Pair Trading in Tehran Stock Exchange based on Smooth Transition GARCH Model

Saeed Bajalan[1]
Reza Eyvazlu[2]
Guilda Akbari[3]

Abstract

In this research, we use a pair trading strategy to make a profit in an emerging market. This is a statistical arbitrage strategy used for similar assets with dissimilar valuations. In the present study, smooth transition heteroskedastic models are used with the second-order logistic function for producing thresholds as trading entry and exit signals. For generating upper and lower bounds, we apply the rolling window approach and one-step-ahead quantile forecasting. Markov chain Monte Carlo sampling method is used for optimizing the parameters. Also, passive strategy in the out-of-sample period is used to compare the profits. The population consists of 36 daily stock returns in Tehran Stock Exchange. Then, we select ten pairs from these stocks and use Minimum Square Distance method, and five pairs from one industrial sector. Finally, we see strategy1 and 2 have positive returns in the out-of-sample period, and they produce higher returns than passive strategy.

Keywords: *Pair trading, Smooth transition GARCH model, Rolling window approach, One-step-ahead quantile forecasting, Out-of-sample-period*

---

1. Assistant Professor of Finance, Faculty of Management, University of Tehran, Iran. corresponding author (Email: saeedbajalan@ut.ac.ir)
2. Assistant Professor of Finance, Faculty of Management, University of Tehran, Iran. (Email: eivazlu@ut.ac.ir)
3. Master of Finance, University of Tehran, Tehran, Iran.

Introduction

Pair trading is one of a market-neutral trading strategy that is used in a pair of highly correlated instruments such as two stocks, exchange-traded funds (ETFs), currencies, commodities or options. This strategy matches a long position with a short position in a pair. When the correlation is weak, traders go long on the under-performer while simultaneously going short on the over-performer, closing the positions as the relationship returns to its statistical norm. The profit is gained from the difference in the price change between the two instruments. Without paying attention to the movement direction, a profit can be realized if the long position goes up more than the short, or the short position goes down more than the long (in a perfect situation, the long position will rise and the short position will fall, but this is not a requirement for making a profit). Pair traders can gain profit during a variety of market conditions, including periods when the market goes up, down or sideways, and during periods of either low or high volatility. As, this strategy is independent of the market movement, it is said to be market neutral. Pair trading is a mean-reverting strategy, assuming that prices will revert to historical trends (Skiena). The pair trading strategy takes advantage of market inefficiency. Pair trading can be described as two different stocks may be exposed to similar underlying market conditions which cause a tendency in their prices to move together (HP Mashele,SE Terblanche & JH Venter, 2015).

Pair trading pioneered by Nunzio Tartaglia's qua t group at Morgan Stanley in the 1980s. It remains an important statistical arbitrage technique used by hedge funds. They found that certain securities were correlated with their day-to-day price movement. Krauss,C. reviewed literature closely related to the umbrella term of pair trading, covering both univariate and multivariate strategies. Clustered by the pair trading approach, they classified their findings and suggestions as: distance approach, co-integration approach, time series approach, stochastic control approach, and other approaches (Krauss, C., 2016). For matching pairs in the."pair trading" strategy Gatev, Goetzmann and Rouwenhorst used the minimum distance approach between normalized historical daily prices over 1962-2002 (Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G., 2006). That strategy was using relative stock price movements as triggers to open or close a pair position. They presented two pieces of empirical evidence. First, although raw returns have fallen, the risk-adjusted returns have continued to persist despite increased hedge fund activity. Second, their results suggested that the change in risk-adjusted returns of pair trading is accompanied by the diminished importance of a common factor that drives the returns to pair strategies.

Grima investigates the long-term pricing relationship among crude oil, unleaded gasoline, and heating oil futures prices, and finds that these commodities future prices are co-integrated. This study finds that the spreads between crude oil and its end products are stationary (Girma, 1999). Vidyamurthy suggests that risk arbitrage in its general connotation relates to trading around corporate events that alter the capital structure of a firm. (Vidyamurthy, 2004). Perlin uses the co-integration method to find two stocks that move together and take long/short positions when they diverge abnormally, hoping that the prices will converge in the future in the Brazilian stock market (Perlin, 2009). Pair selection is executed by maximizing Pearson correlation between standardized price time series. However, a potential problem can occur in the co-integration of pair trading such as occasionally not finding enough co-integration pairs in the sample (see (Chen, C. W. S., Chen, M., & Chen, S. Y., 2014)).
Elliot, Van der Hoek and Malcom apply a Kalman filter in the pair trading strategy to estimate a parametric model of the spread and consider two similar stocks which trade at some spreads (Elliott, R. J., van der Hoek, J., & Malcolm, W. P., 2005). This paper proposes a mean-reverting Gaussian Markov chain model for the spread which is observed in Gaussian noise. The methodology of this research generates wealth from any quantities in financial markets which are observed to be out of equilibrium. This research employs the autoregressive conditional heteroscedastic (ARCH) model and generalized ARCH (GARCH) model for describing dynamic volatility in financial time series (Engle, 1982) and (Bollerslev, 1986). In the GARCH models, volatility is time varying and dependant on both the past volatility and innovations (Jung, 2016).
To use the symmetric GARCH family models in the various types of volatility asymmetry and nonlinearity, these models have been improved and developed. The double threshold GARCH and asymmetric smooth transition GARCH models (Chen, 2006) (Gerlach, R., & Chen, C. W. S., 2008), and many other models select pairs via the minimum squared distance (MSD) method and construct a pair trading strategy with three-regime threshold autoregressive models with GARCH effects (Chen, C. W. S., Chen, M., & Chen, S. Y., 2014). For the stochastic spread method, Liu and Timmermann study optimal convergence trades under both recurring and non-recurring arbitrage opportunities represented by continuing and ‘stopped’ co-integrated price processes and considers both fixed and stochastic (Poisson) horizons (Liu, J., & Timmermann, A., 2013).

Literature Review

The main idea of this research was based on the work of Cathy, Chen and Zona Wang, Songsak Sriboonchitta and Lee, who implemented pair trading strategy and used ST-GARCH model on the return spread of 36 stocks that they took from the Dow Jones Industrial Average (DJIA), the New York Stock Exchange (NYSE) and the NASDAQ stock market. These researchers employed three strategies to generate threshold values as entry and exit signals in the daily trading of the pairs. They also implemented these procedures for three out-of-sample periods and compared them with each other (Cathy W.S. Chen, Zona Wang, Songsak Sriboonchitta, Sangyeol Lee, 2016).

Chan and Tong introduced a smooth transition (ST) autoregressive model to allow for flexibility in model parameters through a smooth transition, which gained popularity via Teräsvirta appear to be the first to discuss the second-order logistic function in ST models. In fact, the ST-GARCH model considered as a model with three regimes, where the first regime was related to extremely low negative shocks; the middle regime represented low absolute returns; and the third regime corresponded to high positive shocks (Chan, K. , & Tong, H., 1986) (Teräsvirta, 1994). (Jorion, 1997) and (Teräsvirta, 1994).

Chen modeled the return spread of potential stock pairs as a three-regime Threshold Autoregressive GARCH (TAR-GARCH) models. The upper and lower regimes in the model were used as trading entry and exit signals (Chen, C. W. S., Chen, M., & Chen, S. Y., 2014).

Data

In this research, data consist of 36 stocks from Tehran Stock Exchange. We obtain the data from Tehran Exchange website and the TSEClient software. Our data span a 10-year period from 21 March 2007 to 20 March 2017. The selection of the dates is based on the beginning and ending dates of the Iranian year. We assume that the in-sample period is 8 years (from 21 March 2007 to 20 March 2016) and the out-of-sample period is one year (from 21 March 2016 to 20 March 2017).

Two methods are employed to select pairs. First, we calculate Minimum Square Distance (MSD) between normalized prices in each pair (630 possible pairs, C_2^{36}) and then selecte 10 pairs with minimum MSD. In Addition, 5 other pairs are selected from separate important industrial sectors.

We calculate daily log returns of all stocks and the daily return spread of stocks in each 15 selected pairs. Consequently, we implemented ST-GARCH model with second-order logistic function along with the in-sample period, and estimate threshold values for upper and lower bounds as entry and exit signals by using the two strategies.

Finally, we assume that if we do not use these strategies and just buy and hold stocks in the out-of-sample period, how would the results be. The reason we use to buy and hold strategy is that in Tehran Stock Exchange, high volatility exists in stock price change. Moreover, emotional behavior affects the process of trading. Also, we examine the question that which method led us to more profit.

Methodology

Pair Trading Procedure

We choose the trading pairs based on two methods: the MSD rule, and the pairs which are at the same industrial sector. To do this, we calculate the minimum square distance between the normalized price series of each pair. At first, calculating the normalized price $P_t^{(j)}$ of asset *j* at time *t* is given by:

$$P_t^{(j)} = \frac{P_t^{(j)} - E(P^{(j)})}{\sigma^{(j)}} \qquad (1)$$

Where $P_t^{(j)}$ is the closing price of asset *j*, $E(P_t^{(j)})$ is the average of $P_t^{(j)}$, and $\sigma^{(j)}$ is the standard deviation of the respective stock price.

Then for calculating MSD, the following formula is used:

$$MSD = \sum_{t=1}^{n} (P_t^A - P_t^B)\,, \qquad (2)$$

Where $P_t^{(j)}$ is the normalized price of asset *j* at time *t*, pairs are selected based on the smallest MSD.

To implement pair trading procedure, we obtain the return spread for the in-sample period for the selected pairs by using MSD method:

$$y_t = r_t^A - r_t^B \qquad , t = 1, \ldots, n \qquad (3)$$

Then the two strategies are implemented as follows:

Strategy1: Fitting the St-GARCH model with a second-order logistic function for the return spread and estimating the threshold values with the rolling window approach.

Strategy2: Fitting the ST-GARCH model with a second-order logistic function for the return spread and obtaining the thresholds values by the one step ahead quantile forecasting. The quantiles which are used here are 20% for a lower bound and 80% for an upper bound as entry and exit signals.

If the return spread is above the upper bound at time *t*, we will sell one share of stock A and buy one share of stock B. On the contrary, if the return spread is below the lower bound at time t, we sell one share of stock B and buy one share of stock A. We assume that we have 245 shares of each pair of stocks (since short selling does not permit in Tehran Stock Exchange, we assume that we have the same number of shares as the number of trading days in the out-of-sample period). If the spread does not cross the thresholds, we do not buy or sell the stocks.

Fitting the smooth transition GARCH model to return spread

In this section, we employ two methods to produce upper and lower bounds as trading entry and exit signals. In both methods, we fit ST-GARCH model with a second-order logistic function and error terms following a standardized Student’s t-distribution to pair return spreads (González-Rivera, 1998). In the first strategy, we use rolling window approach to estimate threshold values in the model as trading entry and exit signals. The second strategy is to estimate one-step-ahead quantile forecast and produce entry and exit signals.

In this research we used ST-GARCH(1,1) model with second-order logistic functions to fit to the return spread, as follows:

$$y_t = \mu_t^{(1)} + F(Z_{t-d};\gamma, C1, C2)\mu_t^{(2)} + a_t \quad (4)$$

$$a_t = \sqrt{h_t}\varepsilon_t \qquad \varepsilon_t \overset{i.id}{\sim} t^*(\nu)$$

$$h_t = h_t^{(1)} + F(Z_{t-d};\gamma, C1, C2)h_t^{(2)}$$

$$\eta_t^{(i)} = \varphi_0^{(i)} + \varphi_1^{(i)} y_{t-1}$$

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} a_{t-1}^2 + \beta_1^{(i)} h_{t-1} \qquad , i = 1,2$$

$$F(Z_{t-d};\gamma, C1, C2) = \frac{1}{1+\exp\{\frac{-\gamma(Z_{t-d}-C1)(Z_{t-d}-C2)}{s_Z}\}}, C1 < C2$$

Where y_t is the return spread, $h_t = Var(y_t \mid F_{t-1})$, F_{t-1} represents the information set at time $t-1$, $t^*(\nu)$ is a standardized t-distribution with ν degrees of freedom, Z_t is the threshold variable, which can be a past observation of y_t, d is a delay lag, s_Z is a sample standard deviation of Z_t, and $F(0)$ is a continuous distribution changing from zero to one.

Some conditions are needed to ensure positive variance and covariance stationarity. These conditions are as follows (Anderson, H.M, Nam, K., &Vahid, F., 1999):

$$\alpha_0^{(1)}, \alpha_1^{(1)}, \beta_1^{(1)} > 0, \quad \alpha_0^{(1)} + \alpha_1^{(2)} > 0, \quad \beta_1^{(1)} + \beta_1^{(2)} > 0, \quad (5)$$

$$\left(\alpha_0^{(1)} + 0.5\alpha_1^{(2)}\right) + \left(\beta_1^{(1)} + 0.5\beta_1^{(2)}\right) < 1$$

Bayesian inference

Here, we define ε_t as a standardized Student’s t-distribution with ν degrees of freedom, so the conditional likelihood function for the ST-GARCH model is:

$$P(\, y^{s+1,n} \mid \theta\,) = \prod_{t=s+1}^{n}\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}}\frac{1}{\sqrt{h_t}}[1+\frac{(y_t-\mu_t)^2}{(\nu-2)h_t}]^{-\frac{\nu+1}{2}}\}, \quad (6)$$

Where θ represents all model parameters, $y^{s+1,n} = (y_{s+1}, \dots, y_n)$, n is the sample size, and h_t is the conditional variance. Note that $\theta = (\emptyset_i^{'}, \alpha_i^{'}, c^{'}, \gamma, \nu, d)^{'}$; $\emptyset_i = \emptyset_0^{(i)} + \emptyset_1^{(i)}$; $\alpha_i = \left(\alpha_0^{(i)}, \alpha_1^{(i)}, \beta_1^{(i)}\right), i = 1{,}2,$ and $c = (c_1, c_2)^{'}$.

In ST model we choose the prior values for parameters and the likelihood function and multiply the priors to give the posterior kernels. The conditional posterior is given by the following formula:

$$P(\,\theta_1 \mid y^{s+1,n}, \theta_{\neq 1}\,) \propto P(\,y^{s+1,n} \mid \theta\,).P(\,\theta_1 \mid \theta_{\neq 1}\,) \quad (7)$$

We use Metropolis-Hastings method to draw MCMC iterates for the parameter groups. For implementing MCMC model by Metropolis-Hastings model, we consider burning in sample period for the first half of each iteration. So we do not consider the first half of each iteration. Parameters are estimated by MHadaptive package in R program (Chivers, 2015).

Threshold values

To determine the threshold values, two strategies are employed; the rolling window approach and one-step-ahead quantile forecasting. We choose the α% quantile level to determine threshold values at 20% and 80%.

Passive strategy

Another strategy is buying and holding the stocks during the out-of-sample period and selling them at the end of the out-of-sample period.

One-step-ahead Value at Risk forecasting

In this section, after fitting ST-GARCH with second-order logistic function model to stock return spread and employing MCMC method for estimating the parameters, we forecast one-step-ahead VaR using the following formula:

$$VaR = -[\mu_{n+1}^{(j)} + t_\alpha\left(\nu^{(j)}\right)\frac{\sqrt{h_{n+1}^{(j)}}}{\sqrt{\nu^{(j)}/(\nu^{(j)}-2)}}] \quad (8)$$

Where, $t_\alpha(\nu^{(j)})$ is the αth quantile of a Student's t-distribution with $\nu^{(j)}$ degrees of freedom, $\sqrt{\nu^{(j)}/(\nu^{(j)}-2)}$ is an adjustment term for a standardized Student's t-distribution with $\nu^{(j)}$ degrees of freedom, $\mu_{n+1}^{(j)}$ is the conditional mean, $h_{n+1}^{(j)}$ is the conditional volatility in the above formula, $j = M+1, \dots, N$

are evaluated upon $y^{s+1,n}$ and the parameter values at MCMC iteration j. The final VaR estimate is the average of this sample (Cathy W. S. Chen, Monica M. C. Weng, Toshiaki Watanabe , 2017).

Empirical results

Pair selection

The stocks which are used in this research are detailed in Table 1.

Table 1

The descriptive statistics of 36 stock returns during the in-sample period.

Code	Company name	Daily return average	Standard deviation	Minimum daily return	Maximum daily return	Jarque Bera test
FOLD1	Mobarakeh Steel	0.107%	0.017	-16.15%	11.61%	<0.001
GOLG1	Gol-E-Gohar	0.089%	0.015	-10.87%	9.44%	<0.001
KAVR1	Iran Khodro D.	0.057%	0.022	-12.86%	10.24%	<0.001
PKOD1	Pars Khodro	0.089%	0.019	-10.53%	7.51%	<0.001
IKCO1	Iran Khodro	0.087%	0.020	-11.41%	11.73%	<0.001
BAHN1	Bahonar Copper	0.058%	0.021	-5.07%	11.10%	<0.001
KRTI1	Iran Carton	0.065%	0.018	-5.10%	11.55%	<0.001
MAPN1	MAPNA	0.116%	0.018	-9.89%	11.52%	<0.001
NAFT1	Oil Ind. Inv.	0.041%	0.018	-10.41%	5.27%	<0.001
AZAB1	Azarab Ind.	0.105%	0.025	-9.28%	7.41%	<0.001
BARZ1	Kerman Tire	0.053%	0.010	-4.29%	10.83%	<0.001
KPRS1	Pars Tile	0.029%	0.012	-8.65%	5.98%	<0.001
RSAP1	Rayan Saipa	0.150%	0.020	-6.92%	7.69%	<0.001
LKGH1	Ghadir Kh. L.	0.071%	0.021	-13.01%	9.95%	<0.001
NOVN1	EN Bank	0.045%	0.013	-12.11%	9.93%	<0.001
BPAR1	Parsian Bank	0.073%	0.015	-4.97%	5.02%	<0.001
MSMI1	I. N. C. Ind.	0.102%	0.016	-5.40%	12.79%	<0.001
TAIR1	Iran Tire	0.061%	0.019	-12.33%	4.88%	<0.001
ARDK1	Ardekan Ceramic	0.188%	0.020	-4.08%	3.92%	<0.001
BIME1	Insurance	0.020%	0.019	-12.37%	4.84%	<0.001

Code	Company name	Daily return average	Standard deviation	Minimum daily return	Maximum daily return	Jarque Bera test
	Inv.					
CHML1	Chadormalu	0.073%	0.015	-9.44%	8.76%	<0.001
GGAZ1	Ghazvin Sugar	0.123%	0.021	-4.64%	14.38%	<0.001
MELT1	Melat Inv.	-0.019%	0.020	-5.07%	10.99%	<0.001
MKBT1	Iran Tele. Co.	0.067%	0.014	-11.49%	11.12%	<0.001
NALM1	Aluminum R.	0.057%	0.022	-15.50%	29.98%	<0.001
PETR1	Petro. Inv.	0.021%	0.021	-7.93%	11.22%	<0.001
PNES1	Isf. Oil Ref.	0.081%	0.016	-5.40%	11.60%	<0.001
SSAP1	Saipa Inv.	0.066%	0.020	-4.49%	4.88%	<0.001
SSHR1	Shargh Cement	0.011%	0.016	-4.87%	4.86%	<0.001
TAMI1	Sand Foundry	0.066%	0.017	-5.10%	11.58%	<0.001
TKIN1	Techinco	0.090%	0.023	-7.18%	4.79%	<0.001
ZMYD1	Zamyad	0.078%	0.022	-6.07%	4.88%	<0.001
ATDM1	Atye Damavand	0.064%	0.018	-4.99%	11.45%	<0.001
GNBO1	Neyshabour S.	0.085%	0.024	-5.06%	13.31%	<0.001
PAKS1	Paxan	0.073%	0.015	-9.00%	4.87%	<0.001
NBEH1	Behran Oil	0.095%	0.014	-15.53%	11.37%	<0.001

The normality assumption is turned down in all cases at the 1% significance level by the Jarque-Bera normality test.

After computing the minimum square distance between each pair of 36 stocks, 10 pairs with minimum MSD are chosen. Five other pairs are selected randomly from one industry. The selected pairs are shown in Table2.

Table 2
List of 15 selected pairs

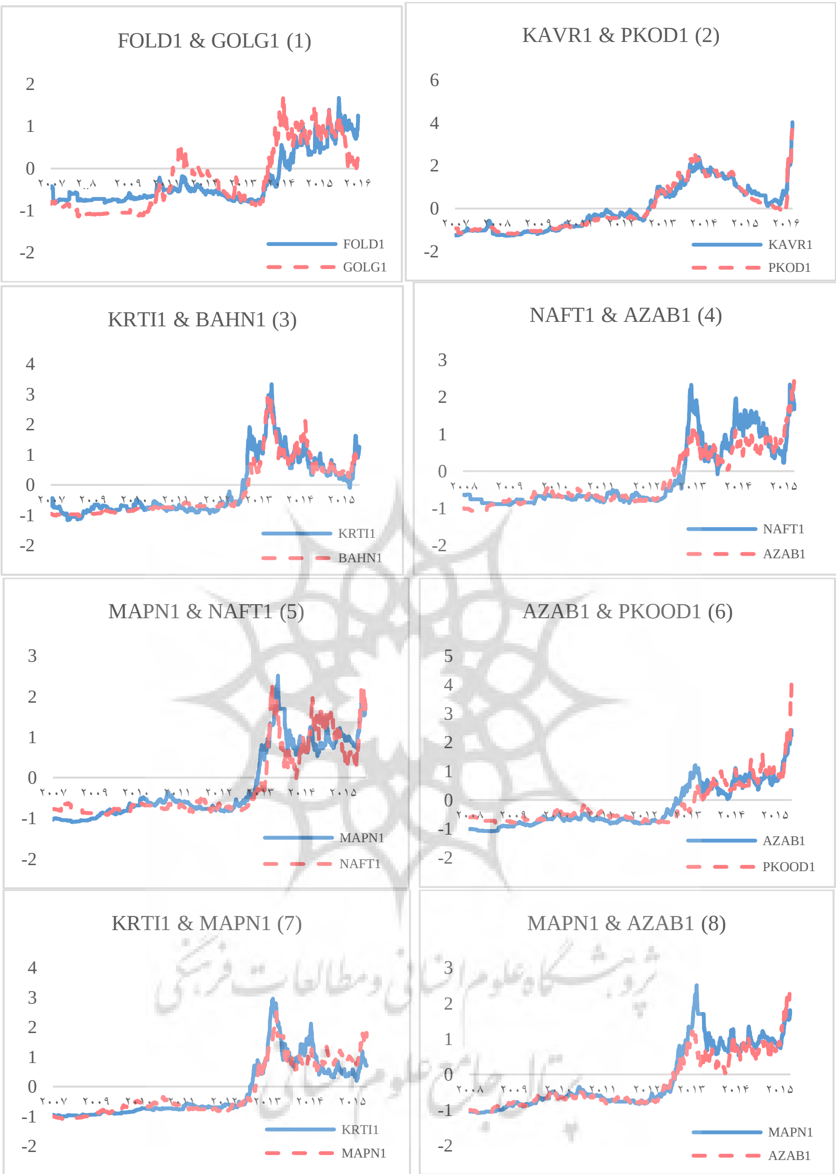
Pairs	Code A	Company name A	Code B	Company name B	MSD	Same industry
1	FOLD1	Mobarakeh Steel	GOLG1	Gol-E-Gohar	99.384	
2	KAVR1	Iran Khodro D.	PKOD1	Pars Khodro	171.254	
3	KRTI1	Iran Carton	BAHN1	Bahonar Copper	224.413	
4	NAFT1	Oil Ind. Inv.	AZAB1	Azarab Ind.	230.421	
5	NAFT2	Oil Ind. Inv.	MAPN1	MAPNA	236.850	
6	AZAB1	Azarab Ind.	PKOD1	Pars Khodro	244.554	
7	KRTI1	Iran Carton	MAPN1	MAPNA	250.437	
8	MAPN1	MAPNA	AZAB1	Azarab Ind.	265.729	
9	KPRS1	Pars Tile	BARZ1	Kerman Tire	272.874	
10	KPRS2	Pars Tile	KRTI1	Iran Carton	273.099	
11	PKOD1	Pars Khodro	IKCO1	Iran Khodro		*
12	BARZ1	Kerman Tire	TAIR1	Iran Tire		*
13	RSAP1	Rayan Saipa	LKGH1	Ghadir Kh. L.		*
14	NOVN1	EN Bank	BPAR1	Parsian Bank		*
15	FOLD1	Mobarakeh Steel	MSMI1	I. N. C. Ind.		*

Note that pairs 1 to 10 are selected by using the MSD method and pairs 11 to 15 are selected each from one industrial sector.

Table 3
The descriptive statistics of the return spreads of 15 pairs during the in-sample period.

Pairs	Company A	Company B	Return spread average	Standard deviation	Minimum	Maximum	Jarque Bera test
1	FOLD1	GOLG1	0.017%	1.866%	-16.146%	12.393%	<0.001
2	KAVR1	PKOD1	-0.037%	2.137%	-13.026%	13.974%	<0.001
3	KRTI1	BAHN1	0.011%	2.630%	-12.769%	11.790%	<0.001
4	NAFT1	AZAB1	-0.075%	2.668%	-14.028%	9.277%	<0.001
5	NAFT2	MAPN1	0.093%	2.379%	-9.981%	11.525%	<0.001
6	AZAB1	PKOD1	0.035%	2.688%	-8.604%	14.449%	<0.001
7	KRTI1	MAPN1	-0.046%	2.426%	-11.525%	11.550%	<0.001
8	MAPN1	AZAB1	0.005%	2.535%	-10.650%	11.525%	<0.001
9	KPRS1	BARZ1	-0.028%	1.535%	-10.832%	7.729%	<0.001
10	KPRS2	KRTI1	-0.054%	2.123%	-14.546%	9.932%	<0.001
11	PKOD1	IKCO1	-0.008%	2.100%	-11.875%	10.425%	<0.001
12	BARZ1	TAIR1	-0.025%	2.071%	-8.471%	12.441%	<0.001
13	RSAP1	LKGH1	0.069%	2.254%	-11.612%	13.618%	<0.001
14	NOVN1	BPAR1	-0.012%	1.797%	-11.340%	9.812%	<0.001
15	FOLD1	MSMI1	0.016%	1.847%	-15.553%	12.255%	<0.001

In Figure 1 we can see the pair normalized price movements in comparison to each other in the in-sample period. It can be seen that in the plots showing pairs which are selected with the MSD method, pair trends are close to each other, but in the pairs which are selected from different industrial sectors, pair trends are not similar to each other. Pairs of RSAP1 & LIKGH1 (13) and BARZ1 & TAIR1 (12) do not show analogous trends as much as the other pairs.


FOLD1 & GOLG1 (1)
FOLD1
GOLG1
KAVR1 & PKOD1 (2)
KAVR1
PKOD1
KRTI1 & BAHN1 (3)
KRTI1
BAHN1
NAFT1 & AZAB1 (4)
NAFT1
AZAB1
MAPN1 & NAFT1 (5)
MAPN1
NAFT1
AZAB1 & PKOOD1 (6)
AZAB1
PKOOD1
KRTI1 & MAPN1 (7)
KRTI1
MAPN1
MAPN1 & AZAB1 (8)
MAPN1
AZAB1

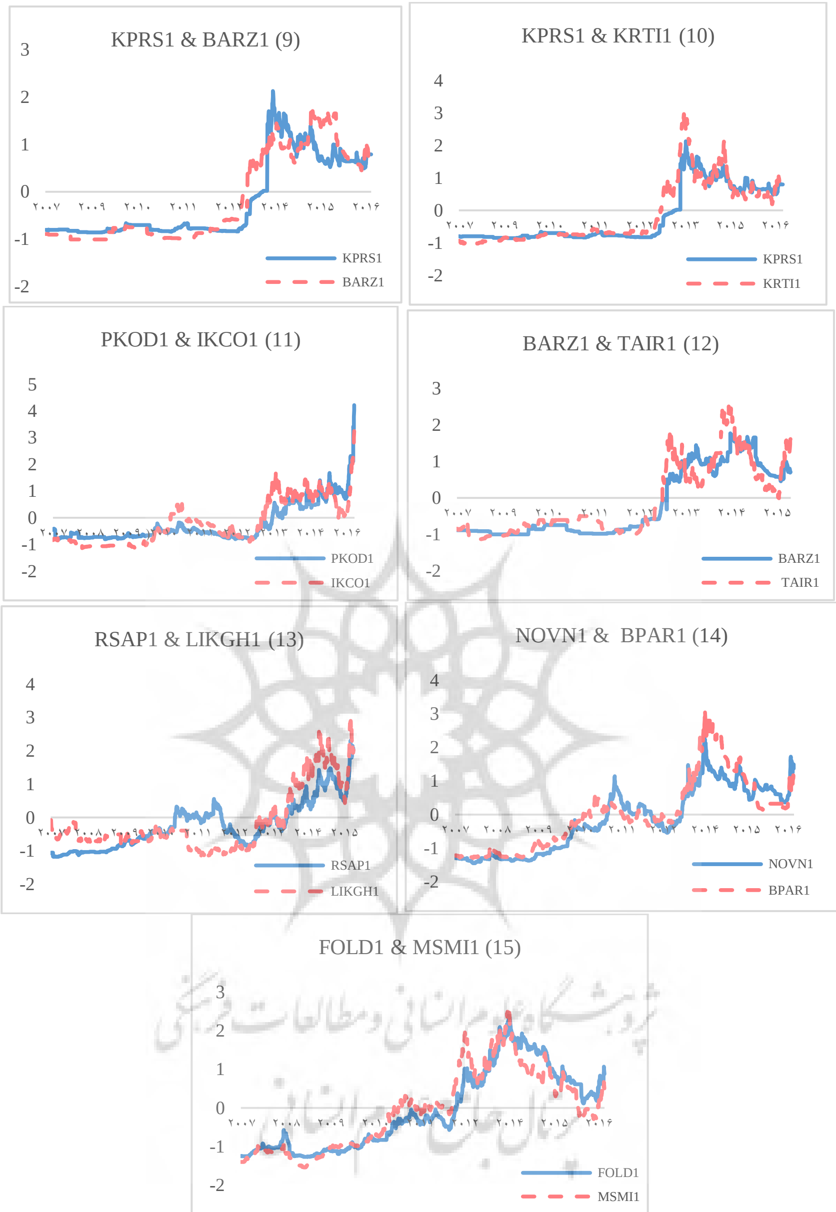


Fig1. Time series plot of normalized prices for selected pairs in the in-sample period**.**

Implementing ST-GARCH model with second-order logistic function and estimating the parameters with MCMC model

First, we implement smooth transition GARCH model with the second-order logistic function on the return spread for the in-sample period using equations 4 and 5. For optimizing the parameters, we consider a sample size of 2,000 for simulating the data using MCMC model. In order to implement this model, we use Metropolis-Hastings method and consider the burn-in sample of 1,000 and a total sample of 50,000 iterations. We only take the second half of each iteration (Martin, A. D., Quinn, K. M., & Park, J. H., 2011).

After implementing ST-GARCH model on the data, we estimate the upper and lower bounds as entry and exit signals for the out-of-sample period. We employ two strategies to produce the thresholds; first, the rolling window approach and second, one-step-ahead quantile forecasting. The results of implementing strategies 1 & 2 are shown in Table 4 and 5.

Table 4

The pair trade annually average profits based on strategy 1 in the out-of-sample period. We employ rolling window approach to produce estimated thresholds.

Pairs	Company A	Company B	Round-trip trades	pair return	Asset value
1	FOLD1	GOLG1	73	5.16%	1,044,714.06
2	KAVR1	PKOD1	104	-23.92%	356,135.61
3	KRTI1	BAHN1	1	35.93%	1,446,762.74
4	NAFT1	AZAB1	92	-15.76%	1,456,715.34
5	NAFT2	MAPN1	107	-8.48%	1,861,657.40
6	AZAB1	PKOD1	92	-17.12%	976,576.49
7	KRTI1	MAPN1	106	-10.74%	2,529,331.61
8	MAPN1	AZAB1	160	-13.88%	1,599,513.98
9	KPRS1	BARZ1	65	24.61%	3,349,386.42
10	KPRS2	KRTI1	11	23.41%	3,266,882.07
11	PKOD1	IKCO1	98	9.56%	839,681.40
12	BARZ1	TAIR1	2	9.46%	757,290.31
13	RSAP1	LKGH1	157	-5.08%	553,103.69
14	NOVN1	BPAR1	163	-5.39%	889,178.63
15	FOLD1	MSMI1	0	8.94%	836,920.00

The transaction cost for selling and buying each stock is respectively 0.975 and 0.464 percent of the stock price.

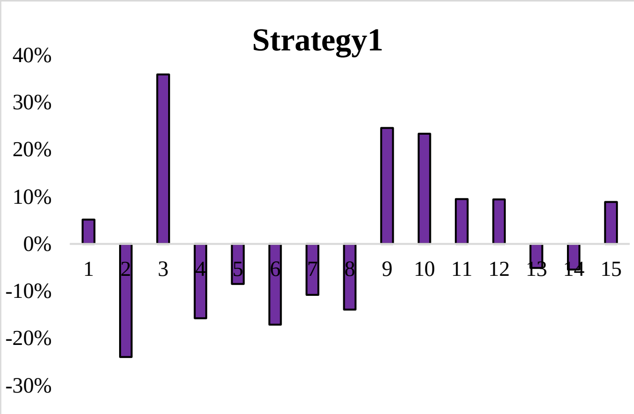


Fig2. Returns for 15 pair trades using strategy 1 in the out-of-sample period.

Table 5
The pair trade annually average profits based on strategy 2 in the out-of-sample period. We employ a one-step-ahead quantile forecasting to produce estimated thresholds. For lower and upper thresholds, we use 20% and 80% quantiles.

Pairs	Company A	Company B	Round-trip trades	pair return	Asset value
1	FOLD1	GOLG1	111	8.18%	978,438.37
2	KAVR1	PKOD1	104	-20.33%	399,817.10
3	KRTI1	BAHN1	140	22.90%	1,105,108.57
4	NAFT1	AZAB1	72	-15.92%	2,377,878.27
5	NAFT2	MAPN1	73	-8.50%	2,085,294.11
6	AZAB1	PKOD1	95	-12.38%	942,025.41
7	KRTI1	MAPN1	47	-15.08%	,121,658.19
8	MAPN1	AZAB1	66	27.94%	,631,728.08
9	KPRS1	BARZ1	178	24.91%	,467,456.46
10	KPRS2	KRTI1	151	23.57%	,388,354.17
11	PKOD1	IKCO1	92	9.43%	56,127.99
12	BARZ1	TAIR1	64	9.45%	1,717,244.21
13	RSAP1	LKGH1	96	-8.78%	690,244.85
14	NOVN1	BPAR1	47	-5.72%	675,760.77
15	FOLD1	MSMI1	103	8.92%	841,124.49

The transaction cost for selling and buying each stock is respectively 0.975 and 0.464 percent of the stock price.

Fig3. Returns for 15 pair trades by using strategy 2 in the out-of-sample period.

Implementing passive strategy

We implemented another strategy to examine the effectiveness of our research. We assume that the investor buys the stocks which exist in each pair at the beginning of the out-of-sample period, and sells them at the end of that period. We then calculate the profits and asset values to compare them with the results from previous strategies.

Table 6

Annually average returns of the analogous pairs in the passive strategy in the out-of-sample period.

Pairs	Company A	Company B	pair return
1	FOLD1	GOLG1	13.27%
2	KAVR1	PKOD1	-51.71%
3	KRTI1	BAHN1	53.82%
4	NAFT1	AZAB1	-27.54%
5	NAFT1	MAPN1	-18.75%
6	AZAB1	PKOD1	-40.02%
7	KRTI1	MAPN1	6.24%
8	MAPN1	AZAB1	-29.30%
9	KPRS1	BARZ1	66.01%
10	KPRS2	KRTI1	62.37%
11	PKOD1	IKCO1	-41.68%
12	BARZ1	TAIR1	24.85%
13	RSAP1	LKGH1	-16.33%
14	NOVN1	BPAR1	-23.67%

15	FOLD1	MSMI1	15.07%

Comparing the result of different strategies

We consider we have the analogous portfolios with the exact same pairs in the pair trading strategies. The results of each strategy are presented in Table 7.

Table 7

Annually returns and asset values are resulted from investing with different strategies.

	Strategy1	**Strategy2**	**Strategy3 (Passive strategy)**
Portfolio return	16.72%	48.59%	2.92%
Asset value	21,763,849.76	24,173,152.49	21,936,933.88

Paired different test

In order to compare population means in strategies 1 and 2, we use paired difference test. The null hypothesis is that there is no significant difference between the two population means. The alternative hypothesis is that significant difference exists between population means, so employing the two strategies will have different results. Table 8 shows the paired difference test results.

Table 8
Results of the paired difference test

Pairs	Company A	Company B	Strategy1 vs. Strategy2	
			Probability	Value
1	FOLD1	GOLG1	19.01%	-1.314
2	KAVR1	PKOD1	64.00%	-0.468
3	KRTI1	BAHN1	2.60%	2.240
4	NAFT1	AZAB1	78.41%	0.274
5	NAFT2	MAPN1	93.41%	0.083
6	AZAB1	PKOD1	35.18%	-0.933
7	KRTI1	MAPN1	48.53%	0.699
8	MAPN1	AZAB1	0.39%	-2.912
9	KPRS1	BARZ1	24.91%	-1.155
10	KPRS2	KRTI1	32.01%	0.996
11	PKOD1	IKCO1	6.63%	1.845
12	BARZ1	TAIR1	92.97%	0.088
13	RSAP1	LKGH1	28.12%	1.080
14	NOVN1	BPAR1	69.51%	0.392
15	FOLD1	MSMI1	0.00%	-55.423

As seen in 4.1. pair trading procedure, in comparing strategies 1 and 2, pairs of KRTI1 & BAHN1, MAPN1 & AZAB1, and FOLD1 & MSMI1 have significantly different means at 5% level.

Optimization Strategy for each pair

The results of the optimized strategy are shown in Table 9. It can be seen that in most cases, strategy 3 (buy and hold) has more return, but considering the return of the whole portfolio, this strategy gives less return than strategies 1 & 2. The reason for this observation is that pair trading is a market-neutral strategy and can be used in the risky market. In other words, the passive strategy is too sensitive to the movement of the market; so if the market is negative, it produces more negative returns than the other two strategies and vice versa.

Table 9
Optimized strategy for each pair

Pairs	Company A	Company B	Strategy1	Strategy2	Strategy3
1	FOLD1	GOLG1			*
2	KAVR1	PKOD1		*	
3	KRTI1	BAHN1			*
4	NAFT1	AZAB1	*		
5	NAFT2	MAPN1	*		
6	AZAB1	PKOD1		*	
7	KRTI1	MAPN1			*
8	MAPN1	AZAB1		*	
9	KPRS1	BARZ1			*
10	KPRS2	KRTI1			*
11	PKOD1	IKCO1	*		
12	BARZ1	TAIR1			*
13	RSAP1	LKGH1	*		
14	NOVN1	BPAR1	*		
15	FOLD1	MSMI1			*

Concluding remarks

Pair trading is a mean-reverting strategy that assumes prices move to their historical trend. In this research, after calculating Minimum Square Distance between all normalized prices of pairs, 10 pairs with the lowest MSD and 5 pairs each in a different industrial sector were selected. Return spreads between selected pairs were calculated, and smooth transition GARCH model with second-order logistic function was fitted to the return spreads. We employed MCMC model for simulating data and estimating the parameters. Consequently, upper and lower bounds were estimated as entry and exit signals. For producing thresholds, we implemented two strategies: first, the rolling window approach and second, one-step-ahead quantile forecasting.

In order to be able to compare the results, another strategy was employed in the out-of-sample period, namely the Passive Strategy. We then compared returns of each pair and return of the whole portfolio in different strategies. Finally, we employed paired difference test between each pair at the 5% level.

In this research, we showed that employing pair trading strategy in case of the second strategy produce more profit in the whole portfolio. In comparing pair trading with passive strategy as pair trading was a market-neutral strategy and could be used in the risky market, it adjusted the return. So we saw despite that the passive strategy return is too much or little, we adjusted return in the pair trading method.

Since the models which are used in this research include a broad class of financial time series models, properly capturing the characteristics of assets liquidity, and also employ the quantile forecasting, these models can be used in practice. While we restricted our attention on stock prices in this study, the proposed method can be applied to other assets and commodities. If this method is applied to a greater number of stocks and in smaller periods (because the return trends in Iran's market have been highly volatile in the past years), the results may be different. So, future researches can focus on implementing the model on a broader set of stocks and using threshold models that do not show asymmetric responses to positive and negative shocks.

References

Anderson, H.M, Nam, K., &Vahid, F. (1999). Asymmetric nonlinear smooth transition GARCH models. *Nonlinear Time Series Analysis of Economic and Financial Data*, 191-207.

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

Cathy W. S. Chen, Monica M. C. Weng, Toshiaki Watanabe . (2017). Bayesian forecasting of Value-at-Risk based on variant smooth transition heteroskedastic models. *Statistics and Its Interface*, 451 – 470.

Cathy W.S. Chen, Zona Wang, Songsak Sriboonchitta, Sangyeol Lee. (2016). Pair trading based on quantile forecasting of smooth transition GARCH models. *North American Journal of Economics and Finance*, 38–55.

Chan, K. , & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 178–190.

Chen, C. W. (2006). On a threshold heteroscedastic model. *International Journal of Forecasting*, 73–89.

Chen, C. W. S., Chen, M., & Chen, S. Y. (2014). Pairs trading via three-regime threshold autoregressive GARCH models. In Huynh et al. (Eds.) Modeling Dependence in Econometrics, Advances in Intelligent Systems and Computing. *Switzerland: Springer International Publishing*, 127–140.

Chivers, C. (2015). *General Markov Chain Monte Carlo for Bayesian Inference using adaptive Metropolis-Hastings sampling.* website: https://goo.gl/s9Db9b.

Elliott, R. J., van der Hoek, J., & Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, 271–276.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica*, 987–1008.

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

Gerlach, R., & Chen, C. W. S. (2008). Bayesian inference and model comparison for asymmetric smooth transition heteroskedastic models. *Statistics and Computing*, 391–408.

Girma, P. B. (1999). Risk arbitrage opportunities in petroleum futures spreads . *Journal of Futures Markets*, 931–955.

González-Rivera, G. (1998). Smooth-transition GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, 61-78.

HP Mashele,SE Terblanche & JH Venter. (2015). Pairs trading on the Johannesburg Stock Exchange. *Investment Analysts Journal*, 13-26.

Jorion, P. (1997). Value at risk: the new benchmark for controlling market risk. *Irwin Professional Pub*.

Jung, Y. C. (2016). Relative performance of VIXC vs. GARCH in predicting realised volatility change. *Investment Analysts Journal*, S1-S16.

Krauss, C. (2016). Statistical arbitrage pairs trading strategies: review and outlook. *Journal of Economic Surveys*, 513–545.

Liu, J., & Timmermann, A. (2013). Optimal convergence trade strategies. *Review of Financial Studies*, 1048–1086.

Martin, A. D., Quinn, K. M., & Park, J. H. (2011). *MCMCpack: Markov chain monte carlo in R.*

Perlin, M. S. (2009). Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives and Hedge Funds*, 122–136.

Ross, S. (1976). The arbitrage theory of capital asset pricing . *Journal of Economic Theory*, 341–360.

Skiena, S. (n.d.). Pair trading. *Department of Computer Science* (p. Lecture 23). State University of New York: Stony Brook, NY 11794-4400.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 208–218.

Vidyamurthy, G. (2004). *Pairs trading: Quantitative method and analysis.* Hoboken, New Jersey: John Wiley and Sons.