed by Bowen, Hutchinson, and O'Sullivan (2010); Do and Faff (2010); Engelberg, Gao, and Jagannathan (2009), and many others.

There are several ways to find matched pairs and pair-trading thresholds. Most choices of pairs adhere to the minimum distance method, the cointegration method (e.g. Law, Li, & Yu, 2018), or expert opinions. Chen and Lin (2017); Chen, Wang, Sribooncitta, and

* Corresponding author.
E-mail address: chenws@mail.fcu.edu.tw (C.W.S. Chen).

https://doi.org/10.1016/j.najef.2020.101295
Received 9 September 2019; Received in revised form 28 July 2020; Accepted 16 September 2020
Available online 22 September 2020

Please cite this article as: Tsai-Yu Lin, *North American Journal of Economics and Finance*, https://doi.org/10.1016/j.najef.2020.101295

Lee (2017) employ both the minimum distance method and expert opinions, in which the latter consider some common pair trades of similar companies in the same industry. Applying the cointegration test for identifying potential pairs may not always make it possible to find enough cointegration pairs in the sample (Chen, Lin, & Huang, 2019). Therefore, our study focuses on pairs in the same sector.

In order to find a transaction entry and exit signal, Chen, Chen, and Chen (2014) propose to use the upper and lower bounds from a three-regime threshold autoregressive model with GARCH specifications. Alternatively, Chen et al. (2017) use a smooth transition GARCH model with a second-order logistic function to generate the lower and upper bounds and further utilize one-step-ahead quantile forecasts to find a possible transaction opportunity. Chen and Lin (2017); Chen et al. (2019) suggest applying nonparametric/semi-parametric one-sided tolerance limits for return spreads in order to find trading entry and exit signals. Most of these studies suggest two pairs of trading strategies.

As a pair trade is a market-neutral strategy, the direction of the overall market does not theoretically affect its gain or loss, yet a diversified investment based on a pair-trading strategy to more practicable contents is still challenging. Combining both "diversification" and "pair trading", this study proposes a statistical learning method to explore the most promising pair among multiple pair assets for each trading time. The idea of statistical learning is to handle the problem of finding a predictive function from data - that is, this study introduces a simultaneous multiple pair-trading strategy through supervised learning, with a focus on semi-parametric tolerance limits that incorporate volatility forecasting. One-step-ahead volatility prediction is either via the exponentially weighted moving average (EWMA) method or the GARCH model of Bollerslev (1986). We then choose the lower and upper tolerance limits from the largest discrepancy return spread as a predictive function.

The advantages of our proposal include the following factors: 1) diversification, as we consider multiple pairs simultaneously; 2) easy to use for practical implications; and 3) we do not need to consider a long time-series period, but rather only a few months of historical datapoints are needed. More precisely, this strategy only needs a short training period for estimation, look-back sample sizes of 60, 80, or 100, based on EWMA volatility, or a 2-year training period for GARCH volatility suggested by Hwang and Pereira (2006). Chen et al. (2014); Chen et al. (2017) require a much longer time period as the training period since they employ nonlinear time series models, but this makes it impossible to execute pair trading from recently listed S&P500 companies. The novelty of this study's contribution with respect to that of Chen, Lin and Huang's proposal (2019) is that it helps investments be more diversified with a multi-asset pair-trading basket and GARCH volatility. We utilize a Gaussian pseudo-likelihood approach for estimating GARCH parameters for which the time series is not actually from a normal distribution.

To demonstrate the proposed strategy, this study selects five companies to represent artificial intelligence (AI) stocks - Alphabet, Amazon.com, Apple, Facebook, and Microsoft - covering April 1, 2017 to March 31, 2019, and extends the pair trading of Chen et al. (2019) from two assets to multiple pair assets. We consider two annual testing periods based on a rolling window training approach for evaluation: April 1, 2017 to March 31, 2018 and April 1, 2018 to March 31, 2019. We call this the "testing period" instead of the common term "testing set" since we deal with time series data and calculate profit from an investment at the end of the testing period, which includes transection cost. The results show that the proposed method generates extremely large returns of 44.96%.

The rest of this paper runs as follows. Section 2 introduces the nonparametric tolerance interval and one-step-ahead volatility prediction. Section 3 illustrates the algorithm for pair trading of multiple assets. Section 4 describes our data for ten baskets of AI stocks in the U.S. equity markets. We perform the pair trading strategy on each basket. Section 5 explains the proposed semi-parametric version of the tolerance interval approach to generate trading entry and exit signals and shows our multi-asset pair-trading profitability. Section 6 provides our concluding remarks.

2. Semi-parametric version of the tolerance interval

A population probability distribution function in practice is usually unknown. Tolerance limits in reality indicate the endpoints of the tolerance interval and are often used as important statistical tools for reference thresholds. Krishnamoorthy and Mathew (2009) note that the tolerance interval widely appears in various practical applications, such as quality control, environmental monitoring, industrial hygiene, and so on. Chen and Lin (2017) are pioneers in applying nonparametric one-sided tolerance limits to find trading entry and exit signals, whereas Chen et al. (2019) combine the nonparametric tolerance interval and volatility forecasting with the EWMA model into semi-parametric version of the tolerance interval. We apply pair trading to multiple assets by using this method and discuss the process in detail below.

If a sample is from a continuous population and does not fit a parametric model or fits a parametric model for which tolerance intervals are difficult to obtain, then one may seek nonparametric tolerance intervals for an intended application. Wilks (1941) is the first to treat the problem of determining tolerance limits and proposes that if a sample is from a continuous distribution, then the distribution of the p-content of the population between two order statistics is independent of the population sampled. Moreover, it is a function of only the particular order statistics chosen.

We assume the random sample $X_1, \ldots, X_n$ is from a continuous, nondecreasing probability distribution. Let $X_{(i)}$ denote the i^{th} smallest observation of $\{X_1, \ldots, X_n\}$ and be the order statistics for the sample. In order to construct a $(p, 1-\alpha)$ two-sided nonparametric tolerance interval, Wilk's result allows us to choose $X_{(m)}$ and $X_{(k)}$, $m < k$, and requires positive integers m and k so that:

$$P_{X_{(m)},X_{(k)}}\left\{P_X\left(X_{(m)} \le X \le X_{(k)} | X_{(m)}, X_{(k)}\right) \ge p\right\} = 1-\alpha. \tag{1}$$

Similarly, a $(p, 1-\alpha)$ one-sided lower nonparametric tolerance interval requires finding a positive integer l such that:

$$P_{X_{(l)}}\left\{P_X\left(X \ge X_{(l)} | X_{(l)}\right) \ge p\right\} = 1-\alpha. \tag{2}$$

On the other hand, a $(p, 1-\alpha)$ one-sided upper nonparametric tolerance interval requires finding a positive integer u such that:

$$P_{X_{(u)}}\left\{P_X\left(X \le X_{(u)} \middle| X_{(u)}\right) \ge p\right\} = 1-\alpha. \tag{3}$$

According to the above, we find $\left(X_{(l)}, X_{(u)}\right)$ as the lower and upper trading signals.

We utilize an R package provided by Young (2010) to obtain the tolerance intervals with content, including discrete and continuous conditions and regression tolerance limits. We further estimate the reference threshold as a predictive function by using the R-package "tolerance" to obtain the nonparametric one-sided tolerance of the spread value.

Forecasting volatility is an important task when dealing with financial trades. We consider two widely used methods to predict volatility and to supervise risks in the stock market. One is RiskMetrics^{TM} developed by J.P. Morgan (1996), and the other is the GARCH model of Bollerslev (1986). The EWMA and GARCH models are as follows:

$$R_t = \mu + a_t$$

$$a_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim NID(0,1),$$

$$\text{EWMA}: \sigma_t^2 = (1-\delta)a_{t-1}^2 + \delta\sigma_{t-1}^2, \tag{4A}$$

$$\text{GARCH}(1,1): \sigma_t^2 = \omega + \alpha a_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{4B}$$

where $NID(0,1)$ stands for normally and independently distributed a standardized normal distribution. Note that we utilize a Gaussian pseudo-likelihood approach for estimating GARCH parameters for which the time series is not actually from a normal distribution. The EWMA model is a special case of the integrated GARCH model, which usually assumes that the decay factor δ for daily data and monthly data is 0.94 and 0.97, respectively. It is thus necessary to impose restricted conditions on the GARCH parameters to ensure positive variance and covariance stationary:

$$\omega > 0,\ 0 \le \alpha,\ \beta < 1,\ \alpha + \beta < 1.$$

Implementation of the EWMA (or the volatility of RiskMetrics) forecast requires an initial value for σ_0^2, which we set to be the mean of the squared of initial observations.

3. A procedure for pair trading of multiple assets

The proposed model in this study only needs the recent n observations in the learning period (named the "look-back" sample size) in order to capture the recent variation. The continuously compounded return is the difference between the current period price and the preceding period price under a natural logarithm. In this study we shall conveniently denote it as the return:

$$r_t^k = \left[\ln\left(P_t^k\right) - \ln\left(P_{t-1}^k\right)\right] \times 100, \quad t = 1, \ldots, n, \tag{5}$$

where P_t^k is the price of stock k on day t, and n is the "look-back" sample size. The return spread based on the selected pair is:

$$R_t = r_t^{A_i} - r_t^{A_j}, \quad t = 1, \ldots, n. \tag{6}$$

The return spread is composed of the return of two assets. To incorporate estimated volatility information into the tolerance limits, we modify the return spread based on X_t, named as the adjusted return spread:

$$X_t = R_t / \widehat{\sigma}_t, t = 1, \ldots, n, \tag{7}$$

where $\widehat{\sigma}_t^2$ is the estimated volatility of R_t obtained from either (4A) or (4B).

We assume h assets $A_1, \ldots, A_h$ and consider all possible $h(h-1)/2$ pairs. We now illustrate the proposed trading procedure as follows.

Step 1: Calculate all pairwise return spreads, $\left\{R_t^{A_i - A_j}\right\}$, where $t = 1, \ldots, n, \quad i \ne j$ as in Eq. (6) for the "look-back" sample sizes n that are specified as 60, 80, and 100 days.

Step 2: Estimate all pairwise volatilities of $\left\{R_t^{A_i - A_j}\right\}$ based on either the EWMA or GARCH model, $\left\{\widehat{\sigma}_t^{A_i - A_j}\right\}$, where $t = 1, \ldots, n, \quad i \ne j$.

Step 3: Calculate all possible pairs' adjusted return spreads, $\left\{X_t^{A_i - A_j}\right\}$, where $t = 1, \ldots, n, \quad i \ne j$, which is given in Eq. (7). Next, choose the maximum value of them at each time t, $\left\{X_t^{max}\right\}$, which we call the maximum adjusted return spreads.

Step 4: Use $\{X_t^{max}\}$ to calculate both lower and upper one-sided tolerance limits with $(p, 1-\alpha) =$ (90%, 0.95) and (95%, 0.95) in Eqs. (2) and (3), which we take as trading entry and exit signals.

Step 5: Calculate the one-step-ahead forecast $\widehat{\sigma}_{n+1}^{A_i - A_j}$, where $i \ne j$ from either the EWMA or GARCH model. Combine the semi-parametric tolerance limits and volatility forecasts as a predictive function. Let $X_{n+1}^{A_i-A_j} = R_{n+1}^{A_i-A_j}/\widehat{\sigma}_{n+1}^{A_i-A_j}$, where $i \neq j$, and then choose the maximum $X_{n+1}^{A_i-A_j}$, where $i \neq j$ at time t, in which we call it X_{n+1}^{max}. When X_{n+1}^{max} is above the upper tolerance limit at time $n+1$, we sell one unit (share) of stock A_i and buy one unit (share) of stock A_j. If X_{n+1}^{max} is below the lower tolerance limit at time $n+1$, then we sell one unit (share) of stock A_j and buy one unit (share) of stock A_i at time $n+1$.

Step 6: The two annual testing periods run Step 1 to Step 5 by a rolling window training approach when a new datapoint is observed, and then the window shifts one-step ahead.

Step 7: Close the trading position and calculate the trading return when the given testing period ends. Suppose there is a total of T days during the period. The formula for realized and unrealized gains goes as follows.

$$\mathrm{Rev} = \sum_{d}\sum_{m=1}^{h}\frac{P_{LB}^{A_m}}{\sum_{k=1}^{h}\left(P_{LB}^{A_k}\right)}\ln\left(\frac{P_{d}^{A_m}}{P_{d-1}^{A_m}}\right), \tag{8}$$

where $d = 1, \cdots T$, $P_d^{A_m}$ stands for the stock price of stock A_m on day d, and $P_{LB}^{A_m}$ represents the total purchase price of all remaining A_m shares as of day $(d-1)$.

It is important to note several points.

1) The profit formula in (8) changes based on how many stocks the portfolio currently holds during each trade.
2) This study reflects the cost spent in the entire testing period, for which the multi-asset profit formula in (8) is different from that of Chen and Lin (2017); Chen et al. (2017), and Chen, Lin, Huang (2019).
3) For volatility strategy, we advise to use at least 500 historical datapoints for GARCH(1,1) models in order to obtain $\widehat{\sigma}_t^2$; otherwise, practitioners use the EWMA volatility based on the look-back sample period.

4. Data analytics from taking three stocks as a sample

AI now impacts most people's daily life, changing the way one lives, works, and plays. Its influence is likely to grow even more in the future. Moreover, the large fluctuations in AI stock prices, coupled with their frequent trading, have become a popular subject for both investors and news media. We select five AI stocks in the U.S. equities market - Alphabet, Amazon, Apple, Facebook, and Microsoft, whose daily opening prices are available from Yahoo Finance. We design three stocks among the five as one basket and obtain a total of ten baskets. Next, we establish the learning period from April 1, 2016, to March 31, 2017 and consider two annual testing periods for evaluation: April 1, 2017 to March 31, 2018 (named H_1) and April 1, 2018 to March 31, 2019 (named H_2). Table 1 lists the baskets of the three stocks as A_1, A_2, and A_3 and displays ten baskets of their stock codes and corresponding companies. These stock code symbols help with descriptive analysis.

Figs. 1 and 2 depict the time series plots of standardized prices for the ten baskets with the learning period and testing period from April 1, 2016 to March 31, 2019. It is very interesting to test the performance of the proposed method, because from the second half of 2018 all stocks fell significantly due to the sharp drop in the U.S. equities market, especially for Facebook.

Investors primarily utilize pair trading on highly correlated assets. Such standardized price changes are more meaningful than just standardized prices alone. Table 2 shows the set of Pearson correlation coefficients for the standardized price changes of two stocks from April 1, 2016 to March 31, 2017. Here, the pair correlation (GOOGL, MSFT) is as high as 0.6635, while the pair correlation (MSFT, AAPL) is as low as 0.2961.

5. Results

We now use the concept of a semi-parametric tolerance interval for the trading entry and exit signals. When X_{n+1}^{max} exceeds the upper limit, we sell the former and buy the latter; on the contrary, when X_{n+1}^{max} is below the lower limit, we sell the latter and buy the former. Fig. 3 shows the trading points for EWMA volatility. Because we employ three stocks, there are three pairs for each transaction. The red square represents the trading point of A_1A_2, the green circle is the trading point of A_2A_3, and the blue triangle is the trading point of

Table 1
Description of the ten baskets.

Basket	CompanyA_1	Code	CompanyA_2	Code	CompanyA_3	Code
1	Apple	AAPL	Amazon	AMZN	Facebook	FB
2	Apple	AAPL	Amazon	AMZN	Alphabet	GOOGL
3	Apple	AAPL	Amazon	AMZN	Microsoft	MSFT
4	Apple	AAPL	Facebook	FB	Alphabet	GOOGL
5	Apple	AAPL	Facebook	FB	Microsoft	MSFT
6	Apple	AAPL	Alphabet	GOOGL	Microsoft	MSFT
7	Amazon	AMZN	Facebook	FB	Alphabet	GOOGL
8	Amazon	AMZN	Facebook	FB	Microsoft	MSFT
9	Amazon	AMZN	Alphabet	GOOGL	Microsoft	MSFT
10	Facebook	FB	Alphabet	GOOGL	Microsoft	MSFT

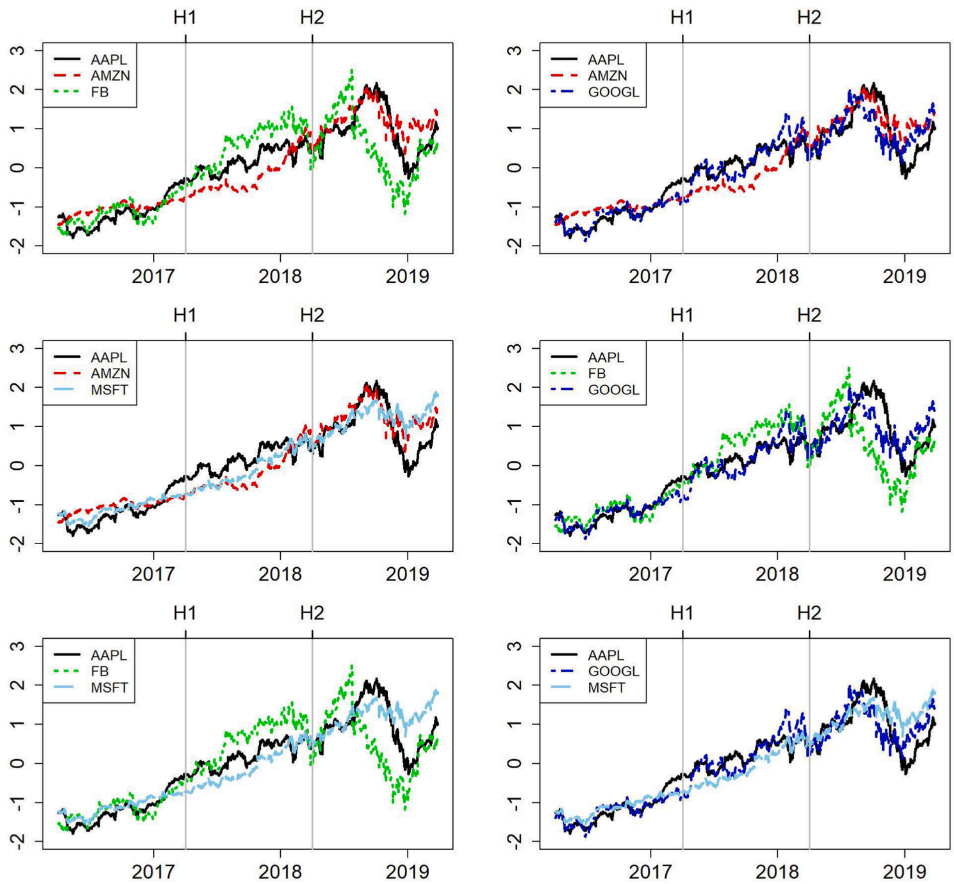

Fig. 1. Time series of standardized prices for baskets (1–6) from April 1, 2016 to March 31, 2019.

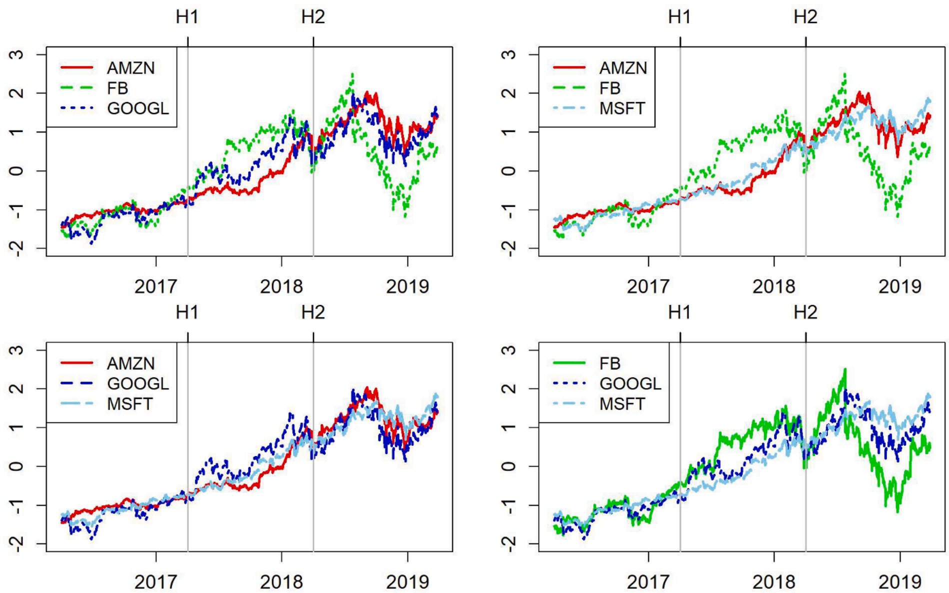

Fig. 2. Time series of standardized prices for baskets (7–10) from April 1, 2016 to March 31, 2019.

Table 2
Pearson correlation coefficient of the standardized price changes of five AI stocks during the training period (April 1, 2016 to March 31, 2017).

Stock	AAPL	AMZN	FB	GOOGL	MSFT
AAPL	1.0000	0.3189	0.3269	0.4022	0.2961
AMZN	0.3189	1.0000	0.4710	0.5708	0.4765
FB	0.3269	0.4710	1.0000	0.5187	0.4121
GOOGL	0.4022	0.5708	0.5187	1.0000	0.6635
MSFT	0.2961	0.4765	0.4121	0.6635	1.0000

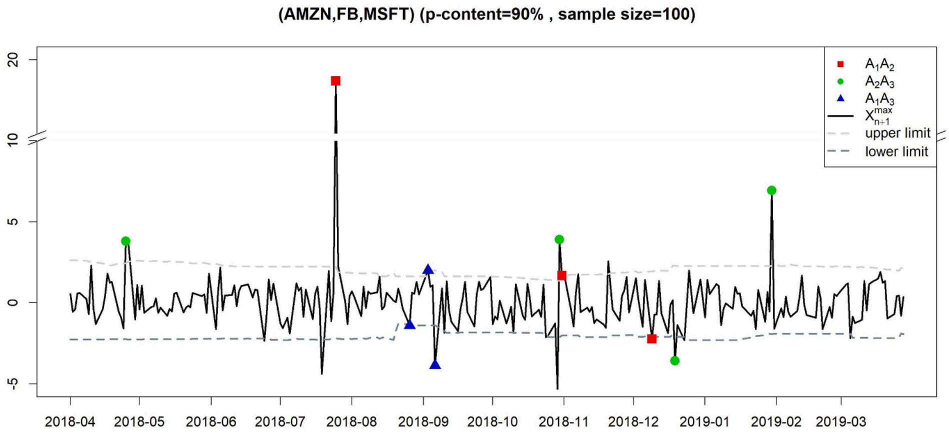

Fig. 3. The trading plots of p-content 90% and a sample size of 100 for the 8th basket (AMZN, FB, MSFT) using the EWMA volatility from April 1, 2018 to March 31, 2019.

A_1A_3.

We compare the efficiency in the number of different investments based on return of investment (ROI). The ROI formula is:

ROI = (current value of investment - cost of investment) / cost of investment. Since ROI is in percentage terms, we can clearly compare it with the returns from various set-ups (p-content and sample sizes), allowing one to measure a variety of time periods and baskets of investments against one another.

To help distinguish between the better strategy at each scenario, Tables 3 and 4 outline the averages of ROIs according to the groups of “p-content” and “sample size” at the end of the testing periods built on EWMA volatility and GARCH volatility, respectively. At the end of H_1, the best average of ROIs occurs at the 95%-content and hits 18.71% (24.79%) according to EWMA volatility (GARCH volatility). In terms of H_2, the best average of ROIs also appears at the 95%-content with 15.58% (14.15%) profits via EWMA volatility (GARCH volatility). The distribution of H_1 is consistently skewed to the right since the corresponding mean is larger than the median. Three out of four have extremely high CV (coefficient of variation) when $(p, n) = (90\%, 100)$. Overall, the CV of 90%-content in fixed sample sizes is larger than that of 95%-content, while the opposite result occurs on the mean. In other words, the ROI of 95%-content is better and more stable than 90%-content.

Fig. 4 sums up the averages of ROIs in the two testing periods for 10 baskets. Generally, the 3rd basket (AAPL, AMZN, MSFT) has the highest profitability, while the 5th basket (AAPL, FB, MSFT) has the lowest one. The proposed strategy incorporating GARCH (1,1) volatility outperforms that of EWMA volatility in H_1 period, but the opposite result occurs in H_2 period. Overall, the ROI over the H_1

Table 3
Summary statistics of ROIs with transaction cost of 0.2% and incorporating EWMA volatility during the two testing periods.

Period	p-content	Sample size	Mean	SD	Min.	Median	Max.	CV (%)
H_1	90%	60	17.93	13.73	1.13	14.28	43.12	76.58
		80	15.96	11.89	1.62	15.28	40.96	74.51
		100	14.36	13.54	1.62	11.51	42.52	94.28
	95%	60	16.29	6.86	7.97	15.21	29.65	42.13
		80	18.71	12.45	1.66	14.43	44.96	66.55
		100	17.79	12.59	1.53	15.17	41.05	70.74
H_2	90%	60	13.56	10.22	−3.08	14.91	29.60	75.35
		80	10.64	11.07	−12.67	12.56	22.49	104.04
		100	7.60	10.94	−10.29	9.70	22.77	144.00
	95%	60	15.56	6.68	1.65	15.07	26.66	42.93
		80	14.24	6.61	−0.76	15.27	23.27	46.40
		100	15.58	4.26	10.32	14.96	23.27	27.35

Note: CV: coefficient of variation.

Table 4
Summary statistics of ROIs with transaction cost of 0.2% and incorporating GARCH (1,1) volatility during two testing periods.

Period	p-content	Sample size	Mean	SD	Min.	Median	Max.	CV (%)
H_1	90%	60	22.79	11.36	7.57	22.63	39.31	49.85
		80	18.18	10.25	5.72	13.90	40.72	56.39
		100	21.87	10.86	9.56	20.12	39.28	49.66
	95%	60	20.38	11.72	0.75	17.21	44.08	57.51
		80	22.33	13.79	0.75	17.21	44.08	61.78
		100	24.79	9.82	14.23	22.03	39.87	39.61
H_2	90%	60	7.42	8.54	−7.15	6.95	20.43	115.24
		80	6.54	6.96	−3.71	5.29	17.33	106.37
		100	4.32	9.89	−10.52	4.72	19.86	228.99
	95%	60	12.60	6.32	−0.27	15.33	18.63	50.19
		80	14.15	9.13	−0.86	14.87	33.69	64.48
		100	7.62	11.07	−16.73	10.36	24.06	145.25

Note: CV: coefficient of variation.

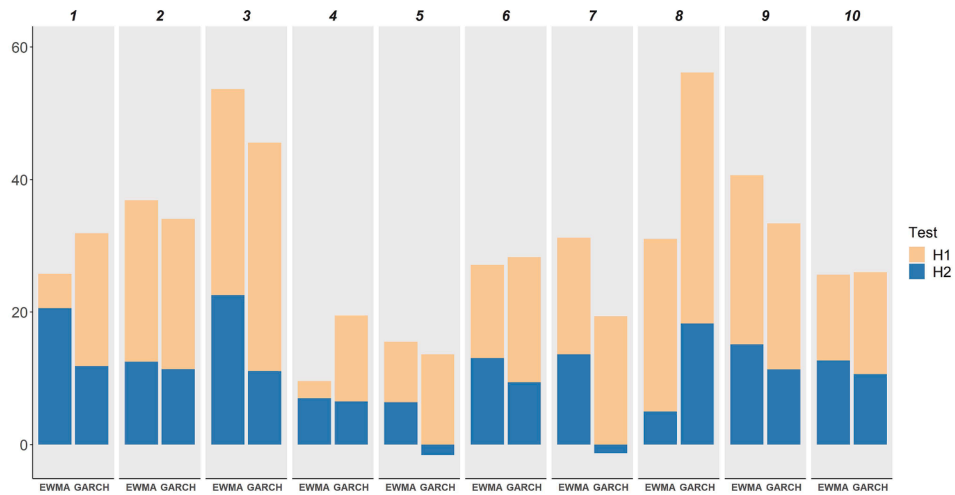

Fig. 4. Sum of averages of ROIs in the two testing periods for 10 baskets.

time period is higher than that over H_2. However, there is less evidence that it can enhance trading profits considerably just due to the prediction from the complicated GARCH model.

Fig. 5 displays the ROIs of each basket under two p-content conditions (90%, 95%) for EWMA volatility during the two testing periods, representing the average ROI on the look-back sample size (60, 80, 100). Except for the 5th and 8th baskets of 90%-content at H_2, all other trading scenarios have positive average returns. All trading scenarios exhibit a positive average ROI over the H_1 time period, with the highest ROI in the 3rd basket and the lowest in the 4th basket. In most cases, the ROI of 95%-content is better than 90%-content. Overall, the variation of ROI for the H_1 period is greater than that of H_2.

Tables 5 and 6 list all baskets of ROIs with a transaction cost of 0.2% for EWMA volatility and the number of trades. We state our findings as follows.

1. The ranges of ROIs are (1.13%, 44.96%) and (-12.67%, 29.60%) for H_1 and H_2, respectively. There are three outliers with negative ROIs that happen in H_2. We can avoid those negative profits if we employ p-content at 95%. The testing time horizon is a critical factor when we want to take profit from the trade.
2. The overall average ROIs are 16.84% and 12.86% for the H_1 and H_2 time periods, respectively.
3. The mean number of trades for H_1 and H_2 is about 4 times and 3.5 times, respectively.
4. In period H_1, the highest ROI occurs at the 8th basket (AMZN, FB, MSFT), $(p,n) = (95\%,80)$, as long as the number of trades is 2 times, and the lowest ROI happens at the 4th basket (AAPL, FB, GOOGL), $(p,n) = (90\%,60)$, where the number of trades is 5 times. Nevertheless, the ROI is still 1.13% for the 4th basket.
5. In period H_2, the highest ROI occurs at the 1st basket (AAPL, AMZN, FB), $(p,n) = (90\%,60)$, as long as the number of trades is 4.5 times, and the lowest ROI occurs at the 8th basket (AMZN, FB, MSFT), $(p,n) = (90\%, 80)$, where the number of trades is 4.5 times.

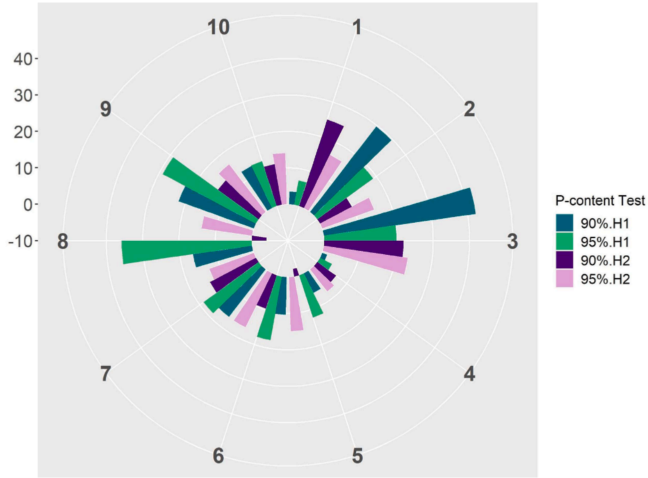

Fig. 5. The averages of ROIs for 10 baskets based on *p*-content (90%, 95%) for EWMA volatility during the two testing periods.

Table 5
Baskets' 1–4: ROI with transaction cost of 0.2% and the number of trades during the two testing periods for the basket of look-back sample sizes (60, 80, 100) and *p*-content (90%, 95%).

No.	Basket	*p*-content	Sample size	H_1		H_2	
				No. of round-trip trades	0.2% cost ROI	No. of round-trip trades	0.2% cost ROI
1	AAPL	90%	60	5.5	3.61	4.5	29.60
	AMZN		80	6.5	3.62	4.5	22.49
	FB		100	7.5	3.36	5.5	22.77
		95%	60	3	7.97	3.5	26.66
			80	3	7.97	1	10.49
			100	3	4.50	2	11.53
2	AAPL	90%	60	5	30.40	4	10.90
	AMZN		80	5.5	28.07	5	8.36
	GOOGL		100	5	30.52	4	10.44
		95%	60	3.5	21.83	2	15.04
			80	1.5	15.91	1.5	15.43
			100	4	19.24	2	15.04
3	AAPL	90%	60	4.5	43.12	3	22.00
	AMZN		80	6.5	40.96	3	22.00
	MSFT		100	7.5	42.52	5	21.51
		95%	60	2.5	19.43	2.5	23.27
			80	1.5	24.56	2.5	23.27
			100	3	15.81	2.5	23.27
4	AAPL	90%	60	5	1.13	3	−1.90
	FB		80	5	1.62	4	10.47
	GOOGL		100	5	1.62	4	10.47
		95%	60	2.5	8.03	2.5	11.85
			80	2.5	1.66	2	−0.76
			100	3	1.53	2.5	11.88

Note: The result is based on EWMA volatility.

6. Conclusion

This study contributes towards utilizing the proposed statistical learning method that monitors deviations in multiple pair assets and triggers buying and selling signals in order to capitalize on market inefficiencies. Different from previous studies in the literature, the statistical learning method herein selects a pair of stocks among multiple pair assets for each trading time. Combining semi-parametric tolerance limits and volatility forecasts as a predictive function, we find trading entry and exit signals. The selection reflects the lower and upper tolerance limits from the largest discrepancy return spread.

We use five AI stocks and take the baskets of look-back sample sizes (60, 80, 100) and *p*-content (90%, 95%) during two annual testing periods. According to the empirical results, this method generates most of the positive excess returns. Even though all selected AI stocks took a bit of a beating at the end of 2018 during the testing period, the ROIs of H_2 are still profitable. We recommend that conservative investors use *p*-content at 95%, which is less adventurous and can generate positive excess profits based on the two annual testing periods. As for the sample size setting, it is not a key factor for profit analysis.

Table 6
Baskets' 5–10: ROIs with transaction cost of 0.2% and the number of trades during the two testing periods for the basket of look-back sample sizes (60, 80, 100) and *p*-content (90%, 95%).

No.	Basket	*p*-content	Sample size	H_1		H_2	
				No. of round-trip trades	0.2% cost ROI	No. of round-trip trades	0.2% cost ROI
5	AAPL	90%	60	3.5	8.65	4	−3.08
	FB		80	6	5.70	5.5	−3.15
	MSFT		100	6.5	3.67	5	−0.08
		95%	60	3	13.69	2.5	15.83
			80	2.5	12.96	2.5	13.98
			100	2.5	10.05	3	14.89
6	AAPL	90%	60	3.5	10.53	4.5	18.35
	GOOGL		80	7	12.26	5	18.56
	MSFT		100	3	8.32	5.5	−7.93
		95%	60	2	14.54	3.5	15.10
			80	2	24.18	3.5	15.10
			100	2	14.54	2	19.24
7	AMZN	90%	60	4	30.41	4	18.83
	FB		80	6.5	15.77	4	15.19
	GOOGL		100	6	3.31	4.5	8.87
		95%	60	2	10.59	3	18.09
			80	1.5	12.90	2.5	10.51
			100	3	32.58	2.5	10.32
8	AMZN	90%	60	5.5	18.03	3.5	10.81
	FB		80	6	16.09	4.5	−12.67
	MSFT		100	7	14.95	5.5	−10.29
		95%	60	2.5	21.29	2.5	1.65
			80	2	44.96	2	20.16
			100	2.5	41.05	3	20.21
9	AMZN	90%	60	3.5	24.47	4	18.62
	GOOGL		80	5.5	20.76	5.5	13.88
	MSFT		100	6.5	20.66	5.5	8.96
		95%	60	2.5	29.65	2.5	13.92
			80	2.5	29.65	2	18.64
			100	2	27.82	1.5	16.84
10	FB	90%	60	5	8.96	3.5	11.47
	GOOGL		80	6	14.78	4.5	11.24
	MSFT		100	6	14.71	4.5	11.24
		95%	60	2	15.88	2.5	14.16
			80	2.5	12.31	2	15.59
			100	3.5	10.84	2	12.54

Note: The result is based on the EWMA volatility.

The advantage of our method is to provide investors with simple trading signal settings and to make investments more efficient. We utilize “tolerance” (Young, 2010) and “rugarch” (Ghalanos, 2020) packages in R for tolerance limits and GARCH volatility to implement this statistical learning base for the pair-trading strategy, and it takes only up to one minute CPU time for one basket in one testing period. Finally, this statistical learning base for multiple asset pair trading is not just for stocks, but can also be applied to bonds, commodities, currencies, and cash equivalents.

CRediT authorship contribution statement

Tsai-Yu Lin: Writing - original draft, Methodology, Supervision, Validation, Visualization. **Cathy W.S. Chen:** Conceptualization, Investigation, Visualization, Writing - review & editing, Funding acquisition. **Fong-Yi Syu:** Software, Validation.

Acknowledgements

The authors thank the Editor-in-Chief and anonymous referee for their valuable time and careful comments, which have improved this paper. Tsai-Yu Lin’s research is funded by the Ministry of Science and Technology, Taiwan (MOST 107-2118-M-035-007) and Feng Chia University (Research Grant No. 19H00808). Cathy W. S. Chen’s research is funded by the Ministry of Science and Technology, Taiwan (MOST107-2118-M-035-005-MY2).

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Bowen, D., Hutchinson, M. C., & O’Sullivan, N. (2010). High-frequency equity pairs trading: Transaction costs, speed of execution, and patterns in returns. *The Journal of Trading, 5*, 31–38.

Chen, C. W. S., Chen, M., & Chen, S. Y. (2014). Pairs trading via three-regime threshold autoregressive GARCH models. *Modeling Dependence in Econometrics, 251*, 127–140.

Chen, C. W. S., & Lin, T. Y. (2017). Nonparametric tolerance limits for pair trading. *Finance Research Letters, 21*, 1–9.

Chen, C. W. S., Lin, T. Y., & Huang, T. Y. (2019). Incorporating volatility in tolerance intervals for pair-trading strategy and backtesting. *Journal of Risk Model Validation, 1*, 63–94.

Chen, C. W. S., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition GARCH models. *North American Journal of Economics and Finance, 39*, 38–55.

Do, B., & Faff, R. (2010). Does simple pairs trading still work? *Financial Analysts Journal, 66*, 83–95.

Engelberg, J., Gao, P., & Jagannathan, R. (2009). An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity. In Proceedings of the 3rd Singapore International Conference on Finance.

Gallmeyer, M. F., Kaniel, R., & Tompaidis, S. (2006). Tax management strategies with multiple risky assets. *Journal of Financial Economics, 2*, 243–291.

Ghalanos, A. (2020). rugarch: Univariate GARCH models. R package version 1.4-2.

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies, 19*, 797–827.

Gibson, R. C. (2004). The rewards of multiple-asset-class investing. *Journal of Financial Planning, 17*, 58–71.

Hwang, S., & Pereira, P. L. V. (2006). Small sample properties of GARCH estimates and persistence. *The European Journal of Finance, 12*, 473–494.

Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys, 31*, 513–545.

Krishnamoorthy, K., & Mathew, T. (2009). *Statistical tolerance regions: Theory, applications, and computation*. John Wiley & Sons.

Morgan, J. P. (1996). RiskMetrics. J. P. Morgan Technical Document (4th ed.) The Journal of Alternative Investments, J. P. Morgan: New York.

Law, K. F., Li, W. K., & Yu, L. H. (2018). A single-stage approach for cointegration-based pairs trading. *Finance Research Letters, 26*, 177–184.

Wilks, S. S. (1941). Determination of sample sizes for setting tolerance limits. *Annals of Mathematical Statistics, 12*, 91–96.

Young, D. S. (2010). Tolerance: An R package for estimating tolerance intervals. R package version 0.1.0.