CHAPTER 5

Pairs Trading Rule with Switching Regression GARCH Model

The contents of this chapter are extracted from the original article named "Pairs Trading Rule with Switching Regression GARCH Model", which was published in "Integrated Uncertainty in Knowledge Modelling and Decision Making" Lecture Notes in Computer Science, pp. 586-598. It can be found in Appendix C.

5.1 Introduction

Many investors have taken a well-known strategy, which is pairs trading and it was invented at Morgan Stanley in 1987. Pairs trading is a market-neutral strategy following two-step process: first of all, we need to identify two stocks whose prices showed a strongly co-movement historically, and second, sell the gain and buy the loss when the price relation is broken. The profit can be made and the prices of the two stocks will converge to a mean if the future reflects all information in the past.

There are a great number of different studies within pairs trading framework, such as distance approach, co-integration approach and time series approach. These can be sorted into three main approaches. Firstly, the distance method utilizes nonparametric distance metrics to calculate the sum of squared deviations between two normalized stock prices as the criterion to form pairs trading opportunities. The most cited paper was published by Gatev, Goetzman and Rouwenhorst (2006). Later, Perlin (2009) furthered the analysis to examine the profitability and risk of the pairs trading strategy for Brazilian stock market. Do and Faff (2010) replicated the original methodology of Gatev et al (2006) and by the sample period extension to June 2008. They confirmed pairs trading strategy to be profitable for a long period of time, despite at a decreasing rate. Secondly, Vidyamurthy (2004) developed a co-integration approach. The co-integration approach describes how to figure out co-moving stocks relying on formal co-integration testing. Applying this method to pairs trading is mostly based on Gatev et al (2006) threshold rule. Vidyamurthy (2004) suggested a univariate co-integration approach, which is employed to preselect the potential co-integrated pairs, and to design the trading rule with nonparametric methods, based on statistical information. Miao (2014) fitted the high frequency and dynamic pairs trading to the co-integration approach. For co-integrated assets in a continuous-time economy, Chiu and Wong (2015) originated the optimal pairs trading strategy in a closed-form solution. Thirdly, the time series approach was developed by Elliott, van der Hoek and Malcolm (2005), which utilizes a Kalman filter for estimating a parametric model of the mean-reverting spread, in which the formation period is ignored and the spread is assumed to follow the state space model. This approach focuses on describing mean-reversion of the spread with other time series methods rather than co-integration. Do et al. (2006) criticized and extended the method of Elliott, van der Hoek and Malcolm (2005) into the stochastic residual spread method to improve the former method.

Admittedly, although the methods aforesaid might have proper outcome, there still exists one problem as to the linear model assumption. Obviously, a great number of studies demonstrated that the financial data are likely to perform asymmetry, volatility clustering and amplitude dependence, which present a non-linear behavior. It cannot be neglected that they switch among different regimes in some cases. Therefore, the linear approach to identify pairs trading signals might be wrong (Bock and Mestel (2009)). The threshold model and Markov Switching model are popular non-linear models, proposed in some of recent literature. For instance, Bock and Mestel (2009) developed the regime-switching rule for pairs trading and showed that it can generate a good performance. Figure 5.1 shows that the means of the spreads during 2004-2010 appear to be higher than the spreads during 2010-2016. It appears that the regime switching model would be more appropriate for the spreads than linear model. With regard to threshold model, Chen, et al. (2014) developed a three-regime threshold GARCH (generalized autoregressive conditional heteroskedasticity) model to capture pairs trading signals and use the threshold value as trading entry and exit signals. They mentioned that the three-regime threshold GARCH model can identify the regime shift and produce an adequate pairs trading signal, resulting in a good return from the Dow Jones 30 stocks. In light of the previous researches, employing the non-linear model to identify the trading signal is more appropriate. It is reasonable to expect our data to feature nonlinearly. Therefore, this study would like to extend the Threshold GARCH into Markov Switching Regression GARCH since the threshold models have several restrictions as discussed in Kuan(2002). Kuan (2002) criticized on the following three aspects. First, it is difficult to do the optimization in the non-linear model since it might fail to find the global or the optimal solution in the parameter space. Secondly, the threshold models may not be useful to explain the certain nonlinear patterns in the data. Thirdly, the threshold model allows the parameter to change the regime according to the exogenous changes. Nevertheless, Kuan (2002) suggested that the Markov switching model is more appropriate for interpreting correlated data, which demonstrates different behavior in unusual economic condition. Hence, this study introduces a Markov Switching Regression GARCH as an alternative tool for detecting pairs trading signals.

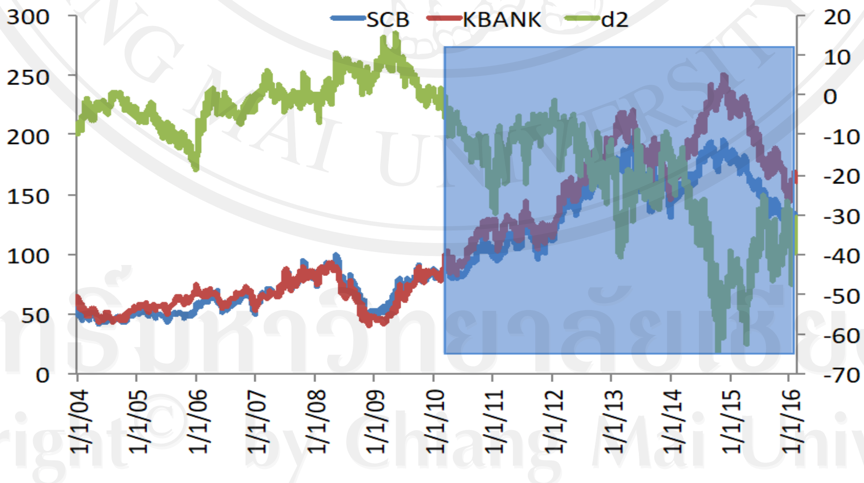

Figure 5.1 Spread of stock prices. The price spreads of SIAM COMMERCIAL BANK (SCB) and KASIKORN BANK (KBANK) are calculated from January 1, 2004 to February 17, 2016. The shaded area indicates the low mean state regime.

5.2 Methodology

The Markov Switching Regression GARCH model is considered to identify the trading signal for stock pair taking into account the structural change in the pair return. In order to select the stock pair, in this thesis, we conduct the Minimum Squared Distance method (MSD) to measure the distance between the two normalized stock price series whereas the first five with the lowest MSD are selected in this empirical study. Note that the lower MSD represents a higher co-movement of the pair returns. The 5 selected pairs are then employed to compute the return spread for the Markov Switching Regression GARCH model. The basic concepts of Minimum Squared Distance method, normalized stock price, the trading rule, and Markov Switching Regression GARCH model are explained in Chapter 2.

5.3 Data

In this chapter, the daily closing prices of 30 stocks in the Stock Exchange of Thailand (SET) SET50 Index are presented and explained. The data are collected from Thomson Reuter data stream, Faculty of Economics, Chiang Mai University. The time period lasts 12 years, from January 1, 2004 to February 17, 2016, covering 3165 observations. We set the in-samples from December 18, 2015 to January 29, 2016. Prior to making the estimation in the Markov Switching Regression GARCH model, this study transforms all the daily data to be log-return and also employ the Augmented Dickey Fuller (ADF) test to check the stationarity of the transformed data. The result shows that all log-returns are stationary at the level. This indicates that the study can further use these log-returns in pair trading analysis.

The in-sample data will be used to produce a simple hedge ratio and then it will be applied to compute a spread return. This study selects 30 companies, namely ADVANCED INFO SERVICE (ADVANC),BANPU (BANPU),BANGKOK BANK (BBL), BANGCHAK PETROLEUM(BCP), BANGKOK DUSIT MED.SVS (BDMS), BUMRUNGRAD HOSPITAL (BH), CENTRAL PLAZA HOTEL (CENTEL), CH KARNCHANG (CK), CHAROEN POKPHAND FOODS (CPF), CENTRAL

PATTANA (CPN), DELTA ELECTRONICS (DELTA), ELECTRICITY GENERATING (EGCO), INTOUCH HOLDINGS (INTUCH), IRPC (IRPC), ITALIAN-THAI DEVELOPMENT (ITD), JASMINE INTERNATIONAL (JAS), KASIKORNBANK (KBANK), KRUNG THAI BANK (KTB), MINOR INTERNATIONAL (MINT), PTT EXPLORATION & PRDN (PTTEP), ROBINSON DEPT.STORE (ROBINS), SIAM COMMERCIAL BANK (SCB), SIAM CEMENT (SCC), SIAM CITY CEMENT (SCCC), TIPCO ASPHALT (TASCO), THANACHART CAPITAL (TCAP), TMB BANK (TMB), TPI POLENE (TPIPL), TRUE CORPORATION (TRUE), THAI UNION FROZEN PRDS (TU). Utilizing the lowest MSD between two normalized stock prices, the five best candidate stock pairs with the lowest MSD are selected for further investigation.

Table 5.1 The descriptive statistics of stock log returns from January 1, 2004 to February 17, 2016

	Mean	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
ADVANC	9.10E-05	0.063669	-0.10146	0.008968	-0.70495	14.86125	18815.55***
BANPU	2.27E-05	0.071928	-0.08082	0.010045	-0.25858	9.942711	6391.809***
BBL	4.61E-05	0.037225	-0.07707	0.007869	-0.64821	11.79778	10428.89***
BCP	7.25E-05	0.071356	-0.07853	0.008813	0.127572	11.21297	8903.933***
BDMS	0.00046	0.09092	-0.05027	0.008448	0.859145	12.56866	12463.75***
BH	0.000371	0.045458	-0.04139	0.008457	0.286056	6.128907	1334.229***
CENTEL	0.000384	0.047861	-0.06652	0.009048	-0.01622	8.265747	3656.781***
CK	-9.73E-06	0.110541	-0.0934	0.012276	0.286937	11.07358	8639.395***
CPF	0.000195	0.050079	-0.0586	0.008273	0.080261	6.63346	1744.416***
CPN	0.000299	0.096035	-0.08864	0.010367	0.042367	11.79175	10194.21***
DELTA	0.000143	0.055715	-0.05552	0.009009	-0.09688	6.467042	1590.139***
EGCO	8.28E-05	0.043225	-0.05451	0.006162	-0.09415	8.706873	4299.633***
INTUCH	4.87E-05	0.084901	-0.11998	0.009059	-0.84254	26.59489	73791.75***
IRPC	-5.79E-05	0.113599	-0.1391	0.012166	-0.58941	24.22751	59607.06***
ITD	-9.74E-05	0.114239	-0.14641	0.014273	-0.09592	12.3205	11461.05***
JAS	0.000136	0.102662	-0.12885	0.014948	-0.1418	11.00746	8466.364***
KBANK	0.000125	0.049892	-0.08852	0.008764	-0.26678	8.60818	4185.232***
KTB	5.63E-05	0.062994	-0.10588	0.009262	-0.42538	11.70244	10082.67***
MINT	0.00037	0.068308	-0.06695	0.01063	0.187745	7.790945	3045.541***
PTTEP	3.15E-05	0.060504	-0.0816	0.009638	0.006979	9.485061	5546.163***
ROBINS	0.000298	0.080823	-0.14613	0.0097	-0.71472	26.01534	70124.44***

Table 5.1 (Continued)

	Mean	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
SCB	0.000119	0.060521	-0.10095	0.008878	-0.33546	12.37629	11653.12***
SCC	6.97E-05	0.051416	-0.05799	0.007367	0.101702	7.755889	2988.268***
SCCC	3.25E-05	0.069254	-0.072	0.008031	-0.17466	10.917	8281.874***
TASCO	0.000214	0.109264	-0.10266	0.011395	0.187132	13.15749	13624.63***
TCAP	0.000108	0.105	-0.10175	0.009315	0.08861	17.56641	27985.41***
TMB	-6.85E-05	0.118845	-0.08388	0.010579	0.442598	14.44391	17374.1***
TPIPL	-0.0001	0.101458	-0.15679	0.012492	-0.32984	17.49664	27771.26***
TRUE	3.91E-05	0.09374	-0.15212	0.015324	-0.33765	11.76253	10185.75***
TU	0.000126	0.053317	-0.04433	0.007688	0.279347	6.631871	1780.658***

Table 5.1 provides the descriptive statistics of all log returns. We can observe that the means of all log returns are close to zero, and this result coincides with the mean reversion theory. In addition, there are 4 companies' stock returns that show a negative mean.

Before illustrating the pairs trading strategy, this study computes the MSD between the two normalized stock price series for all possible stock pairs and it is found that the five pairs trading candidates are the following:

Pair 1: SIAM COMMERCIAL BANK (SCB) vs KASIKORN BANK (KBANK)
Pair 2: CENTRAL PATTANA (CPN) vs CENTRAL PLAZA HOTEL (CENTEL)
Pair 3: INTOUCH HOLDINGS (INTUCH) vs ADVANCED INFO SERVICE (ADVANC)
Pair 4: CENTRAL PLAZA HOTEL (CENTEL) vs BANGKOK DUSIT MED.SVS (BDMS)
Pair 5: CENTRAL PATTANA (CPN) vs BANGKOK DUSIT MED.SVS (BDMS).

Table 5.2 Pair selection

Pair	**Stock 1**	**Stock 2**	**MSD**
1	SCB	KBANK	84.8114
2	CPN	CENTEL	128.0145
3	INTUCH	ADVANCE	164.3700
4	CENTEL	BDMS	197.3506
5	CPN	BDMS	205.1679

And the value of the MSD of each pair is shown in Table 5.2. Then this study fits the Markov switching regression model with GARCH effect to these five selected pair returns. Then, the upper and lower threshold values, which are computed from the standard deviation of return spread of the stock pair, are presented as trading signals. When the return spread exceeds upper threshold, we will sell the stock with a higher price and buy the one with lower price. Similarly, when the return spread drops below the lower threshold, we buy the stock with a higher price and sell the the one with lower price.

5.4 Empirical Results

As mentioned earlier, six different error distributions, , are considered as a conditional volatility assumption in the model. To find the best fit distribution assumption among these six distributions the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used as the criterion and the lowest value is preferred. Moreover, the number of regime, namely one-regime, two-regime, and three- regime, are also selected using the same criterion as distribution selection. Table 5.3 reports an evidence that student-t performs the best fit distribution assumption in the model for all pairs. However, the different two results are obtained for regime selection. It is showed that CPN-CENTEL, CENTEL-BDSM AND BDSM-CPN pairs prefer 2-regime Markov switching regression GARCH(1,1) while 3-regime Markov switching regression GARCH(1,1) is preferred for SCB-KBANK and INTUCH-ADVANCE pairs.

Table 5.3 Model selection

1-Regime					
AIC/BIC	SCB-KBANK	CPN-CENTEL	INTUCH-ADVANCE	CENTEL-BDMS	BDMS-CPN
Normal	-21115.38	-20549.8	-20011.5	-21016.88	-19812.83
	-21085.03	-20549.8	-19975.07	-20986.53	-19776.41
student-t	-21113.38	-20566.98	-21952.04	-21504.72	-20566.9
	-21076.95	-20530.55	-21909.54	-21468.29	-20530.47
skew-T	-21351.44	-20735.94	-21756.88	-21425.48	-20735.44
	-21308.94	-20693.44	-21714.38	-21395.13	-20692.94
Skew-normal	-20450.44	-15819.16	-16276.87	-15057.39	-14737.1
	-20413.97	-15782.73	-16234.38	-15020.96	-14694.6
GED	-19761.44	-19670.74	-20392.5	-20025.82	-19669
	-19725.02	-19640.39	-20356.07	-19995.47	-19638.64
skew GED	-19798.56	-20079.5	-18278.24	-20373	-20079.82
	-19756.07	-20037	-18235.74	-20330.5	-20037.32

Table 5.3 (Continued)

2-Regime					
AIC/BIC	SCB-KBANK	CPN-CENTEL	INTUCH-ADVANCE	CENTEL-BDMS	BDMS-CPN
Normal	-21350.9	-20810.66	-21872.36	-21493.02	-20810.66
	-21278.05	-20737.81	-21799.51	-21420.17	-20737.81
student-t	-21340.32	**-21828.4**	-22746.72	**-22868.84**	**-21228.36**
	-21261.4	**-21743.41**	-22661.73	**-22783.85**	**-21143.37**
skew-T	-21359.24	-20631.48	-22522.7	-21535.06	-20370.98
	-21262.11	-20534.35	-22437.71	-21437.93	-20273.85
skew-normal	-19096.79	-20631.86	-21583.52	-21487.52	-20631.84
	-19017.86	-20546.87	-21498.53	-21390.39	-20546.85
GED	-21396.7	-20798.12	-21831.74	-21437.2	-20798.12
	-21317.78	-20713.13	-21746.75	-21352.21	-20713.13
skew GED	-15029.48	-20627.86	-21583.52	-21487.52	-20627.86
	-14944.49	-20530.73	-21498.53	-21390.39	-20530.73
3-Regime					
AIC/BIC	SCB-KBANK	CPN-CENTEL	INTUCH-ADVANCE	CENTEL-BDMS	BDMS-CPN
Normal	-22313.72	-20601.18	-21599.46	-20981.16	-20592.94
	-22186.23	-20473.69	-21471.97	-20853.67	-20465.45
student-t	**-23951.52**	-21024.4	**-23283.14**	-21748.22	-21003.84
	-23805.82	-20878.7	**-23137.44**	-21602.52	-20858.14
skew-T	-21138.62	-21720.68	-23279.78	-21813.8	-21094.72
	-20974.71	-21556.77	-23115.87	-21649.89	-20930.81
skew-normal	-22294.4	-20558.28	-21502.06	-20993.22	-20603.96
	-22148.7	-20412.58	-21356.36	-20847.52	-20458.26
GED	-22244.98	-20529.62	-21533.18	-20192.46	-20572.98
	-22099.28	-20383.92	-21387.48	-20046.76	-20427.28
skew GED	-20599.96	-19932.18	-21616.8	-20192.46	-20133.58
	-20436.05	-19768.26	-21452.89	-20046.76	-19969.67

Table 5.4 Estimation results of MS-reg-GARCH of pair SCB-KBANK

SCB-KBANK pair with std distribution			
Parameter	Regime 1 $S_t = 1$	Regime 2 $S_t = 2$	Regime 3 $S_t = 3$
φ_0	0.0002	0.0001	0.0002
φ_1	0.4392***	0.0701*	0.6058***
ω	0.0001*	0.0001*	0.0005***
α	0.0001	0.0001	0.0917***
β	0.5436***	0.5436***	0.7896*
v	2.1000***	2.1000***	2.841***
Transition matrix			
	Regime1	Regime2	Regime3
Regime1	p_{11} =0.8108	p_{12} =0.1893	p_{13} =0.0411
Regime2	p_{21} =0.1837	p_{22} = 0.8106	p_{23} =0.0125
Regime3	p_{31} =0.0055	p_{32} =0.0001	p_{33} =0.9464
Log Likelihood	11310.49		
AIC	-23951.52		
BIC	-23805.82		

Table 5.4 shows the estimated results of 3 regimes MS-reg-GARCH(1,1) with student-t distribution for SCB-KBANK pair. The model contains two equations namely, mean equation and variance equation. Consider the mean equation, we interpret φ_0 as the hedge ratio and the result shows that the hedge ratio in regime 3 performs the highest value compared with the others. The value φ_1 in this regime is very close to 1, thus this indicates that SCB-KBANK pair tend to move together in regime 3. Conversely, SCB-KBANK pair seems not to move together in regime 2. Consider the variance equation, it is important concerning the persistence of volatility shocks. To measure this volatility persistence, the study measures it by $\alpha_{s_t=i}+\beta_{s_t=i}$ and the higher value of $\alpha_{s_t=i}+\beta_{s_t=i}$ refers to the higher unconditional variance of the process. The result of variance equation shows that the value of $\alpha_{s_t=i}+\beta_{s_t=i}$ of regime 3 displays the highest persistence of volatility shock, indicating the covariance stationarity with high degree of volatility persistence in this regime. According to this result, we can interpret regime 3 as high volatility regime, while for regime 1 and 2 we interpret as low and moderate volatility regime, respectively. Moreover, this table also shows the result of the transition probabilities obtained from the model and we can observe that all regimes are persistent since the probability of staying in their own regime is larger than 80%, while the probability of switching between these regimes is less than 20%. As we observe in the Figure 5.2, it illustrates filtered probabilities in three regimes. We can observe that regimes 1 and 2 present a similar pattern and the probabilities of staying in those two regimes are not much different, meaning that the stocks in this pair have an equal chance to stay in these two regimes. Consider regime 3, we can observe that there are three periods of time that present a high volatility which are corresponding to the global financial crisis in 2007-2008, European sovereign debt crisis in 2010-2012, and Russian financial crisis in 2014.

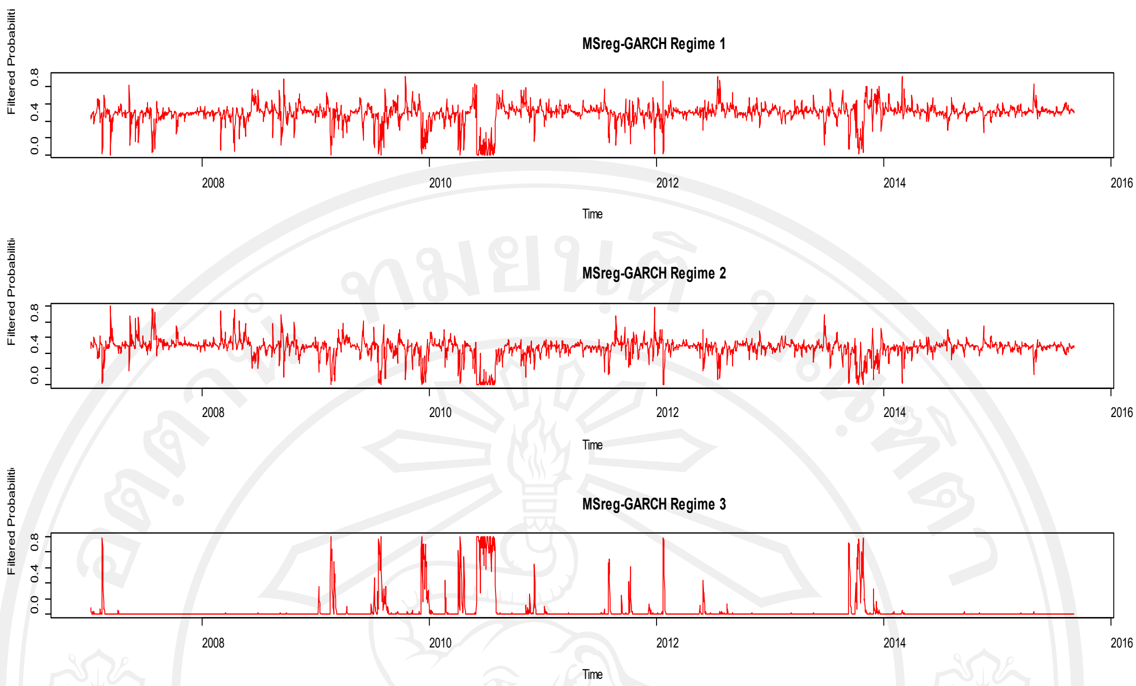

Figure 5.2 SCB-KBANK pair return spread filtered probabilities

Table 5.5 Estimation results of MS-reg-GARCH of pair CPN-CENTEL

CPN-CENTEL pair with std distribution		
Parameter	Regime 1 $S_t = 1$	Regime 2 $S_t = 2$
φ_0	0.0001	0.0001
φ_1	0.1614***	0.0806***
ω	0.0001**	0.0001
α	0.1006***	0.0050
β	0.8001	0.4004***
v	3.0993	2.1142***
Transition matrix		
	Regime1	Regime2
Regime1	p_{11} =0.9000	p_{12} =0.1000
Regime2	p_{21} =0.1000	p_{22} = 0.9000
Log Likelihood	10651.15	
AIC	-21828.4	
BIC	-21743.**41**	

Source: calculation

*, **, and *** are significant at 10%, 5% and 1%, respectively.

Table 5.5 shows the estimated results of 2-regime MS-reg-GARCH(1,1) with student-t distribution for CPN-CENTEL pair. The model also has two equations namely, mean equation and variance equation. Again, in the mean equation, we can interpret φ_1 as the hedge ratio and the result shows that the hedge ratio in regime 1 is relatively larger. Therefore, this indicates that CPN-CENTEL pair has a stronger co-movement in regime 1 than regime 2. Consider the variance equation, the volatility persistence can be measured by the sum $\alpha_{s_t=i} + \beta_{s_t=i}$ and the higher value of $\alpha_{s_t=i} + \beta_{s_t=i}$ indicates the higher unconditional variance of the process. The result of variance equation shows that the value of $\alpha_{s_t=i} + \beta_{s_t=i}$ of regime 1 is close to 1, thus, illustrating the high persistence of volatility shock, and that there is a stationary variance with high degree of volatility persistence in this regime. According to this result, we also interpret regime 1 as high volatility regime, while low volatility regime is the interpretation for regime 2. Moreover, the table also show the result of the transition probabilities and we can observe that both regimes are persistent since the probability of staying in their own regime is equal to 90%, while the probability of switching is less than 10%. As illustrated in Figure 5.3, we can observe a high fluctuation of regime switching along the sample period.

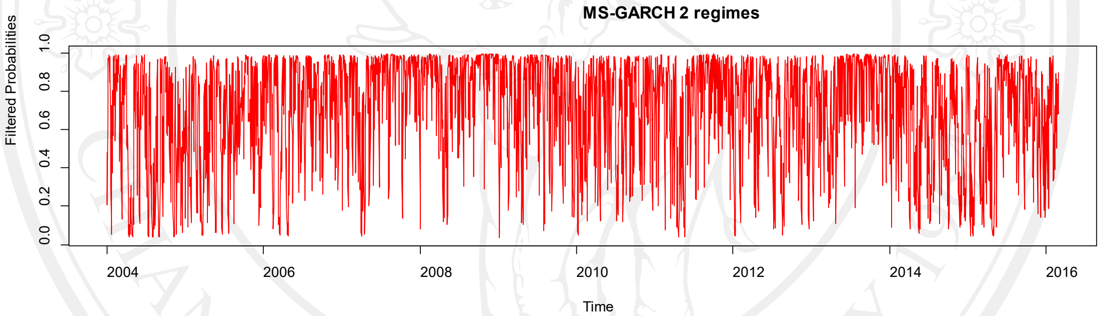

Figure 5.3 CPN-CENTEL pair return spread filtered probabilities

Table 5.6 Estimation results of MS-reg-GARCH of pair INTUCH-ADVANCE

INTUCH-ADVANCE pair with std distribution			
Parameter	Regime 1 $S_t = 1$	Regime 2 $S_t = 2$	Regime 3 $S_t = 3$
φ_0	0.0001	-0.0003	0.0001
φ_1	0.9227***	0.5246*	0.3883***
ω	0.0001*	0.0001*	0.0001***
α	0.2125	0.2584	0.3184***
β	0.6192***	0.3491***	0.3026*
v	5.4264***	2.8354***	2.1613***

Table 5.6 (Continued)

Transition matrix			
	Regime1	Regime2	Regime3
Regime1 Regime2 Regime3	p_{11} =0.6424 p_{21} =0.2543 p_{31} =0.1033	p_{12} =0.1614 p_{22} = 0.8096 p_{32} =0.0290	p_{13} =0.0132 p_{23} =0.0156 p_{33} =0.9712
Log Likelihood AIC BIC	11700.77 -23283.14 -23137.44		

Source: calculation

*, **, and *** are significant at 10%, 5% and 1%, respectively.

Table 5.6 provides the estimated results of 3-regimes MS-reg-GARCH(1,1) with student-t distribution for INTUCH – ADVANCE pair. Consider the mean equation, we interpret φ_0 as the hedge ratio and the result show that the hedge ratio in regime 1 is the largest compared with the others. The value φ_1 in this regime is very close to 1, thus this indicates that INTUCH – ADVANCE pair tends to move together in regime 1. Conversely, INTUCH – ADVANCE pair seems not move together much in regime 3. Consider the variance equation, it is important concerning the persistence of volatility shocks. Generally, the volatility persistence can be measured by the sum $\alpha_{s_t=i}+\beta_{s_t=i}$ and the higher value of $\alpha_{s_t=i}+\beta_{s_t=i}$ represents the higher unconditional variance of the process. The result of variance equation shows that the value of $\alpha_{s_t=i}+\beta_{s_t=i}$ of regime 1 displays the highest persistence of volatility shock, indicating the covariance stationarity with high degree of volatility persistence in this regime. According to this result, we can interpret regime 1 as high volatility regime, while for regimes 3 and 2 we interpret as low and moderate volatility regime, respectively. In addition, the table also provides the result of the transition probabilities and reveals that these three regimes are persistent since the probabilities of staying in their own regime are larger than 60%, 80% and 90% respectively, while the probabilities of switching between these regimes are less than 40%, 20% and 10% respectively. This confirms that only a severe or extreme event will lead to a regime switching. Figure 5.4 also illustrates filtered probabilities in three regimes. We can observe that the regimes 1 and 2 show a similar pattern and the probabilities of staying in those two regimes are not much different, thus indicating that the stocks in this pair have an equal chance to stay in these two regimes. Consider regime 3, we observe that there are two periods of time that present a high volatility which are corresponding to the European sovereign debt crisis in 2012, and Russian financial crisis in 2014.

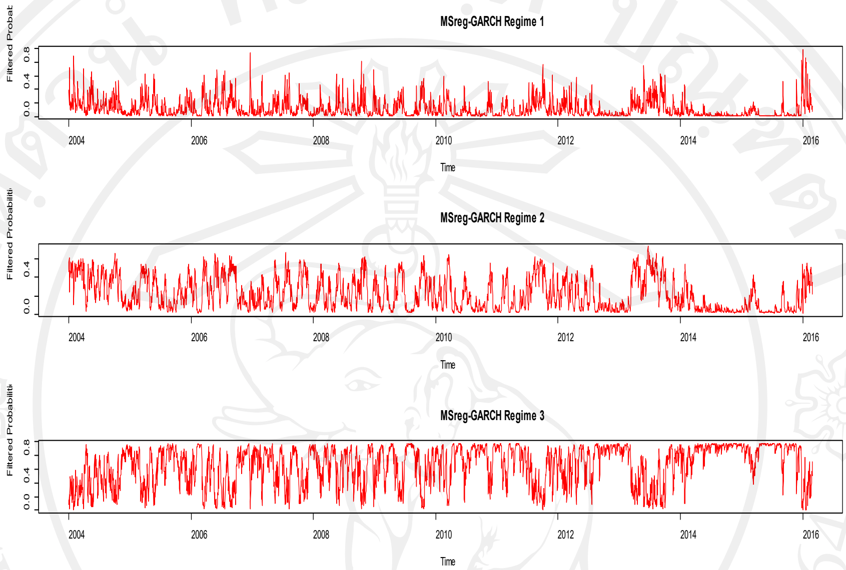

Figure 5.4 INTUCH-ADVANCE pair return spread filtered probabilities

Table 5.7 Estimation results of MS-reg-GARCH of pair CENTEL-BDMS

CPN-BDMS pair with std distribution		
Parameter	Regime 1 $S_t = 1$	Regime 2 $S_t = 2$
φ_0	0.0001	-0.0001
φ_1	0.3434***	0.1716
ω	0.0001**	0.0001
α	0.1007	0.0507
β	0.8142	0.4004***
v	3.0963***	2.1372***
Transition matrix		
	Regime1	Regime2
Regime1	p_{11} =0.9003	p_{12} =0.1002
Regime2	p_{21} =0.0997	p_{22} = 0.8998
Log Likelihood	10509.2	
AIC	-22868.84	
BIC	-22783.85	

Source: calculation

*, **, and *** are significant at 10%, 5% and 1%, respectively.

The 2-regime MS-reg-GARCH(1,1) with student-t distribution for CENTEL-BDMS pair is shown in Table 5.7. Consider the mean equation, here, we interpret φ_1 as the hedge ratio and the result shows that the hedge ratio in regime 1 is larger than in regime 2. Thus this indicates that CENTEL-BDMS pair tends to move together in regime 1 more than regime 2. On other words, CENTEL-BDMS pair seems not to move together in regime 2. Consider the variance equation, it is important concerning the persistence of volatility shocks. Generally, the volatility persistence can be measured by the sum $\alpha_{s_t=i}+\beta_{s_t=i}$ and the higher value of $\alpha_{s_t=i}+\beta_{s_t=i}$ represents the higher unconditional variance of the process. The result of variance equation shows that the value of $\alpha_{s_t=i}+\beta_{s_t=i}$ of regime 1 is close to 1 and displays the high persistence of volatility shock, indicating the covariance stationarity with high degree of volatility persistence in this regime. According to this result, we can interpret regime 1 as high volatility regime, while for regime 2 we interpret as low volatility regime. In addition, the result of the transition probabilities is also shown in Table 5.7. The study finds that both regimes are highly persistent since there is a high probability of staying in their own regime, namely 90%, whereas the probability of switching between both regimes is only 10%. This brings us to confirm that only a severe and extreme event can switch the structural change between regimes. Figure 5.5 illustrates the state process of regime 1. The result shows that CENTEL-BDMS pair has mostly stayed in regime 1 rather than regime 2.

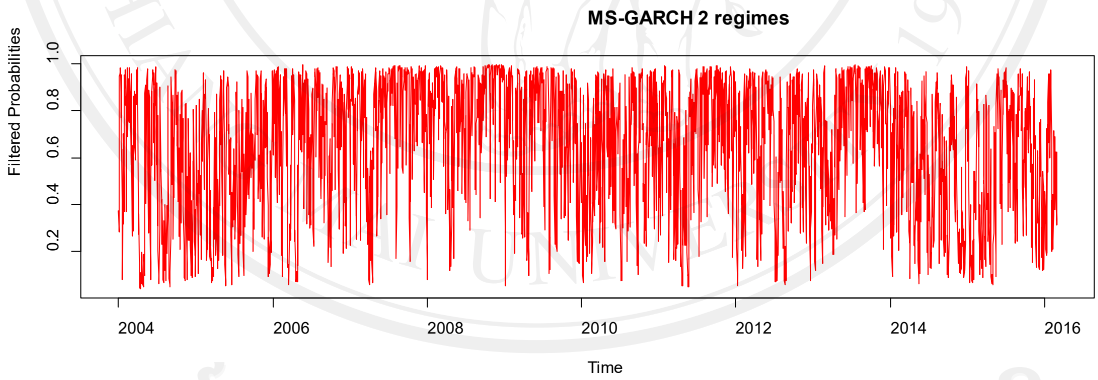

Figure 5.5 CENTEL-BDMS pair return spread filtered probabilities

Table 5.8 Estimation results of MS-reg-GARCH of pair CPN-BDMS

CPN-BDMS pair with std distribution		
Parameter	Regime 1 $S_t = 1$	Regime 2 $S_t = 2$
φ_0	0.0002	0.0001
φ_1	0.2532***	0.1265
ω	0.0001**	0.0001
α	0.1008	0.0054
β	0.7451	0.4005***
v	3.0995***	2.1431***
Transition matrix		
	Regime1	Regime2
Regime1	p_{11} =0.9001	p_{12} =0.1000
Regime2	p_{21} =0.0999	p_{22} = 0.9000
Log Likelihood	10932.7	
AIC	-21228.36	
BIC	-21143.37	

Source: calculation

*, **, and *** are significant at 10%, 5% and 1%, respectively.

The 2-regime MS-reg-GARCH(1,1) with student-t distribution for CENTEL-BDMS pair is shown in Table 5.7. Consider the mean equation, here, we interpret φ_1 as the hedge ratio and the result shows that the hedge ratio in regime 1 is larger than that in regime 2. Thus this indicates that CPN—BDMS pair tends to move together in regime 1 more than regime 2. Conversely, CPN—BDMS pair seems not to move together in regime 2. Consider the variance equation, it is important with respect to the persistence of volatility shocks. Generally, the volatility persistence can be measured by the sum $\alpha_{s_t=i} + \beta_{s_t=i}$ and the higher value of $\alpha_{s_t=i} + \beta_{s_t=i}$ represents the higher unconditional variance of the process. The result of variance equation shows that the value of $\alpha_{s_t=i} + \beta_{s_t=i}$ of regime 1 is close to 1 and displays the high persistence of volatility shock, indicating the covariance stationarity with high degree of volatility persistence in this regime. According to this result, we can interpret regime 1 as high volatility regime, while for regime 2 we interpret as low volatility regime. In addition, the result of the transition probabilities is also shown in Table 5.8. The study finds that both regimes are highly persistent with a high probability of staying in their own regime, namely 90%, whereas the probability of switching between both regimes is only 10%. This bring us to confirm that only a severe and extreme event can switch the structural change between regimes. As we observe in Figure 5.6, it illustrates the state process of regime 1. The result shows that CPN-BDMS pair has mostly stayed in regime 1 rather than regime 2.

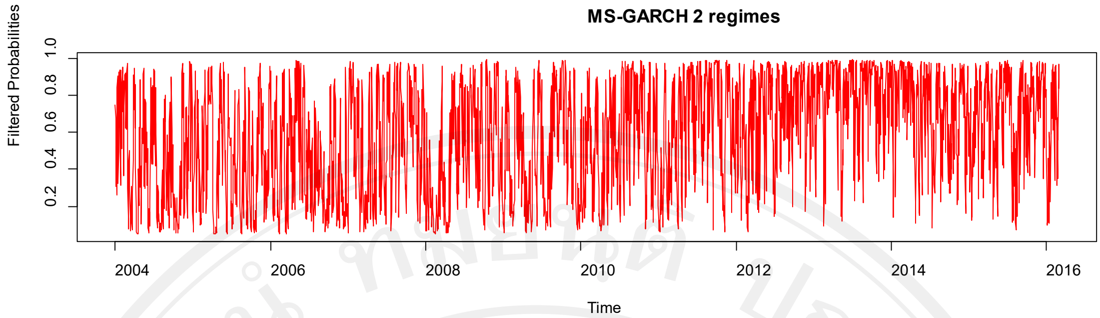

Figure 5.6 CPN-BDMS pair return spread filtered probabilities

We plot a trading signal for our five pair returns in Figures 5.7-5.11. In line with the following Figures 5.7-5.11, the results plot the five pair returns spreads during December 18, 2015 - January 29, 2016, covering 30 trading days. Two blue lines in each figure are interpreted as threshold values or trading lines which are considered as trading entry and exit signals. Whenever the spread goes beyond the upper threshold line, we can sell $stock_t^1$ and buy $stock_t^2$. However, whenever the spread goes from below the lower threshold line, we can buy $stock_t^1$ and sell $stock_t^2$. For example, SCB-KBANK pair return spread, there exist four trading signals for selling and buying $stock_t^1$ and $stock_t^2$, respectively, and the same number of trading signals for buying and selling $stock_t^2$ and $stock_t^1$, respectively.

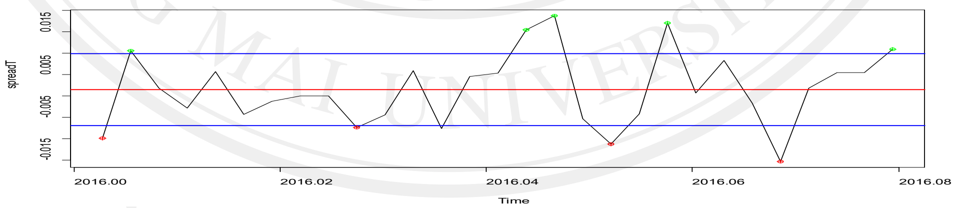

Figure 5.7 SCB-KBANK pair return spread

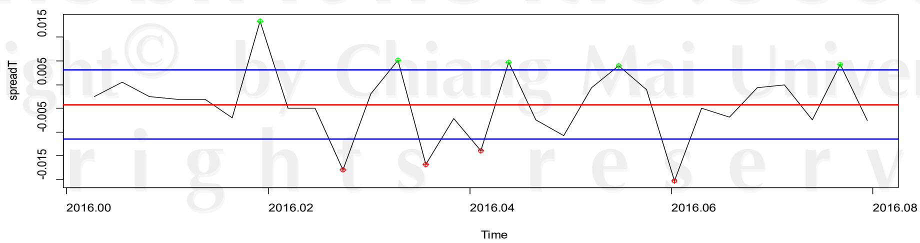

Figure 5.8 CPN-CENTEL pair return spread

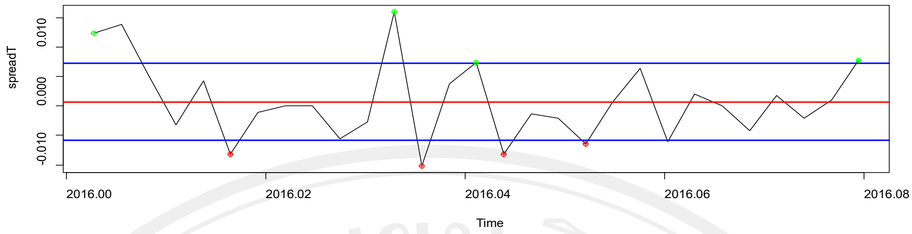

Figure 5.9 INTUCH-ADVANCE pair return spread

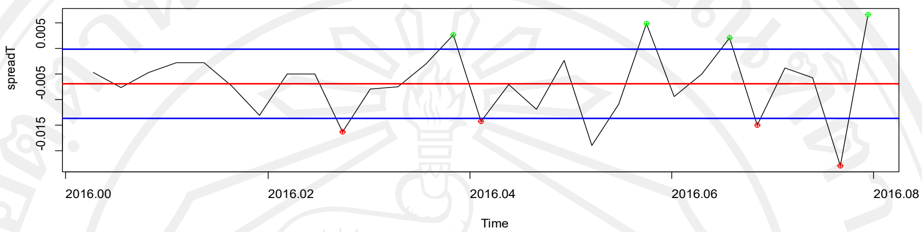

Figure 5.10 CENTEL-BDMS pair return spread

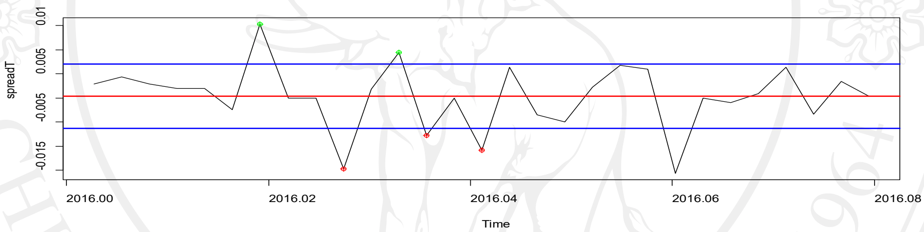

Figure 5.11 CPN-BDMS pair return spread

Table 5.9 Stock returns in five pairs and pair returns from December 12, 2015 to January 29, 2016

Pair	Stock1	Mean return	Stock2	Mean return	No. of trading	Pair return
1	SCB	3.3222%	KBANK	3.5981%	7	9.9553%
2	CPN	-2.6185%	CENTEL	-4.9613%	9	5.4998%
3	INTUCH	2.4359%	ADVANCE	4.3466%	8	16.9454%
4	CENTEL	-4.9613%	BDMS	-2.4022%	9	-2.6121%
5	CPN	-2.6185%	BDMS	-2.4022%	6	1.2678%

According to Table 5.9, we summarize a trading strategy and the return of stock pair as well as individual return during our in-sample period. For the number of trading signals in these five pairs, we can count by looking at the points where value of spread exceeding either upper or lower threshold line. If the spread value exceeds either upper or lower threshold line, we can count it as a trading signal. When it comes to Table 5.9, it is obvious that there are 7.9 round trips trading on average along the in sample period.

Moreover, we can see that our trading signals make a positive return to all pairs, except for CENTEL-BDMS pair. However, the value of loss from negative return of CENTEL-BDMS pair is lower in comparison with single loss from individual stock. We can observe that INTUCH and ADVANCE shows the highest pair return, followed by SCB-KBANK, CPN-CENTEL, and CPN-BDSM. In addition, we also make a comparison between the individual and pair returns. We find that the returns from our pairs trading strategy produce a higher return than the individual trade.

In sum, our pairs trading signal generated from the Markov switching approach works fairly well in this application study. It can generate a higher return when compared with the single return in individual stock.

5.5 Conclusions

In this thesis, a Markov Switching Regression GARCH is introduced to detect a pairs trading strategy. First of all, all possible stock pair combinations are selected by considering the level of correlation. The study uses a Minimum Squared Distance method (MSD) approach to measure the correlation between all possible pairs. The correlation of each pair is measured by two normalized stock prices. The study selects five lowest MSDs for empirical application. Then, the five selected pairs are used to calculate the return spread through the Markov Switching Regression GARCH. We observe that CPN-CENTEL, CENTEL-BDSM and BDSM-CPN are compatible with 2-regime Markov switching regression GARCH(1,1), while 3-regime Markov switching regression GARCH(1,1) is preferred for SCB-KBANK and INTUCH-ADVANCE pairs. To define the trading strategy, the threshold value is computed as a trading signal detector. In this study, the threshold value is computed by the standard deviation of the expected return spread obtained from the Markov Switching Regression GARCH model. In other words, the upper and lower threshold values can be defined as $Uthres = u + sd$ and $Lthres = u - sd$, respectively. Following the trading rule, this study finds that there exist 7.9 round trips trading on average in the 30 trading days from the period December 18, 2015 to January 29, 2016. The average 5 pairs' profit is 6.20% where INTUCH and ADVANCE pair performs the highest return.

In the future research, the study suggests considering the Copula approach with regime switching to determine the pairs trading strategy. Obviously, it would play a big role to capture the marginal distributions and also measure the dependency between the pair stock returns. With a better understanding of the joint distribution of the two stocks, the study expects that investors will enjoy more trading opportunities.