OPTIMAL PRICES AND LOT SIZES FOR EUR/USD CURRENCY TRADING

AFTER NEWS ANNOUNCEMENTS

Mr. Pavarich Suwanpetai

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Financial Engineering

Department of Banking and Finance

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2016

ราคาและจำนวนสัญญาที่ดีที่สุดสำหรับการซื้อขายเงินตราสกุลยูโร/ดอลล่าร์สหรัฐ
หลังจากการประกาศข่าว

นายปวริศ สุวรรณเพทาย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมการเงิน ภาควิชาการธนาคารและการเงิน
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2559

Thesis Title	OPTIMAL PRICES AND LOT SIZES FOR EUR/USD CURRENCY TRADING AFTER NEWS ANNOUNCEMENTS
By	Mr. Pavarich Suwanpetai
Field of Study	Financial Engineering
Thesis Advisor	Assistant Professor Thaisiri Watewai, Ph.D.

Accepted by the Faculty of Commerce and Accountancy, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

............................Dean of the Faculty of Commerce and Accountancy

(Associate Professor Pasu Decharin, Ph.D.)

THESIS COMMITTEE

....Chairman

(Anant Chiarawongse, Ph.D.)

....Thesis Advisor

(Assistant Professor Thaisiri Watewai, Ph.D.)

....Examiner

(Peerapong Dhangwatnotai, Ph.D.)

ปวริศ สุวรรณเพทาย : ราคาและจำนวนสัญญาที่ดีที่สุดสำหรับการซื้อขายเงินตราสกุลยูโร/ดอลล่าร์สหรัฐ หลังจากการประกาศข่าว (OPTIMAL PRICES AND LOT SIZES FOR EUR/USD CURRENCY TRADING AFTER NEWS ANNOUNCEMENTS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร.ไทยศิริ เวทไว, 72 หน้า.

ในวิทยานิพนธ์ฉบับนี้เราหาราคาเปิด-ปิดสัญญา และ จำนวนสัญญาที่ดีที่สุดในการซื้อ-ขาย เงินตราสกุลยูโร/ดอลล่าร์สหรัฐฯ ในระยะเวลา 15 นาทีหลังจากที่มีการเปิดเผยตัวเลขการว่างงานของสหรัฐฯ เราใช้ข้อมูลอัตราแลกเปลี่ยน 10 วินาทีและข้อมูลการประกาศข่าวตั้งแต่เดือนมกราคม พ.ศ 2553 ถึงมกราคม พ.ศ 2557 ในการศึกษานี้ เราวัดพลวัตของราคา เงินตราสกุลยูโร/ดอลล่าร์สหรัฐฯ ด้วยแบบจำลอง 2-Regime mean-reverting jump-diffusion โดยใช้อัลกอริทึมExpectation-Maximization การทดสอบด้วยข้อมูลนอกกลุ่มตัวอย่างแสดงให้เห็นว่า ควรหลีกเลี่ยงการซื้อ-ขายเงินตราในบางสถานการณ์ นอกจากนี้การเลือกจำนวนสัญญาที่เหมาะสมสามารถลดความสูญเสียที่เกิดจากการที่สัญญาซื้อ-ขายถูกบังคับปิดและเพิ่มผลตอบแทนได้ ผลการวิจัยยังชี้ให้เห็นว่าการใช้กลยุทธ์แบบผสมทั้งการซื้อและการขาย สำหรับข่าวแต่ละประเภทสามารถช่วยเพิ่มประสิทธิภาพในการซื้อ-ขายได้

ภาควิชา การธนาคารและการเงิน ลายมือชื่อนิสิต ....
สาขาวิชา วิศวกรรมการเงิน ลายมือชื่อ อ.ที่ปรึกษาหลัก ....
ปีการศึกษา 2559

5682966726 : MAJOR FINANCIAL ENGINEERING
KEYWORDS: OPTIMAL PRICES, OPTIMAL LOT SIZES, EUR/USD CURRENCY, UNEMPLOYMENT CLAIMS ANNOUNCEMENTS

PAVARICH SUWANPETAI: OPTIMAL PRICES AND LOT SIZES FOR EUR/USD CURRENCY TRADING AFTER NEWS ANNOUNCEMENTS. ADVISOR: ASST. PROF. THAISIRI WATEWAI, Ph.D., 72 pp.

In this thesis, we find optimal entry and exit prices and optimal lot size to trade the EUR/USD currency in the 15-minute period after the releases of US unemployment claims numbers. We use 10-second exchange rate data and the news announcement data from January 2010 to January 2014 in this study. We fit the dynamic of the EUR/USD prices to a 2-Regime mean-reverting jump-diffusion model using the Expectation-Maximization algorithm. The out-of-sample tests show that it is better to avoid some trades and choosing an appropriate lot size can reduce losses from force closure and improve returns. The results also suggest that using a combination of long and short strategies for each type of news may improve the trading performance.

Department:	Banking and Finance	Student's Signature	....
Field of Study:	Financial Engineering	Advisor's Signature	....
Academic Year:	2016		

ACKNOWLEDGEMENTS

This thesis accomplishment would never occur without the assistance of many people whose names may not all be mentioned. My problem-solving skill, writing skill, programming skill and analytical skill have been significantly improved from their suggestions. All of their comments, suggestions, and time that they have given to me are necessary for this accomplishment. However, there are some people those I especially want to give them the thankfulness for their assistance.

First and the most important one is my thesis advisor, Asst. Prof. Dr. Thaisiri Watewai. My gratitude, and respect are especially expressed to him for his obligingness and kindness on my work. Learning from his intensive guidance as an expert in this field, I have improved many of my skills. With my respect, I really appreciate his sacrifice for this work and realize that lot of his time was spent on this thesis. As a result, this thesis is finally accomplished with an acceptable standard.

Furthermore, I would like to offer my gratitude to my committee, Dr. Anant Chiarawongse and Dr. Peerapong Dhangwatnotai, for their effective comments to improve my thesis's quality.

Finally, I am forever grateful for my family for their never ending support and encouragement.

CONTENTS

Page

Chapter 1 Introduction

The foreign exchange market is one of the biggest financial markets in the world with a huge trading value from several pairs of currencies. Over one trillion dollars in daily transactions are executed in this market over 24 hours of weekdays.

Traders always have to deal with the dynamic of currency pairs for speculation in the FX market. The evidence that supports the existence of a speculation in the FX market is discovered by (Frankel and Froot 1985) who detect that the speculation in the FX market comes from the use of technical analysis. As there are tons of traders speculating in this market, the currency exchange rates can change rapidly due to speculation from many traders following and contrasting the releases of the important economic news.

Historical data suggest that the currency price may not move in one certain direction immediately after a news release but is likely to fluctuate for a while after the announcement period. The reason could be that there are some big players speculating in the market who try to hedge their positions in their preferred direction after the news announcement. The price can temporarily become highly fluctuating after high-impact economic news relevant to the currency pair is released. This state of the market will be called the transient state in this context. The study from (Allen and Taylor 1990) supports this event that the short-term volatility is driven by a large number of speculators using technical analysis. The rapid change in the exchange rate price is accounted to the momentum of the use of technical analysis in the short term trading by (Schulmeister 1988). The studies from (Froot, Scharfstein et al. 1990) point out that using the technical analysis to speculate in the short period, when new information arrives to the market, is preferred to the traders than using the fundamental analysis.

Also, there is a study from (Goodman 1979) that shows that a prediction result from technical analysis in a short time period is more accurate than using the fundamental; this can support a reason why the technical analysis is preferred. There is a confirmation that using the technical analysis to trade in the FX market is profitable from the study of (Levich and Thomas 1993).

After the transient state, the dynamic of the currency pair will try to stabilize itself into a steady state and the price will follow the news result. With a high variance during a short time period after the news release, traders who open a position may face gain or loss due to their betting direction. For instance, the traders will gain a huge amount if they bet on the right direction. On the other hand, the wrong direction of the bet may force the traders to close their positions due to the use of leverage.

Another stylized fact that can be observed is that the FX bid-ask spreads are usually widening when news is announced. This is another reason that could result in a forced closure due to the insufficient buffered money in their portfolios. The study by (Glosten and Milgrom 1985) is the evidence supporting this fact. They state that the bid-ask spread will immediately widening when news is released and narrowing when the market absorbs the result of an announcement. The studies of (Bollerslev and Melvin 1994) state that the exchange rate volatility increases price risk. The changes in bid-ask spreads occur because the FX dealers try to gain some information advantage by strategically vary the bid-ask price to prevent losses from the uncertain situations as discovered by (Evans and Lyons 2002). With an immediate widening of the spread, the forced close condition can be met before prices begin trending and this can be disastrous for the traders. The violation of the forced closure rule is a huge ruin to the traders' wealth because the position will be forced to close immediately with a market best price without caring how much losses are.

Fortunately, the traders can prepare for the arrival of high-impact news since the announcement time and the level of the impact are scheduled and listed in *www.forexfactory.com*. For example, the traders know from the website in advance that the unemployment claims of USA will be announced on 02 Jan 2014 at 9:30 A.M. with a high impact level to the change of variance.

Figure 1: Bid-Ask movement of EUR/USD after Unemployment claims is announced on 09/01/2014 at 20.30 (+7 GMT)

Figure 1 shows an example of how to make a profit based on an announcement of bad news in the EUR/USD market by placing an appropriate short position after the news is announced (line A) and closing the short position when a 200 points gain is reached (line B).

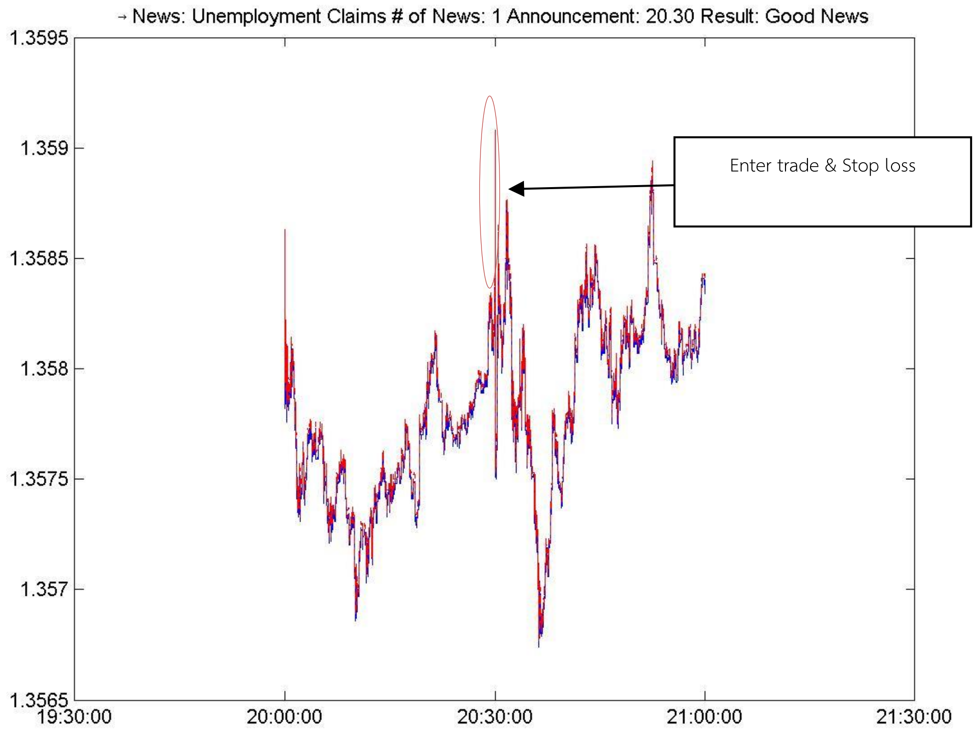

Figure 2: Bid-Ask movement of EUR/USD after Unemployment claims is announced on 30/01/2014 at 20.30 (+7 GMT)

Figure 2 shows the case in which traders enter the trade and immediately exit due to the stop loss triggered by spread widening. The reason is that under an uncertain condition large orders are more likely to raise dealer risk, and so the spread is widened by market makers to prevent the risk caused by the uncertainty (Biais 1993).

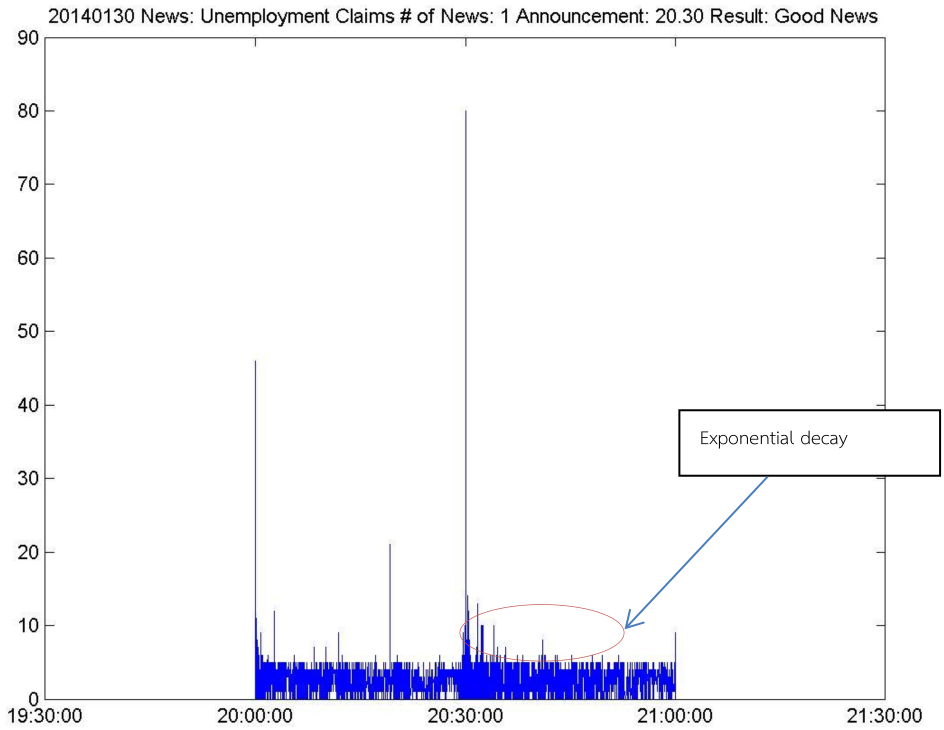

Figure 3: Spread size of EUR/USD after Unemployment Claims is announced on 30/01/2014 at 20.30 (+7 GMT)

Figure 3 shows an example of a spike in the spread after a news release, and the spread is narrowing exponentially fast in time.

Unfortunately, it is impossible to know how much the spreads might spike with a given news release. Trading in news release environments is so dangerous due to many uncertain factors, but the potential rewards could be huge if the traders can find themselves on the right side of the position.

A variety of entry strategies can be helpful for traders to enter the right initial position. However, how the traders take the profit from the matched order by closing the contract at the right price is also important.

In this thesis we will try to improve the strategies for FX trading during news announcement period by using estimation and optimization techniques.

The goal of this thesis is inspired by the view of individual traders who mainly speculate on the release of news. To gain some benefit from this situation, traders may set their market price order of either buy or sell at some optimal price and optimal lot size right after the news is announced.

To find the optimal price to open a trading position with an optimal lot size, traders need to know that if the order is opened far from the current exchange rate level, they can be ensured that their order, if matched, could be a low risk trade. However, there is a high chance that the order is not matched and that gives zero gain to the traders. If the order is matched with a small lot size, this trade may not generate enough gain to compensate the low chance of matching. On the other hand, if traders place the order closer to the current exchange rate level, the traders can be more confident that their order will be matched. However, this trade has a higher risk, so a smaller lot size should be used. Therefore, not only the optimal place of opening the order but also the optimal lot size must be taken into account in this situation.

Intuitively, exchange rate level should go up after the good high-impact news is released and traders can decide to send whether a limit order or a stop order to try to make a profit. Using the limit order is effective in the situation that a reversal on currency price movement occurs after the order is matched. For example, in the situation that the exchange rate is initially moving in the downward direction, the Limit Long order must be placed below the current exchange rate level. After the Limit Long order is matched, the direction of exchange rate is starting to change reversely to the upward direction, and when the price reaches a certain level, the trader can close the position. On the other hand, the stop order is more effective in the situation where the exchange rate is moving in the trending fashion after it passes through some level. For example, in the situation where the exchange rate is initially moving in the upward direction, the Stop Long order must be placed above the current exchange rate level. The Stop Long order is matched when the exchange rate level is passing through where the order is placed. The Stop Long order will be effective if the exchange rate keeps moving in the upward direction, and the trader can close the position when he enjoys enough profit.

In conclusion, after a news release, traders may set up limit orders and expect for the price reversal after orders are matched, or to set up the stop orders in the direction that is consistent with the news and expect the price to move accordingly. However, due to the high volatility during the transient period, the forced closure risk is also high when the lot size is large.

The primary objective of this thesis is to find the optimal entry price, lot size, and exit price for trading EUR/USD during the 15-minute period after the “Unemployment claims” of US is announced. The scope of this study consists of

- Proposing models of exchange rate dynamic of the EUR/USD price and its spread.

- Estimating the parameters of the FX model and the spread model.

- Finding the most suitable strategy to trade in each situation of news announcement. We consider the following strategies: Stop Long order, Stop Short order, Limit Long order and Limit Short order.

- Using an optimization method to find the optimal entry price, target exit price lot size for the chosen strategies.

- Testing each strategy out-of-sample.

Chapter 2 Literature Review

There are many models and techniques we use in this thesis especially in model selection and model estimation. This chapter provides a review of related literature on those models and techniques. Based on the study from (Johnson and Schneeweis 1994) that uses the empirical test to compare many models fitting the movement of 4 currency pairs in FX market when the macroeconomic news is released, they found that in the period that the news had an impact to the market, the dynamics of the currency pairs is better fitted to the complex model with jump diffusion than the simple model without jump.

Their paper uses the weekly data collected once every Wednesday (or Thursday if Wednesday is a holiday) for 4 currencies which are UK pound, German mark, Japanese yen and French franc, all of which are compared with the US dollar. The weekly data cover the period for 20 years starting from January 7, 1976 to January 17, 1996. In contrast, we use 10-second data in our study. However, we expect that a model with jumps should still provide a better fit with more recent data as high-frequency trading has become normal in the FX market. So the response from news releases should be quick and may drive the currency prices to move sharply which is expected to be captured by jumps.

In addition to improve the fit, there is also a suggestion from (Engel and Hamilton 1990) who observe the behavior of the exchange rate using the quarterly FX data to conclude that using the regime switching is a good approximation for FX market.

Instead of using quarterly data, (Marsh 2000) uses the daily exchange rates for three currencies against US dollar and concludes that Markov regime switching model is well fitted to the data although the performance in out-of-sample parameter forecasting is low.

Moreover, (Stephane and Zou 2011) had concluded that the Cox-Ingersoll-Ross regime switching model is better in parameter estimation than non-regime switching model for daily exchange rate data ranging from January 1 ,2000 to October 30 ,2011. They also suggest that using the regime switching model for daily data helps the investors detect some economic events especially when the dynamic of exchange rate is significantly different. With these backgrounds, we may expect that the model with the regime switching may give a better fit than the model without the regime switching. However, we rely on the 10-second data rather than the daily data. So we will consider models with and without regime switching and choose the best model for our study.

To estimate the regime switching model, many parameter values are concealed especially the state variable representing the market condition: the transient state and the steady state. The EM algorithm which is developed by (Dempster, Laird et al. 1977) is an answer for this problem. The EM algorithm is widely used for parameter estimation when the data have missing values. The EM algorithm reduces the difficult task of optimizing the non-separable log-likelihood function by considering a sequence of simpler sub-problems for which the log-likelihood value continuously improves in each step. This algorithm was used for estimating the maximum likelihood estimators in the cases where there are some missing values. The algorithm iterates between an expectation-step (E-step), which constructs an expectation function of log-likelihood using the currently estimated parameters, and a maximization-step (M-Step), which computes the best set of parameters that maximizes the expected complete log-likelihood. The algorithm is terminated when the number of iterations exceeds a limit or the change in the set of parameters is less than a tolerance. There is also the comparison between the EM algorithm with Newton's method by (Springer and Urban 2014). They conclude that EM is slower to converge in terms of the number of iterations but the computation costs for each of iteration is lower. However, the pro of using the EM algorithm is the simplicity for implementation with the consistent and unbiased estimators. In the cons side, EM has a slow linear convergence rate in some cases.

In conclusion, we expect that the model with jump and regime switching may be a great choice for the scope of the thesis and we will develop an EM-based algorithm to estimate the parameter values of our models.

Chapter 3 Data and Model

3.1 EUR/USD data

In this thesis there are two types of data used which are EUR/USD bid-ask quote data and news data.

3.1.1 Tick-data bid-ask quote

EUR/USD bid-ask quote data is published by www.truefx.com. This website gathers it from the real market quote. The data downloaded from this website is a tick-data consisting of bid-ask quote, date and time-stamp for 5 years ranging from January 7, 2010 to March 6, 2014. From the tick data, we construct the 10-second data and use them in our study.

3.1.2 News data

The study from (Galati and Ho 2001) investigates that the releases of macroeconomic news of either the euro area or the United States have impacts on the EUR/USD currency. This thesis considers good and bad unemployment claims news which can be a proxy to indicate the current situation of the macroeconomics of the United States.

Sometimes, there are other types of news announced close to the unemployment claims news. In this thesis we will consider only the period with a single news announcement which can be obviously classified into good and bad news.

Note that: the good news means that the news released may strengthen the EUR currency or weaken the USD currency. Similarly, news will be characterized as bad news if the news released may weaken the EUR currency or strengthen the USD currency. For example, if the actual number of unemployment claims number turns out to be higher than forecast, it may weaken the USD currency. This shows the weakening US economy. Therefore, this news is characterized as good news for EUR/USD currency.

Good news dataset is filtered from the cases in which actual unemployment claim number is higher than the forecast number and there are no other macroeconomic announcements within the range of 2 hours after the news is announced. The news is collected by hand ranging from January 7, 2010 to Jan 30, 2014, which consists of the announcement times, the forecasts of unemployment claim numbers and actual unemployment claim numbers. There are 38 good single news announcements during this period.

Bad news dataset is filtered similarly to the good news cases except that the actual numbers are lower than the forecast numbers in this case. The news is collected by hand ranging from January 7, 2010 to December 26, 2013, which consists of the announcement times, the forecast numbers and the actual numbers. There are 43 bad single news announcements during this period.

Note that unemployment claims news is announced weekly but the time of announcement is not fixed, but is known ahead of time. The time of news announcement is provided by www.forexfactory.com. Good and bad news datasets are shown in Table 22 and Table 23 respectively in Appendix 1.

3.1.3 Daily closing exchange rate for EUR/USD

EUR/USD daily closing rate data is pulled from Bloomberg. The data covers 5 years ranging from January 1, 2010 to December 31, 2014. This daily data is used for volatility adjustments due to different market conditions as discussed later.

3.2 Model

We have two models in our study. The first model is for the exchange rate dynamic, which is assumed to follow a mean-reverting process with regime switching. The other model is for the bid-ask spread process. We assume a random spike in the spread after the news release with an exponential decay.

3.2.1 Mean Reverting with regime switching model

The study of (Johnson and Schneeweis 1994) assumes that the volatility of the currency price is high at the time of news announcement but it immediately goes back to its normal value in the next subsequent periods. To be more realistic, the volatility of the currency price after the news announcement is extremely high and stays high for a certain time period. However, the volatility of the price will move to its fundamental value in the steady state in the longer term. So, traders should not underestimate the volatility because it is related to the entry price and forced close of their position. For example, after the high-impact news is announced, the price dynamic may turn into the transient state which has a high volatility. If we do not adjust the entry price to deal with the market state, we may submit a buy order at a relatively high price, and once it is matched, the volatility and bring the price further down, causing us to close the position due to the forced close rule. Therefore, we expect that the switching of the parameters due to the changing of the market state may give a better result than a pricing model with fixed volatility. This idea supports the use of a regime-switching model.

(Engel and Hamilton 1990) shows that the currency price has mean-reverting effect in long horizon. Moreover, the study from (Johnson and Schneeweis 1994) also suggests that the complex model with jump is preferred than the simple non-jump model. With these evidences we expect that using a mean reverting model with jump for the currency price movement will improve a fit.

Before news is announced, dealers may try to adjust the spread to be wider a little bit for their own benefit by adjusting the bid-ask prices (lower the bid price or increasing the ask price) to be prompt for the arrivals of news to the market. When high-impact news arrives, the spreads width will increase. Once the market becomes clear on the direction of the price movement, the spreads will automatically be adjusted to be narrower and finally go back to the normal state.

The mean-reverting with jump and regime-switching model is an extended form of the mean-reverting with jump process (Clewlow and Strickland 2000).

Consider a mean-reverting with jump process:

$$\frac{dS_t^{(m)}}{S_t^{(m)}} = \kappa\left[\alpha - \ln S_t^{(m)}\right]dt + \sigma dW_t^{(m)} + \theta_t^{(m)} dN_t^{(m)} \quad (1)$$

In this model, $S_t^{(m)}$ is the asset price at time t after the m^{th} news is announced. The process is mean reverting to the long-term log-price level α with the mean reversion rate κ when jumps are ignored. $W_t^{(m)}$ is a standard Brownian motion. The size of jumps for m^{th}news at time t, $\theta_t^{(m)}$,is the extra parameter added to the usual mean-reverting process. We assume that $\theta_t^{(m)}$ is a log-normal distributed random variable where $ln(1+\theta) \sim Nor(\eta, \omega^2)$; η is the mean jump size, $\omega^2$is the variance of the jump and N_t is a Poisson process with rate λ .

In this thesis, we allow all parameter values to be switched between the transient and steady states. The state variable,$y_t^{(m)}$, representing the state of a Markov chain is introduced to make all of the parameter's values state-dependent. The price process is now given by

$$\frac{dS_t^{(m)}}{S_t^{(m)}} = \kappa\left(y_t^{(m)}\right)\left[\alpha\left(y_t^{(m)}\right) - \ln S_t^{(m)}\right]dt + \sigma\left(y_t^{(m)}\right)dW_t^{(m)} + \theta_t^{(m)}\left(y_t^{(m)}\right)dN_t^{(m)} \quad (2)$$

where $y_t^{(m)}$ is assigned to be 1 or 2 for transient or steady state, respectively. Using the transformation $a_t^{(m)} = ln\, S_t^{(m)}$ and Ito's lemma to obtain the return process for the m^{th}news at time t, we have the following process:

$$da_t^{(m)} = \kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - a_t^{(m)}\right]dt + \sigma\left(y_t^{(m)}\right)dW_t^{(m)} + ln(1 + \theta_t^{(m)}\left(y_t^{(m)}\right))\, dN_t^{(m)} \quad (3)$$

where,

$$\bar{\alpha}\left(y_t^{(m)}\right) = \alpha\left(y_t^{(m)}\right) - \frac{\sigma^2\left(y_t^{(m)}\right)}{2\kappa(y_t^{(m)})} \quad (4)$$

In practice continuous measurements of $S_t^{(m)}$ are not possible. We consider the discrete time process $\Delta a_t^{(m)}$ insteads of $da_t^{(m)}$:

$$\Delta a_t^{(m)} = \Delta z_t^{(m)}\left(y_t^{(m)}\right) + \sum_{i=0}^{\Delta N_t^{(m)}} \delta_{t,i}^{(m)}\left(y_t^{(m)}\right) \quad (5)$$

where,

$$\Delta z_t^{(m)}\left(y_t^{(m)}\right) = \kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - a_t^{(m)}\right]\Delta t + \sigma\left(y_t^{(m)}\right)\Delta W_t^{(m)} \quad (6)$$

$$\delta_{t,i}^{(m)}\left(y_t^{(m)}\right) = ln(1+\theta_t^{(m)}\left(y_t^{(m)}\right)) \sim^{iid} Nor\left(\eta\left(y_t^{(m)}\right), \omega^2\left(y_t^{(m)}\right)\right) \tag{7}$$

$\delta_{t,i}^{(m)}\left(y_t^{(m)}\right)$ denote for the size of i^{th}jump that occurs in $(t-1,t]$ of the regime $y_t^{(m)}$.

With the unit time interval $(t-1,t]$, let $x_t^{(m)} = \Delta a_t^{(m)}$. So we can rewrite it as

$$x_t^{(m)} = \Delta z_t^{(m)}\left(y_t^{(m)}\right) + \sum_{i=0}^{\Delta N_t^{(m)}} \delta_{t,i}^{(m)}\left(y_t^{(m)}\right) \tag{8}$$

It is obvious that the discrete-time processes $\Delta z_t^{(m)}$, $\Delta N_t^{(m)}$ and $\delta_{t,i}^{(m)}$ are mutually independent given $y_t^{(m)}$ and the term $\sum_{i=0}^{\Delta N_t^{(m)}} \delta_{t,i}^{(m)}\left(y_t^{(m)}\right)$ can be interpreted as zero when $\Delta N_t^{(m)} = 0$. Now, we may estimate the parameters of process $S_t^{(m)}$ which are $\kappa, \bar{\alpha}, \sigma^2, \eta, \omega^2, \lambda$ by using process $x_t^{(m)}$.

By given the number of jumps in the interval $(t-1,t]$ or $\Delta N_t^{(m)}$, and the hidden state $y_t^{(m)}$, $x_t^{(m)}$ is a Normal random variable with mean $\kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - ln\, S_t^{(m)}\right] + \Delta N_t^{(m)}\eta(y_t^{(m)})$ and variance $\sigma^2(y_t^{(m)}) + \Delta N_t^{(m)}\omega^2(y_t^{(m)})$. Furthermore, we may say that by given hidden state $y_t^{(m)}$, $x_t^{(m)}$ are independently and identically distributed with the density of Poisson-Normal as follows:

$$f\left(x_t^{(m)}|y_t^{(m)}\right) = \sum_{j=0}^{\infty}\left[\phi\left(x_t^{(m)}; \kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - ln\, S_t^{(m)}\right] + j\eta\left(y_t^{(m)}\right), \sigma^2\left(y_t^{(m)}\right) + j\omega^2\left(y_t^{(m)}\right)\right)\right] e^{-\left(\lambda\left(y_t^{(m)}\right)\right)} \frac{\left(\lambda\left(y_t^{(m)}\right)\right)^j}{j!} \tag{9}$$

where $\phi(x\,; a, b)$ denote a Normal density at x with mean a and variance b.

To be able to estimate for the Geometric Brownian motion with jump diffusion model, we define a new parameter $\rho\left(y_t^{(m)}\right) = \kappa(y_t^{(m)})\bar{\alpha}\left(y_t^{(m)}\right)$ and rearrange the model as follows:

$$\begin{aligned}
da_t^{(m)} &= \kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - a_t^{(m)}\right] dt + \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1+\theta_t^{(m)}\left(y_t^{(m)}\right))\, dN_t^{(m)} \\
da_t^{(m)} &= \left[\kappa\left(y_t^{(m)}\right)\bar{\alpha}\left(y_t^{(m)}\right) - \kappa\left(y_t^{(m)}\right) a_t^{(m)}\right] dt + \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1+\theta_t^{(m)}\left(y_t^{(m)}\right)) \\
da_t^{(m)} &= \left[\rho\left(y_t^{(m)}\right) - \kappa\left(y_t^{(m)}\right) a_t^{(m)}\right] dt + \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1+\theta_t^{(m)}\left(y_t^{(m)}\right))
\end{aligned} \tag{10}$$

We can see that by setting the parameter $\kappa\left(y_t^{(m)}\right) = 0$, the model will be changed into the 2-regimes Geometric Brownian motion with jump diffusion model as follows:

$$da_t^{(m)} = \rho\left(y_t^{(m)}\right) dt + \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1+\theta_t^{(m)}\left(y_t^{(m)}\right)) \tag{11}$$

Similarly, we can derive 12 special-case (or nested) models in total which will be considered in this thesis as shown in Table 1.

Table 1: parameters of each model in the scope of the thesis

Model	Number of regimes	π	λ	ω^2	η	σ^2	κ	ρ	q
Pure Diffusion (PD)	1	×	×	×	×	✓	×	×	×
Geometric Brownian Motion (GBM)	1	×	×	×	×	✓	×	✓	×
Mean-Reverting (MR)	1	×	×	×	×	✓	✓	✓	×
Pure Diffusion with Jump (PDJ)	1	×	✓	✓	✓	✓	×	×	×
Geometric Brownian Motion with Jump (GBMJ)	1	×	✓	✓	✓	✓	×	✓	×
Mean-Reverting with Jump (MRJ)	1	×	✓	✓	✓	✓	✓	✓	×
2-Regime Pure Diffusion (2-PD)	2	✓	×	×	×	✓	×	×	✓
2-Regime Geometric Brownian Motion (2-GBM)	2	✓	×	×	×	✓	×	✓	✓
2-Regime Mean-Reverting (2-MR)	2	✓	×	×	×	✓	✓	✓	✓
2-Regime Pure Diffusion with Jump (2-PDJ)	2	✓	✓	✓	✓	✓	×	×	✓
2-Regime Geometric Brownian Motion with Jump (2-GBMJ)	2	✓	✓	✓	✓	✓	×	✓	✓
2-Regime Mean-Reverting with Jump (2-MRJ)	2	✓	✓	✓	✓	✓	✓	✓	✓

3.2.2 Spread Model

The spread model can be divided into two parts which are the spike phase and the stable phase. For the spike phase, we assume a constant spread estimated using the average spread of tick-data which occurs in the first ten seconds after each news announcement. For the stable spread, we will use exponential decay function to model the spread size of the bid-ask quote after the spike phase using the tenth seconds spread as a starting spread. Therefore, a model will have mathematical representations as follow:

$$\Phi_t^{(m)} = \begin{cases} \xi^{(m)} \; if \; t < 10 \\ \xi^{(m)} e^{-\zeta^{(m)}(t-10)} \; if \; t \geq 10 \end{cases} \tag{12}$$

where,

$\Phi_t^{(m)}$ denote the spread of the bid-ask quote for the m^{th} news at time t after the announcement.

$\xi^{(m)}$ denote the spread of the bid-ask quote for the m^{th} news in the spike phase.

$\zeta^{(m)}$ denote the exponential decay rate of the bid-ask spread size for m^{th} news after the spike phase.

Chapter 4 Methodology

This section has four sub-sections which are model estimation, simulation, optimization and out-of-sample trading.

In the estimation sub-section, we estimate the parameters of the currency price dynamic using the EM algorithm based on the ten-second exchange rate data following the first 20 news announcements. This is done separately for good news and bad news. We fit each of the models outlined in Table 1, and use the AIC and BIC to choose the best models. Due to the changing market conditions between the in-sample period (first 20 news announcements) and the out-of-sample period (the remaining news announcements), we adjust the model parameters during the out-of-sample period by using a volatility ratio. In particular, the model parameters are scaled up (down) if forecasted volatility during the out-of-sample period is higher (lower) than that during the in-sample period. We use a GARCH model to estimate the volatility.

The parameters of the spreads will have 2 parts to estimate which are the spike spreads and the decaying part. The spike spreads measured by the 10^{th} second bid-ask spreads are collected from each news announcement during the in-sample period. For the decaying part, the decay rates will be estimated using linear regression on the linearized model of equation (12) to match each decay rate with a given spike spread. The kernel smoothing function will be applied to find a joint probability density function between the spike spreads and the decay rates.

In the simulation sub-section, the Monte-Carlo simulation technique will be used to generate mid-prices for every ten seconds starting from the time of the news announcement until the next 15 minutes. The parameters, after the scaling, from the estimation step will be used to simulate the prices. This is done for each of the news announcements in the out-of-sample period.

Also, the spreads will be simulated in this step. Each consists of two parts which are spike part and decaying part. The size of the spike spread will be simulated using a fitted distribution. The decaying part is simulated using a fitted exponential decay rate parameter.

After the simulation step, we will have simulated bid-ask prices for EUR/USD which are calculated by adding and subtracting the simulated spread to the mid-price. This simulated bid-ask prices will be used in the optimization step.

In the optimization sub-section, an optimization method will be used to find the best price to place the order of buy or sell with the optimal volume and also the target price based on the simulated bid-ask prices. We do this for each of the out-of-sample trading scenarios. There will be 4 strategies to be considered in this sub-section which are Stop Long order, Stop Short order, Limit Long order and Limit Short order. The best strategy which is judged by the best Sharpe ratio of the in-sample trading with an appropriate chance of order matching will be used to trade in the out-of-sample sub-section.

In the out-of-sample trading sub-section, the trading strategy and the optimal solution from the optimization sub-section will be applied to the out-of-sample data. To measure the performance of the strategy, the open-and-hold strategy will be used as a benchmark. This strategy will immediately open the position at the time that the economic news is announced to the market, and hold the position until the end of the trading period which is 15 minutes after news is announced. The opened position of both the chosen strategy and the benchmark will be marked to market at the end of trading period. The performance is measured by the Sharpe ratio.

We now provide the details of each step.

4.1 Price Estimation

The purpose of this section is to estimate the parameters using news and the set of T observable returns $(x_1^{(m)}, x_2^{(m)}, \ldots., x_T^{(m)})$ for $m = 1,2,3, \ldots, M$.

The reason that using maximum likelihood estimation for the case of a Poisson jump process is an unsuccessful approach is because the probability density function of log-returns turns out to be an infinite series due to the number of jumps (Beckers 1981). (Ball and Torous 1985) uses the simplifier model which relaxes the condition of the number of jumps using the Bernoulli random variable for the approximated form of the Poisson-Normal density.

Instead of using maximum likelihood estimation, the estimation method which will be used in this thesis is the EM algorithm based on the fact that the EM recursion will be more numerically robust than direct maximum likelihood.

4.1.1 EM algorithm

Consider the two sets of data which are *Incomplete Dataset* and *Complete Dataset* defined by the vector $X_m$and C_m respectively where,

$$Incomplete\ Dataset:\ X_m = (x_1^{(m)}, x_2^{(m)}, \ldots., x_T^{(m)}) \quad , m = 1,2,3, \ldots, M$$

$$Complete\ Dataset: C_m = \begin{pmatrix} x_1^{(m)} & \ldots & x_T^{(m)} \\ z_1^{(m)} & \ldots & z_T^{(m)} \\ \Delta N_1^{(m)} & \ldots & \Delta N_T^{(m)} \\ \delta_{\Delta N_1}^{(m)} & \ldots & \delta_{\Delta N_T}^{(m)} \\ y_1^{(m)} & \ldots & y_T^{(m)} \end{pmatrix}, \delta_{\Delta N_T}^{(m)} = (\delta_{\Delta N_T^{(m)},1}^{(m)}, \delta_{\Delta N_T^{(m)},2}^{(m)}, \ldots, \delta_{\Delta N_T^{(m)},T}^{(m)})$$

By given the hidden parameter set $\Theta = (\pi, \lambda, \eta, \omega^2, \kappa, \bar{\alpha}_r, \sigma^2, q)$ where $\pi = P(y_1 = 1)$, the complete log-likelihood of the complete dataset $C_1, C_2, \ldots, C_M$can be defined by (See Appendix 2)

$$\begin{aligned} ln(f(C_1, C_2, \ldots, C_M|\Theta_r)) = &\sum_{m=1}^{M} ln\left(P\left(y_1^{(m)}\right)\right) + \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\Delta N_t^{(m)} \middle| y_t^{(m)}\right)\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta_{\Delta N_t^{(m)}}^{(m)} \middle| \Delta N_t^{(m)}, y_t^{(m)}\right)\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(z_t^{(m)} \middle| y_t^{(m)}, \Delta N_t^{(m)}\right)\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T-1} ln(P(y_{t+1}^{(m)}|y_t^{(m)})) \ , m = 1,2,3, \ldots, M \end{aligned} \quad (13)$$

where, $f(C_1, C_2, \ldots, C_M)$ is the joint distribution of complete dataset.

The parameter set Θ is hidden; the only observable data are returns $\mathbb{X} = (X_1, X_2, \ldots, X_M)$. The EM algorithm starts from an initial parameter set Θ_0. Then it computes the associated expectations and probabilities of the expected complete log-likelihood (Expectation step), and them maximizes the expected complete log-likelihood to obtain the new parameter set

Θ_1 (Maximization step). The algorithm iterates between the Expectation and Maximization steps until the parameter set converges. In particular, we define the expected log complete likelihood at the r^{th} loop as

$$Q(\Theta_r|\Theta_{r-1}) = \mathbb{E}(ln(f(C_1, C_2, \dots, C_M)|\Theta_r)\,|\mathbb{X}, \Theta_{r-1}) \;\;, m = 1,2,3, \dots, M$$

where,

$$\Theta_r = (\pi_r, \lambda_r, \eta_r, \omega_r^2, \kappa_r, \bar{\alpha}_r, \sigma_r^2, q_r)$$

is the parameter set from the r^{th} loop, and the expecation is taken over the unobserved data.

Given the parameter set from the previous loop Θ_{r-1}, the new parameter can be found by maximizing $Q(\Theta|\Theta_{r-1})$ with respect to the parameter set Θ. As the basis of an EM algorithm, the function $Q(\Theta|\Theta_{r-1})$ can be determined by the computation in the Expectation-step. With the properties of EM, it can be shown that recursive computation of Θ_r yields monotonic increasing in log-likelihood and the estimated parameters will finally converge to the maximum likelihood estimator $\hat{\Theta}$ for the original incomplete dataset $\mathbb{X}$.

4.1.1.1 Maximization-step

This section will show how to compute each of parameters which are $\pi, \lambda, \eta, \omega^2, \bar{\alpha}, \kappa, \sigma^2$ and q respectively using the first-order condition and Lagrange multiplier method in the maximization step (See Appendix 3)

$$\hat{\pi}_k = \frac{1}{M}\sum_{m=1}^{M} P\left(y_1^{(m)} = k|\mathbb{X},\Theta_{r-1}\right) \tag{14}$$

$$\hat{\lambda}_k = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}\left[\Delta N_t^{(m)}\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right] P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T} P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)} \tag{15}$$

$$\hat{\eta}_k = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} j\,\mathbb{E}\left[\delta_{t,i}^{(m)}\middle|\mathbb{X},\Theta_{r-1},\Delta N_t^{(m)} = j, y_t^{(m)} = k\right] P\left(\Delta N_t^{(m)} = j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right) P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} jP\left(\Delta N_t^{(m)} = j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right) P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)} \tag{16}$$

$$\hat{\omega}_k^2 = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} j\,\mathbb{E}\left[\left(\delta_{t,i}^{(m)} - \eta_k\right)^2\middle|\mathbb{X},\Theta_{r-1},\Delta N_t^{(m)} = j, y_t^{(m)} = k\right] P\left(\Delta N_t^{(m)} = j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right) P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} jP\left(\Delta N_t^{(m)} = j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right) P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)} \tag{17}$$

$$\hat{\bar{\alpha}}_k = \frac{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2^{(m)}} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2^{(m)}} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right)^2} \tag{18}$$

$$\hat{\kappa}_k = \frac{\hat{\bar{\alpha}}_k\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2^{(m)}} P_{(j,t,k)}^{(m)}\right)} \tag{19}$$

$$\hat{\sigma}_k^2 = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}\left(\mathbb{V}ar\left[z_t^{(m)}\middle|\mathbb{X},\Theta_{r-1},\Delta N_t^{(m)} = j, y_t^{(m)} = k\right] + \left(Z_t^{(m)} - \left(\hat{\kappa}_k(\hat{\bar{\alpha}}_k - x_t^{(m)})\right)\right)^2\right) P_{(j,t,k)}^{(m)}}{\sum_{m=1}^{M}\sum_{t=1}^{T} P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right)} \tag{20}$$

$$\hat{q}_{i,l} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T-1} P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X},\Theta_{r-1})}{\sum_{m=1}^{M}\sum_{t=1}^{T-1} P(y_t^{(m)} = i|\mathbb{X},\Theta_{r-1})} \quad for\ i,l = 1,2 \tag{21}$$

where $Z_t^{(m)} = \mathbb{E}\left[z_t^{(m)}\middle|\mathbb{X},\Theta_{r-1},\Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$ (22)

$$P_{(j,t,k)}^{(m)} = P\left(\Delta N_t^{(m)} = j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)} = k\right) P\left(y_t^{(m)} = k\middle|\mathbb{X},\Theta_{r-1}\right) \tag{23}$$

4.1.1.2 Expectation-Step

This section will show how to compute each of the expected values which are needed in the maximization step. The conditional mean and variance for normal distribution are applied in this step.

Consider the jump intensity $\hat{\lambda}_k$, it can be observed that the distribution of $x_t^{(m)}$ given $\Delta N_t^{(m)}$ and regimes $y_t^{(m)}$is Gaussian. The calculation for the r^{th} loop of $\mathbb{E}\left[\Delta N_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k\right]$ can be shown by

$$\mathbb{E}\left[\Delta N_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k\right] = \sum_{s=0}^{\infty} sP\left(\Delta N_t^{(m)} = s \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k\right)$$

$$= \sum_{s=0}^{\infty} sP\left(\Delta N_t^{(m)} = s \middle| x_t^{(m)}, \Theta_{r-1}, y_t^{(m)} = k\right)$$

$$= \sum_{s=0}^{\infty} s\left(\frac{P\left(x_t^{(m)} \middle| \Delta N_t^{(m)} = s, \Theta_{r-1}, y_t^{(m)} = k\right) P\left(\Delta N_t^{(m)} = s \middle| \Theta_{r-1}, y_t^{(m)} = k\right)}{P\left(x_t^{(m)} \middle| \Theta_{r-1}, y_t^{(m)} = k\right)}\right)$$

$$= \sum_{s=0}^{\infty} s\left(\frac{\phi\left(x_t^{(m)}; \rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln\left(S_t^{(m)}\right) + s\eta_k^{(r-1)}, \sigma_k^{2(r-1)} + s\omega_k^{2(r-1)}\right) \frac{\left(\lambda_k^{(r-1)}\right)^s}{s!} e^{-\lambda_k^{(r-1)}}}{\sum_{u=0}^{\infty} \phi\left(x_t^{(m)}; \rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln\left(S_t^{(m)}\right) + u\eta_k^{(r-1)}, \sigma_k^{2(r-1)} + u\omega_k^{2(r-1)}\right) \frac{\left(\lambda_k^{(r-1)}\right)^u}{u!} e^{-\lambda_k^{(r-1)}}}\right)$$

where, $\rho_k^{(r-1)} = \kappa_k^{(r-1)} \bar{\alpha}_k^{(r-1)}$

Now consider $\hat{\eta}_k$ and $\hat{\omega}_k^2$ which are the mean and variance of jump size respectively.

The value which will be computed in this part are $\sum_{j=0}^{\infty} j\mathbb{E}\left[\delta_{t,i}^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$ and $\sum_{j=0}^{\infty} j\mathbb{E}\left[\left(\delta_{t,i}^{(m)} - \hat{\eta}_k\right)^2 \middle| \mathbb{X}, \Theta, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$. Using conditional normal distribution, we have, for the r^{th} loop,

$$\sum_{j=0}^{\infty} j\mathbb{E}[\delta_{t,i}^{(m)} | \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k]$$
$$= \sum_{j=0}^{\infty} j\left(j\eta_k^{(r-1)} + \left(x_t^{(m)} - \left(\rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln S_t^{(m)} + j\eta_k^{(r-1)}\right)\right)\left(\frac{\omega_k^{2(r-1)}}{\sigma_k^{2(r-1)} + j\omega_k^{2(r-1)}}\right)\right)$$

and

$$\sum_{j=0}^{\infty} j\mathbb{E}\left[\left(\delta_{t,i}^{(m)} - \hat{\eta}_k\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$$
$$= \sum_{j=0}^{\infty} j\left(\mathbb{V}ar\left[\delta_{t,i}^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] + \left(\mathbb{E}\left[\delta_{t,i}^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] - \hat{\eta}_k\right)^2\right)$$

where,

$$\mathbb{V}ar\left[\delta_{t,i}^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] = \omega_k^{2(r-1)} - \omega_k^{2(r-1)}\left(\frac{\omega_k^{2(r-1)}}{\sigma_k^{2(r-1)} + j\omega_k^{2(r-1)}}\right).$$

Consider $\hat{\alpha}_k$, $\hat{\kappa}_k$ and $\hat{\sigma}_k$ which are the long-term mean of stock price, speed of mean reversion and volatility of stock respectively. The common terms which are needed for the

r^{th} loop are $Z_t^{(m)} = \mathbb{E}\left[z_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$ and $\mathbb{V}ar\left[z_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$.

The distribution of $z_t^{(m)}$given $x_t^{(m)}$,$y_t^{(m)}$ and $\Delta N_t^{(m)}$ is Gaussian, so conditional normal distribution will be applied to compute both values:

$$\begin{aligned}\mathbb{E}&\left[z_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]\\ &= (\rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln S_t^{(m)})\\ &+ \left(\frac{\sigma_k^{2(r-1)}}{\sigma_k^{2(r-1)} + j\omega_k^{2(r-1)}}\right)\left(x_t^{(m)} - \left(\rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln S_t^{(m)} + j\eta_k^{(r-1)}\right)\right)\end{aligned}$$

$$\mathbb{V}ar\left[z_t^{(m)} \middle| \mathbb{X}, \Theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] = \sigma_k^{2(r-1)} - \sigma_k^{2(r-1)} \left(\frac{\sigma_k^{2(r-1)}}{\sigma_k^{2(r-1)} + j\omega_k^{2(r-1)}}\right)$$

Lastly, for $P\left(y_t^{(m)} = k \middle| \mathbb{X}, \Theta_{r-1}\right)$ the forward-backward algorithm is applied to compute this probability value.

4.1.1.3 Forward-Backward Algorithm

The forward–backward algorithm is an algorithm used to find the probability of being in the unobservable state k at time t of the process given certain information. Formally, this algorithm can be classified into 3 sub-parts which are Forward Procedure, Backward Procedure and Smoothing Procedure. In the Forward procedure, the probability of being in state k at time t is computed by the past information from time 1 to time t. In the Backward procedure, the probability of being in state k at time t is computed backward from the future information from time T to time $t+1$. In the Smoothing Procedure, the process of this part is to normalize the product of the probability given from the Forward and Backward parts.

Using this algorithm to find the value of $P\left(y_t^{(m)} = k \middle| \mathbb{X}, \Theta\right)$ for the r^{th} loop, we have

$$P\left(y_t^{(m)} = k \middle| \mathbb{X}, \Theta_{r-1}\right) = \frac{P\left(\mathbb{X}, y_t^{(m)} = k \middle| \Theta_{r-1}\right)}{\sum_{k=1}^{2} P\left(\mathbb{X}, y_t^{(m)} = k \middle| \Theta_{r-1}\right)} \propto P\left(\mathbb{X}, y_t^{(m)} = k \middle| \Theta_{r-1}\right).$$

Define the forward probability as

$$F_t(y_t^{(m)} = k) = P(y_t^{(m)} = k | x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \dots, x_t^{(m)}, \Theta_{r-1})$$

and the backward probability as

$$B_t(y_t^{(m)} = k) = P(x_T^{(m)}, x_{T-1}^{(m)}, x_{T-2}^{(m)}, \dots, x_{t+1}^{(m)} | y_t^{(m)} = k, \Theta_{r-1}).$$

Using the Bayes' rule, we get

$$\begin{aligned}&P\left(\mathbb{X}, y_t^{(m)} = k \middle| \Theta_{r-1}\right)\\ &\propto P\left(y_t^{(m)} = k \middle| x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \dots, x_t^{(m)}, \Theta_{r-1}\right) P\left(x_T^{(m)}, x_{T-1}^{(m)}, x_{T-2}^{(m)}, \dots, x_{t+1}^{(m)} \middle| y_t^{(m)} = k, \Theta_{r-1}\right)\\ &\propto F_t\left(y_t^{(m)} = k\right) B_t(y_t^{(m)} = k)\end{aligned}$$

Note that $\mathbb{X} = (\mathbb{X}^{(1)}, \mathbb{X}^{(2)}, \dots \mathbb{X}^{(m)})$ where $\mathbb{X}^{(m)} = \left(x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \dots, x_T^{(m)}\right)$ and $\mathbb{X}^{(m)}$ is independent with $\mathbb{X}^{(m\prime)}$ for all $m \neq m'$.

Now, the *Forward Procedure* is applied to find the value of $F_t\left(y_t^{(m)} = k\right)$ by first computing probability $F_1\left(y_1^{(m)} = k\right)$

$$F_1(y_1^{(m)} = k) \propto \sum_{n=0}^{\infty} P(y_1^{(m)} = k) P(x_1^{(m)} | y_1^{(m)} = k, \Delta N_1^{(m)} = n, \Theta_{r-1})\, P(\Delta N_1^{(m)} | y_1^{(m)} = k)$$
$$\propto \sum_{n=0}^{\infty} \pi_k^{(r-1)} \phi(x_1^{(m)}; \rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln(S_1^{(m)}) + n\eta_k^{(r-1)}, \sigma_k^{2(r-1)} + n\omega_k^{2(r-1)})\, \frac{\left(\lambda_k^{(r-1)}\right)^n}{n!} e^{-\lambda_k^{(r-1)}}$$

and then using the recursively transition to find the value of $F_t\left(y_t^{(m)} = k\right)$

$$F_t(y_t^{(m)} = k) \propto \sum_{n=0}^{\infty} \sum_{j=1}^{2} F_{t-1}(y_{t-1}^{(m)} = j) P(y_t^{(m)} = k | y_{t-1}^{(m)} = j) P(x_t^{(m)} | y_t^{(m)} = k, \Delta N_t^{(m)} = n, \Theta_{r-1})\, P(\Delta N_t^{(m)} | y_t^{(m)} = k)$$
$$\propto \sum_{n=0}^{\infty} \sum_{j=1}^{2} F_{t-1}(y_{t-1}^{(m)} = j) q_{jk}^{(r-1)} \phi(x_1^{(m)}; \rho_k^{(r-1)} - \kappa_k^{(r-1)} \ln(S_1^{(m)}) + n\eta_k^{(r-1)}, \sigma_k^{2(r-1)} + n\omega_k^{2(r-1)})\, \frac{\left(\lambda_k^{(r-1)}\right)^n}{n!} e^{-\lambda_k^{(r-1)}}$$

After we get the value of $F_T(y_T^{(m)} = k)$ from the *Forward Procedure,* it is more convenient to find $P\left(y_t^{(m)} = k \middle| \mathbb{X}, \Theta_{r-1}\right)$ by using the value of $F_t(y_t^{(m)} = k)$ and compute the backward probabilities from time T back to time 1. Define

$$G(y_t^{(m)} = k) \propto P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1}) = \textstyle\sum_{g=1}^{2} P(y_t^{(m)} = k, y_{t+1}^{(m)} = g | \mathbb{X}, \Theta_{r-1})$$
$$\propto \sum_{g=1}^{2} P(y_{t+1}^{(m)} = g | \mathbb{X}, \Theta_{r-1}) P(y_t^{(m)} = k | y_{t+1}^{(m)} = g, \mathbb{X}, \Theta_{r-1})$$
$$\propto \sum_{g=1}^{2} G(y_{t+1}^{(m)} = g) P(y_t^{(m)} = k | y_{t+1}^{(m)} = g, \mathbb{X}^{(m)}[1{:}t], \Theta_{r-1})$$
$$\propto \sum_{g=1}^{2} G(y_{t+1}^{(m)} = g) \left(\frac{P(y_{t+1}^{(m)} = g | y_t^{(m)} = k, \mathbb{X}^{(m)}[1{:}t], \Theta_{r-1}) P(y_t^{(m)} = k | \mathbb{X}^{(m)}[1{:}t], \Theta_{r-1})}{P(y_{t+1}^{(m)} = g | \mathbb{X}^{(m)}[1{:}t], \Theta_{r-1})} \right)$$
$$\propto \sum_{g=1}^{2} G(y_{t+1}^{(m)} = g) \left(\frac{q_{kg}^{(r-1)} F_t(y_t^{(m)} = k)}{\sum_{h=1}^{2} P(y_{t+1}^{(m)} = g | y_t^{(m)} = h)\, F_t(y_t^{(m)} = h)} \right)$$
$$\propto \sum_{g=1}^{2} G(y_{t+1}^{(m)} = g) \left(\frac{q_{kg}^{(r-1)} F_t(y_t^{(m)} = k)}{\sum_{h=1}^{2} q_{hg}^{(r-1)}\, F_t(y_t^{(m)} = h)} \right)$$

where $G\left(y_T^{(m)} = k\right) = F_T\left(y_T^{(m)} = k\right)$ and $\mathbb{X}^{(m)}[1{:}t] = (x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, \dots, x_t^{(m)})$

Therefore, $P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1}) \propto \sum_{g=1}^{2} G(y_{t+1}^{(m)} = g) \left(\frac{q_{kg}^{(r-1)} F_t(y_t^{(m)} = k)}{\sum_{h=1}^{2} q_{hg}^{(r-1)} F_t(y_t^{(m)} = h)} \right)$

4.1.2 Akaike Information Criterion (AIC)

The Akaike Information Criterion method is a statistical method that used to compare the efficiency of the model. This method will compute the log-likelihood of the model as a reward and penalize it with the number of parameters fitted in the model. The main source of penalty is from the over-fitting parameter because the increase in the number of parameters in the model will almost always improve the log-likelihood. Given a log-likelihood $loglik$ and the number of parameters $nParam$, AIC can be calculated by

$$AIC = -2(loglik) + 2(nParam)$$

Based on the formulation, the model with the lowest AIC is preferred.

4.1.3 Bayesian Information Criterion (BIC)

Similar to AIC, the Bayesian Information Criterion method can be used to compare the goodness-of-fit of the model. Given a log-likelihood $loglik$, the number of observations $nObs$ and the number of parameters $nParam$, BIC can be calculated by

$$BIC = -2(loglik) + (nParam)(ln(nObs))$$

Based on the formulation, the model with the lowest BIC is preferred.

As shown from the mathematical expressions of the AIC and BIC, we can see that the AIC framework tries to select the best model accounting for the number of parameters but ignoring the number of observations. On the contrary, the BIC framework tries to find the best fitted model taking into the number of parameters and the number of observations.

Study of (Acquah 2010) shows that the AIC framework overcomes the BIC under the unstable data conditions (i.e. small sample size or large noise level).

4.1.4 Volatility Scale Ratio

As the parameters are fitted based on the in-sample data, the volatility of the currency movement during the out-of-sample period can be changed due to the change in the market condition, such as ECB (Euro central Bank) stimulates the economy system by QE (Quantitative Easing). Therefore, we use a volatility adjustment ($\Delta^{(m)}$) to adjust the model parameters for each of the out-of-sample news announcements

By assuming that in the out-of-sample data the m^{th} news will be announced on the date k, the volatility scale ratio for that news will be computed using GARCH(1,1) by

$$\Delta^{(m)} = \sqrt{\frac{v_{k-1}^{(m)}}{\bar{v}_{in-sample}}}$$

where $v_{k-1}^{(m)}$ is the variance of the log-return on date $k-1$ based on the GARCH(1,1) model fitted using the daily data from the first date of the in-sample period to the day before the m^{th} news is announced, and $\bar{v}_{in-sample}$ is the mean of the variances of the log-return of the daily data during the in-sample period estimated from a GARCH(1,1) model using the data from January 1, 2010 to December 31, 2014.

4.2 Spread Estimation

The estimation of the spread will be divided into two parts which are the spike part and decaying part. The steps to estimate the spread parameters are as follow:

1. The estimation for the spike spread can be done by averaging all of the spreads of the tick-data during the first ten seconds after each of the news announcements. Then, we will store all of the averaged values in the M-dimensional vector ξ where M is the number of news. Therefore, $\xi = [\xi^{(1)}\ \xi^{(2)}\ \xi^{(3)} \dots \xi^{(M)}]$.

2. Recall that the spread at time $t > 10$ is $\Phi_t^{(m)} = \xi^{(m)} e^{-\zeta^{(m)}(t-10)}$. The estimation for the decaying spread can be done by using linear regression on the historical spreads for each news announcement after the spike phase (i.e. 10 seconds after the announcement time). The regression equation is $ln\Phi_t^{(m)} = ln\,\xi^{(m)} - \zeta^{(m)}(t-10)$. Then, we will store all of the parameter values in the M-dimensional vector ζ where M is the number of news. Therefore, $\zeta = [\zeta^{(1)}\ \zeta^{(2)}\ \zeta^{(3)} \dots \zeta^{(M)}]$.

3. Use a Kernel smoothing function for estimating the joint probability density of ξ and ζ. Then, apply the t −copula fitting between the smoothed ξ and smoothed ζ which is characterized by the correlation matrix $\hat{\mathrm{P}}$ and the degree of freedom $\hat{\mathrm{N}}$ for creating the joint distribution between ξ and ζ.

4.3 Simulation

In this section price dynamic of the currency will be simulated with the model:

$$\frac{dS_t^{(m)}}{S_t^{(m)}} = \kappa\left(y_t^{(m)}\right)\left[\alpha\left(y_t^{(m)}\right) - ln\,S_t^{(m)}\right]dt + \Delta^{(m)}\sigma\left(y_t^{(m)}\right)dW_t^{(m)} + \theta_t^{(m)}\left(y_t^{(m)}\right)dN_t^{(m)}$$

where $y_t^{(m)}$ is assigned to be 1 or 2 according to transient state and steady state

(Ball and Torous 1985) come up with a simplified model for Poisson jump process using Bernoulli jump process. They state that if the return is computed for the very small time interval the Bernoulli model would converge to the Poisson model. Therefore, Bernoulli jump process will be used in the simulation for jumps instead of Poisson jump model to reduce the complexity of the infinite series representation of probability density function of log-return.

In the optimization step, we need to compute the expected value of the terminal wealth for a given model and trading strategy. This quantity is complex, and hence we approximate the expectation by using a simulation technique. With the parameters from the estimation step, we will simulate the mid-price of the currency (after adjusting the volatility with a scaling ratio) to get 15 minutes of ten-seconds time-step data. We repeat this to obtain 5000 paths for each news announcement scenario. At the first time step the mid-price of the currency will be scaled to 1, so that the dynamic of the price will be the percentage change from the price right before the news announcement.

Furthermore, we will also simulate spreads. Spreads are divided into two phases after the news announcement which are spike phase and decaying phase.

First, we will use t −copula with the correlation matrix $\hat{\mathrm{P}}$ and the degree of freedom $\hat{\mathrm{N}}$ to create the joint distribution between spike spread size ξ and decay rate ζ as the vector $\hat{T}$ = $[\ \xi\ \ \zeta\ ]$.

We simulate the spike spread size and the decay rate following these steps:

1. The spike spread size will be simulated by using inverse cumulative density function from the first column of $\hat{T}$ with the average spike spread sizes from historical data for each scenario.

2. We will use the kernel smoothing to find the decay rate. Given the simulated spread size, we use the inverse cumulative density function with the second column of vector $\hat{T}$.

The results from these steps are the simulated spike spread and the exponential decay rate which will be used in the decaying phase. In the decaying phase, spreads after the first ten seconds will be calculated using the equation (12). Declining of spread will start at the tenth second after the new announcement and continue until 15 minutes after the news announcement had passed.

By adding to and subtracting from the simulated currency mid-prices with the half of the spreads, we will get the simulated ten-second bid-ask prices for EUR/USD for good and bad news.

4.4 Optimization

Suppose a trader has initial wealth W_0at time 0 before the news is announced. When the news is announced, the price will start to fluctuate rapidly, and the trader will set an open position with a desired lot size at this time. When the order is placed the target price is also set and traders will do nothing until the currency price move up/down to the target price and then close the position. In another case that the price is unable to hit the target price before the 15-minute period has passed, the trader must close the position and realize the profit or loss. Also, there is a chance that the trader may hold too excessive position that the opened position will be forced to close when the buffered cash left in the portfolio is not sufficient.

In this thesis, we will find the best strategy to trade the EUR/USD currency during the news announcement period. The strategies covered in this thesis are the Stop Long order, Stop Short order, Limit Long order and Limit Short order. The strategy with the best Sharpe ratio of the in-sample trading with an appropriate chance of order matching from in-sample test will be used in the out-of-sample trading. To compute the Sharpe ratio, we will calculate the Sharpe ratio only from the cases in which the order is matched. The reason is that the limit orders are rarely matched and the wealth data seem to be overly stable, making the Sharpe ratio computed from all cases too extreme.

In this section, we will use one-step optimization to find the optimal price to open the position (β_1), lot size of the opened position (V) and the target price to close the position (β_2). The objective function of this optimization is based on maximizing expected wealth of trader when the position is closed by using the average wealth from each path due to the complexity in computing the expected value of wealth. The optimization problem for the good news is modeled as

$$max_{(\beta_1,\beta_2,V)\in\mathbb{R}^3}\left\{\frac{1}{Q}\sum_{q=1}^{Q} W_q\right\}$$
$$subject\ to\quad \beta_1,\beta_2,V \geq 0$$
$$\beta_2 \geq \beta_1$$

where Q is the total number of simulated paths, W_q is the trader's wealth after the position is closed for price path q (Note that for bad news we also use a similar model except for the last constraint which will be $\beta_2 \leq \beta_1$ instead).

In news trading, there are many factors to consider. The most important factor to consider is the possibility of a forced close in their position due to the high volatility after news is announced. Traders will be forced to close when the cash buffered in their portfolio, called free margin, is less than or equal to 30 percent of the total margin that traders pay for opening their trading position.

To be precise, denote the free margin by F_t, which is the cash buffered at time t in the portfolio. It can be calculated by subtracting the total margin denoted by K_t from the sum of cash left in portfolio ψ_t and unrealized profit U_t:

$$F_t = \psi_t + U_t - K_t$$

The position will be forced to close immediately if the cash buffered in the portfolio is less than 30 percent of the total margin, so we can set an inequality for the forced closure constraint as follows:

$$\frac{F_t}{K_t} \geq 0.3 \tag{14}$$

Traders will need to set their leverage (L) for their portfolio due to the costly contract price which is fixed at 100,000 US dollar per contract. In the scope of this thesis, the leverage will be set at 2,000 times which is a normal case for the traders with less amount of money. Denote by B the margin that traders need to place for each contract. Now cash left and total margins are fixed. Therefore, cash left ψ_t can be calculated by $W_0 - \frac{VB}{L}$.

Instead of viewing a free margin as a forced closure constraint, it is easier to view a forced closure constraint as the amount of unrealized return. Let r_t denote the unrealized return for EUR/USD without leverage. The unrealized leveraged profit U_t can be expressed as the product of the opened lot size V with the unrealized leveraged return for each contract Br_t. Therefore, the unrealized profit U_t can be represented by VBr_t. Thus, the inequality constraints for forced closure can be rewritten from equation (24) as:

$$\left(W_0 - \frac{VB}{L}\right) + (VBr_t) \geq\ 0.3\frac{VB}{L}\beta_1$$
$$r_t \geq \frac{0.3\beta_1}{L} + \frac{1}{L} - \frac{W_0}{VB}$$

With this inequality, it is shown that the return of the portfolio for all time t must be greater than $\frac{0.3\beta_1}{L} + \frac{1}{L} - \frac{W_0}{VB}$ to prevent position forced closure.

There is also another constraint that prevents an immediate forced closure after the position is matched; due to the spread widening. It can be written as follows:

$$\frac{\left(W_0 - \frac{VB}{L}\right) - VB\left(\frac{\Phi_0}{\beta_1}\right)}{K_t} \geq 0.3$$

where $VB\left(\frac{\Phi_0}{\beta_1}\right)$ denote the unrealized loss due to the spread.

Consequently, to calculate the realized return, let τ denote the time when the position is closed, then

$$r_\tau = \frac{P_\tau - \beta_1}{\beta_1} \qquad P_\tau = \begin{cases} \beta_2 & if\ \tau_{\beta_2} = min(\tau_{\beta_2}, T, \tau_d) \\ P_T & if\ T = min(\tau_{\beta_2}, T, \tau_d) \\ P_d & if\ \tau_d = min(\tau_{\beta_2}, T, \tau_d) \end{cases} \qquad \begin{aligned} \tau &= min(\tau_{\beta_2}, T, \tau_d) \\ \tau_{\beta_1} &= inf\{s \geq 0 : P_s \leq \beta_1\} \\ \tau_{\beta_2} &= inf\{s \geq \tau_o : P_s \geq \beta_2\} \\ \tau_d &= inf\{s \geq \tau_o : P_s \leq P_d\} \end{aligned}$$

where,

r_τ denote the currency return if the position is closed at time τ

P_T denote the price at time T

P_d denote the price when the position is forced closed

P_τ denote the price when the position is closed

τ_0 denote the time that the trading order is matched

τ_{β_1} denote the time when the position is opened, or infinity otherwise

τ_d denote the time when the position is forced close , or infinity otherwise

τ_{β_2} denote the time when the price reaches the target price β_2 after the position is opened

In conclusion, the trader's wealth W_q can be calculated by the sum of cash left ψ_τ and leveraged realized profit/loss R_τ (in dollars amount)

$$W_q = \psi_\tau + R_\tau$$

where

$$R_\tau = VBr_\tau \mathbb{I}\{\tau_0 < T\},$$

and $\mathbb{I}$ is an indicate function taking value one if the limit order is matched at time $\tau_0 < T$ and zero otherwise.

In Section 5.2 the results are reported as the currency returns. The percentage changes of the trader's wealth are reported in Appendix 4. Note that this is for the case of the Limit Long position. For the other strategy, there will be a small change in the calculation. The differences are shown in Table 2.

Table 2 : Calculation of the currency return for the optimal trading for each order type

Limit Long position order	Limit Short order
$\tau_{\beta_1} = inf\{s \geq 0 : P_s \leq \beta_1\}$ $\tau_d = inf\{s \geq \tau_o : P_s \leq P_d\}$ $r_\tau = \dfrac{P_\tau - \beta_1}{\beta_1}$	$\tau_{\beta_1} = inf\{s \geq 0 : P_s \geq \beta_1\}$ $\tau_d = inf\{s \geq \tau_o : P_s \geq P_d\}$ $r_\tau = \dfrac{\beta_1 - P_\tau}{\beta_1}$
Stop Long position order	**Stop Short position order**
$\tau_{\beta_1} = inf\{s \geq 0 : P_s \geq \beta_1\}$ $\tau_d = inf\{s \geq \tau_o : P_s \leq P_d\}$ $r_\tau = \dfrac{P_\tau - \beta_1}{\beta_1}$	$\tau_{\beta_1} = inf\{s \geq 0 : P_s \leq \beta_1\}$ $\tau_d = inf\{s \geq \tau_o : P_s \geq P_d\}$ $r_\tau = \dfrac{\beta_1 - P_\tau}{\beta_1}$

4.5 Out of Sample Test

We will compare our strategy with an open-and-hold strategy which chooses to open either a long or short position at the current market price when the news is announced. The position will be closed at price P_τ when the time reaches the limit (15 minutes). We consider 3 open-and-hold benchmarks which are the Always Long Strategy, the Always Short Strategy and the Mixed Long and Short Strategy.

The Always Long Strategy will immediately open the long position at the best ask price when the news is announced and hold the position until the end of the trading period (15-minute holding).

The Always Short Strategy will immediately open the short position at the best bid price when the news is announced and hold the position until the end of the trading period (15-minute holding).

The Mixed Long and Short Strategy will choose to open either the long or the short position, whichever provides a higher return, immediately when the news is announced and hold the position until the end of the trading period (15-minute holding).

Let P_0^{ask} denote the market best ask price at the time that news is announced, and P_0^{bid} denote the market best ask price at the time that news is announced. Let Cur_τ denote the currency return when the position is closed at time τ.

We use the currency return to measure the performance of the suggested strategy and benchmarks. The currency return from the benchmark strategies are computed as shown in Table 3. The results are reported in the Section 5.2 as the currency returns.

Table 3 : Calculation of currency return for the benchmark for each order type

Long position order	Short position order
$\tau_d = inf\{s \geq \tau_o : P_s \leq P_d\}$ $Cur_\tau = \dfrac{P_\tau - P_0^{ask}}{P_0^{ask}}$	$\tau_d = inf\{s \geq \tau_o : P_s \leq P_d\}$ $Cur_\tau = \dfrac{P_0^{bid} - P_\tau}{P_0^{bid}}$

Chapter 5 Result and Discussion

5.1 Parameters Estimation and model selection

Table 4 : Log-likelihood, AIC and BIC Value for good news

Model	Log-likelihood	# of parameter	# of observation	AIC	BIC
2-Regime Pure Diffusion with Jump (2-PDJ)	15,100	10	1,800	-30,181	-30,126
2-Regime Mean-Reverting with Jump (2-MRJ)	15,101	14	1,800	-30,174	-30,097
2-Regime Geometric Brownian Motion with Jump (2-GBMJ)	15,055	12	1,800	-30,087	-30,021
2-Regime Mean-Reverting (2-MR)	15,024	8	1,800	-30,033	-29,989
2-Regime Pure Diffusion (2-PD)	15,017	4	1,800	-30,026	-30,004
2-Regime Geometric Brownian Motion (2-GBM)	15,017	6	1,800	-30,022	-29,989
Geometric Brownian Motion with Jump (GBMJ)	15,000	5	1,800	-29,991	-29,963
Mean-Reverting with Jump (MRJ)	15,000	6	1,800	-29,989	-29,956
Pure Diffusion with Jump (PDJ)	14,698	4	1,800	-29,388	-29,366
Pure Diffusion (PD)	14,625	1	1,800	-29,248	-29,242
Geometric Brownian Motion (GBM)	14,625	2	1,800	-29,246	-29,235
Mean-Reverting (MR)	14,625	3	1,800	-29,244	-29,228

Table 4 shows the performance of the model fitting for the good-news in-sample data using the AIC and BIC values. The model with the least AIC and BIC values is preferred which is the 2-Regime pure diffusion with jump (2-PDJ). The dynamic of this model can be expressed by setting the parameters κ and $\bar{\alpha}$ to zero in full model given by equation (10).

The AIC and BIC values from Table 4 obviously show that allowing regime switching improves the fit of the model as expected. We can see that all of the models with two regimes give the better result in model fitting, and the models with jump are better fitted to the exchange rate than the non-jump-model although the regime switching is allowed or not. We can see from the result that in the good-news announcement, allowing the regime switching to each parameter in the model is preferred to adding jump factor to the model. The proposed 2-MRJ model gives the best log-likelihood value as expected due to higher number of parameters. This value shows that the proposed model has the best potential in estimation comparing with all others. However, the AIC and BIC values of the proposed model cannot overcome the simpler 2-PDJ model due to the penalty from the complexity of the model. Using the AIC and BIC values as a benchmark to justify the model in simulation, the model that will be used is 2-PDJ instead of the proposed model.

Table 5 : Log-likelihood, AIC and BIC Value for bad news

Model	Log-likelihood	# of parameter	# of observation	AIC	BIC
2-Regime Mean-Reverting with Jump (2-MRJ)	14,948	14	1,800	-29,869	-29,792
2-Regime Pure Diffusion with Jump (2-PDJ)	14,926	10	1,800	-29,833	-29,778
2-Regime Geometric Brownian Motion with Jump (2-GBMJ)	14,921	12	1,800	-29,818	-29,752
Pure Diffusion with Jump (PDJ)	14,876	4	1,800	-29,745	-29,723
Geometric Brownian Motion with Jump (GBMJ)	14,877	5	1,800	-29,745	-29,717
Mean-Reverting with Jump (MRJ)	14,877	6	1,800	-29,743	-29,710
2-Regime Pure Diffusion (2-PD)	14,867	4	1,800	-29,727	-29,705
2-Regime Geometric Brownian Motion (2-GBM)	14,866	6	1,800	-29,720	-29,687
2-Regime Mean-Reverting (2-MR)	14,811	8	1,800	-29,606	-29,562
Pure Diffusion (PD)	14,518	1	1,800	-29,035	-29,029
Geometric Brownian Motion (GBM)	14,518	2	1,800	-29,033	-29,022
Mean-Reverting (MR)	14,519	3	1,800	-29,032	-29,015

Table 5 shows the performance of the model fitting for the bad-news in-sample data using AIC and BIC values. The model with the least AIC and BIC values is preferred which is the proposed model: 2-Regime Mean-Reverting with Jump (2-MRJ).

The AIC and BIC values from the Table 5 show that adding the jump parameters improves the fit of the model. We can see that all of the models with jumps are better fitted to the exchange rate than the non-jump-model. We can also see from the Table 5 that in the bad-news announcement, adding jump parameters to the model is preferred than the regime switching. The proposed model 2-MRJ gives the best log-likelihood, AIC and BIC values as expected. Using the AIC and BIC values as a benchmark to justify the model in simulation, the model that will be used is 2-MRJ.

Table 6 : Good news parameters values of 2-regime pure diffusion with jump (2-PDJ)

	Regime-1	Regime-2
Parameter	**Values**	**Values**
$\boldsymbol{\pi}$	1.0000	0.0000
$\boldsymbol{\lambda}$	0.1544	0.3906
$\boldsymbol{\omega^2}$ **(x$\mathbf{10^{-8}}$)**	0.8111	4.0887
$\boldsymbol{\eta}$ **(x$\mathbf{10^{-6}}$)**	-0.4668	-5.9176
$\boldsymbol{\sigma^2}$ **(x$\mathbf{10^{-9}}$)**	1.0057	2.0884
$\boldsymbol{q}$	0.9963	0.0037

Table 6 shows the parameter value of the 2-regime pure diffusion with jump (2-PDJ) model. By setting the parameters κ and $\bar{\alpha}$ to zero in equation (10), the dynamic of the model is expressed as:

$$da_t^{(m)} = \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1 + \theta_t^{(m)}\left(y_t^{(m)}\right)) \, dN_t^{(m)}$$

The estimated parameter values suggest that the dynamic of the currency after the news announcement is mainly driven by the volatility $\sigma\left(y_t^{(m)}\right)$ and the jump size $ln(1 + \theta_t^{(m)}\left(y_t^{(m)}\right))$. Since the jump size $ln(1 + \theta_t^{(m)}\left(y_t^{(m)}\right))$ is normally distributed with η as the mean jump size and ω^2 as the variance of the jump size, the expected return from the model for each given regime $y_t^{(m)}$ can be calculated by $\mathbb{E}\left[da_t^{(m)}|\Theta\right] = \lambda\eta dt$ which are the negative values for both regimes. Therefore, the suitable strategy for the good news announcement is to open a short position.

The trading period last 15 minutes after the news is announced and is divided into 90 10-second time-steps. Given that the regime is initially a transient regime, we expect that the regime will be changed from the transient regime to the steady regime during this period with probability of $1 - (0.9963)^{90} = 0.2837$ or more. After the regime has been changed from the transient regime to the steady regime, it is obviously seen from Table 6 that the parameter values are significantly changed (e.g. the mean jump size η is changed from -0.4668 to -5.9176, as well as other parameter values.) The result in changing the regime makes the expected return to be more negative value.

The Limit Short order needs to be placed above the current price and waits until the price moves up to where the order is placed. According to the suggestion from the model that the exchange rate is expected to instantly move in downward direction after the news announcement and drop faster after the regime switches, we expect that using the Limit Short order strategy, the submitted order will be rarely matched. With this reason the Limit Short order strategy is expected to be placed with a huge lot size to improve the Sharpe ratio when the order is matched. The advantage of using this strategy can have a huge gain from this strategy in the case that the order is matched which gives a large Sharpe ratio to this strategy as a trade-off with the order matching occurrence.

However, using the Stop Short order strategy may give a lower Sharpe ratio than the Limit Short order strategy, given the order is matched, but can significantly improve the chance of order matching since the stop order can instantly follow the direction of the exchange rate at the current price. The opened order may have a lower lot size compared with the Limit Short order strategy because the Stop Short order strategy needs to follow the exchange rate direction, so the entry price of the stop order will be worse than the limit order. Therefore, we expect that using the Stop Short order strategy will improve the chance of order matching with a lower lot size as a trade-off compared with the Limit Short order strategy. With this reason the Stop Short order strategy may give a lower Sharpe ratio if we are comparing with the Limit Short order strategy due to the smaller lot size and lower entry price.

Table 7 : Bad news parameters values of 2-Regime Mean-Reverting with Jump (2-MRJ)

	Regime-1	Regime-2
Parameter	**Values**	**Values**
$\boldsymbol{\pi}$	0.7427	0.2573
$\boldsymbol{\lambda}$	0.1842	0.3557
$\boldsymbol{\omega^2}$ **(x10^{-8})**	0.7302	3.0501
$\boldsymbol{\eta}$ **(x10^{-6})**	5.8969	4.3869
$\boldsymbol{\sigma^2}$ **(x10^{-9})**	0.7843	1.3855
$\boldsymbol{\kappa}$	0.0068	0.0046
$\boldsymbol{\bar{\alpha}}$ **(x10^{-7})**	6.2796	7.4857
$\boldsymbol{q}$	0.9964	0.0036

Table 7 shows the parameter values of the 2-Regimes Mean-Reverting with Jump (2-MRJ). The dynamic of the model is expressed as:

$$da_t^{(m)} = \kappa\left(y_t^{(m)}\right)\left[\bar{\alpha}\left(y_t^{(m)}\right) - a_t^{(m)}\right] dt + \sigma\left(y_t^{(m)}\right) dW_t^{(m)} + ln(1 + \theta_t^{(m)}\left(y_t^{(m)}\right)) \, dN_t^{(m)}$$

The estimated parameter values, the model suggest that the dynamic of the currency after the news announcement has a mean reversion effect due to the positive value of mean convert speed parameter $\kappa\left(y_t^{(m)}\right)$. With a positive long-run mean $\bar{\alpha}\left(y_t^{(m)}\right)$ and positive mean jump size $\eta\left(y_t^{(m)}\right)$, we expect that the long position can take advantage in this scenario.

However, the choice that traders will choose the Stop Long order or Limit Long order to take action in this case is not yet obvious because the Limit Long order strategy can intuitively take advantage from the mean reversion effect but there is also a trade-off with the chance of order matching. On the other hand, the Stop Long order strategy is a great strategy for a trending markets but the entry may not be as low as that from the limit order. So to choose the strategy to trade during the bad news scenario, we will use the simulation.

5.2 In-sample Trading Simulation

This section will show the simulation trading results of the good and bad news for each scenario. In each scenario, the exchange rate movement will be simulated for 5,000 paths and based on the simulated paths, we find the optimal open, lot size and target price.

Each table reports the detailed simulation trading result for each strategy with each news result. It provides the volatility scale ratio (Scale Ratio), average currency returns (Currency Return), standard deviation (Std.), Sharpe ratio (Sharpe Ratio), optimal open prices (Open), optimal lot sizes (Lot), optimal target prices (Target), percentage matched orders (%Matched), and percentage forced close order (%Force).

There are 18 scenarios for the good news and 23 scenarios for the bad news. Each scenario will be sorted in descending order by the volatility scale ratio in *Scale Ratio* column.

Table 8 and Table 9 will report the results from the simulated trades using the Stop order strategies for good news scenarios. Table 10 and Table 11 will report the results from the simulated trades using the Limit order strategies for good news scenarios. Table 12 and Table 13 will report the results from the simulated trades using the Stop order strategies for bad news scenarios. Table 14 and Table 15 will report the results from the simulated trades using the Limit order strategies for bad news scenarios.

There are also Table 16 and Table 17 which show the summary result from applying each strategy in the simulation trading. (Good and Bad news result respectively)

Table 8: In-sample result for each good news scenario using Stop Long order strategy trading on simulation

Stop Long Strategy

Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	-0.008	0.055	-0.139	1.000	0.010	1.001	97.330	0.000
2	0.978	-0.008	0.050	-0.156	1.000	0.010	1.001	96.440	0.000
3	0.918	-0.009	0.049	-0.176	1.000	0.010	1.001	96.560	0.000
4	0.881	-0.008	0.037	-0.207	1.000	0.023	1.000	96.780	0.000
5	0.860	-0.007	0.029	-0.255	1.000	0.024	1.000	97.000	0.000
6	0.842	-0.008	0.030	-0.270	1.000	0.023	1.000	96.330	0.000
7	0.842	-0.008	0.030	-0.271	1.000	0.023	1.000	96.560	0.000
8	0.840	-0.008	0.030	-0.277	1.000	0.022	1.000	96.560	0.000
9	0.838	-0.009	0.031	-0.278	1.000	0.022	1.000	96.560	0.000
10	0.814	-0.008	0.028	-0.292	1.000	0.013	1.000	97.220	0.000
11	0.795	-0.008	0.027	-0.308	1.000	0.022	1.000	96.560	0.000
12	0.783	-0.008	0.023	-0.338	1.000	0.023	1.000	96.670	0.000
13	0.757	-0.010	0.007	-1.387	1.001	0.010	1.001	6.110	0.000
14	0.727	-0.008	0.036	-0.230	1.000	0.022	1.000	99.330	0.000
15	0.713	-0.009	0.043	-0.214	1.000	0.010	1.001	96.780	0.000
16	0.691	-0.009	0.042	-0.211	1.000	0.010	1.001	99.330	0.000
17	0.673	-0.008	0.041	-0.202	1.000	0.010	1.001	96.890	0.000
18	0.672	-0.008	0.041	-0.204	1.000	0.010	1.001	96.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 9: In-sample result for each good news scenario using Stop Short order strategy trading on simulation

Stop Short Strategy

Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	-0.001	0.055	-0.017	1.000	0.010	0.999	99.110	0.000
2	0.978	0.000	0.048	-0.008	1.000	0.010	0.999	94.440	0.000
3	0.918	0.000	0.048	-0.008	1.000	0.010	0.999	98.110	0.000
4	0.881	0.000	0.046	0.003	1.000	24.330	0.999	94.560	0.000
5	0.860	0.001	0.046	0.011	1.000	32.997	0.999	96.670	0.000
6	0.842	0.001	0.042	0.028	1.000	57.316	0.999	91.110	0.240
7	0.842	0.001	0.047	0.023	1.000	58.708	0.999	98.780	0.340
8	0.840	0.001	0.047	0.031	1.000	33.034	0.999	99.670	0.000
9	0.838	0.001	0.047	0.025	1.000	34.328	0.999	99.890	0.000
10	0.814	0.001	0.046	0.019	1.000	32.811	0.999	98.330	0.000
11	0.795	0.001	0.046	0.018	1.000	32.855	0.999	99.890	0.000
12	0.783	0.000	0.046	0.011	1.000	32.300	0.999	99.890	0.000
13	0.757	0.000	0.043	0.002	1.000	32.100	0.999	97.560	0.000
14	0.727	0.000	0.042	0.002	1.000	0.856	0.999	97.890	0.000
15	0.713	0.000	0.042	-0.008	1.000	0.010	0.999	98.440	0.000
16	0.691	0.000	0.041	0.001	1.000	1.566	0.999	98.440	0.000
17	0.673	0.000	0.039	-0.003	1.000	0.110	0.999	98.440	0.000
18	0.672	0.000	0.040	-0.004	1.000	0.096	0.999	99.670	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 10: In-sample result for each good news scenario using Limit Long order strategy trading on simulation

Limit Long Strategy									
Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	0.012	0.015	0.782	0.999	67.201	1.001	8.000	0.000
2	0.978	0.021	0.014	1.463	0.999	79.471	1.001	5.330	0.000
3	0.918	0.009	0.012	0.763	0.999	82.822	1.001	4.560	2.440
4	0.881	0.000	0.009	0.003	0.999	26.810	1.001	4.780	0.000
5	0.860	0.008	0.012	0.665	0.999	86.833	1.001	4.670	2.380
6	0.842	0.010	0.013	0.795	0.999	92.839	1.001	5.110	8.700
7	0.842	0.009	0.013	0.700	0.999	92.243	1.001	5.220	6.380
8	0.840	0.009	0.013	0.730	0.999	92.911	1.001	5.000	13.330
9	0.838	0.004	0.013	0.311	0.999	110.669	1.001	4.780	48.840
10	0.814	0.005	0.010	0.462	0.999	92.951	1.001	4.890	6.820
11	0.795	0.005	0.011	0.433	0.999	101.973	1.001	5.110	26.090
12	0.783	0.020	0.011	1.816	0.999	91.042	1.000	4.670	2.380
13	0.757	0.005	0.008	0.680	0.999	89.850	1.001	4.330	2.560
14	0.727	0.001	0.005	0.155	0.999	88.231	1.001	3.330	0.000
15	0.713	0.006	0.005	1.121	0.999	97.669	1.001	3.110	0.000
16	0.691	0.015	0.007	2.260	0.999	104.142	1.001	2.890	11.540
17	0.673	0.018	0.006	3.022	0.999	104.331	1.001	2.560	4.350
18	0.672	0.018	0.006	3.015	0.999	104.273	1.001	2.560	4.350

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 11: In-sample result for each good news scenario using Limit Short order strategy trading on simulation

Limit Short Strategy

Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	0.004	0.012	0.316	1.001	67.031	0.999	6.670	0.000
2	0.978	0.004	0.010	0.378	1.001	73.814	0.999	5.670	0.000
3	0.918	0.009	0.011	0.853	1.001	83.042	0.999	4.670	2.380
4	0.881	0.006	0.008	0.671	1.001	89.001	0.999	3.890	5.710
5	0.860	0.006	0.008	0.808	1.001	83.736	0.999	3.780	2.940
6	0.842	0.016	0.010	1.632	1.001	86.872	0.999	3.670	3.030
7	0.842	0.018	0.010	1.883	1.001	87.090	0.999	3.560	0.000
8	0.840	0.017	0.010	1.697	1.001	87.108	0.999	3.440	3.230
9	0.838	0.015	0.010	1.553	1.001	87.225	0.999	3.560	0.000
10	0.814	0.012	0.008	1.552	1.001	88.256	0.999	3.110	0.000
11	0.795	0.007	0.008	0.894	1.001	88.571	0.999	3.330	0.000
12	0.783	-0.005	0.026	-0.199	1.000	30.000	1.000	67.890	0.000
13	0.757	0.016	0.008	2.005	1.001	87.926	0.999	2.890	0.000
14	0.727	0.018	0.008	2.120	1.001	86.707	0.999	3.000	0.000
15	0.713	0.017	0.008	2.000	1.001	92.335	0.999	2.890	0.000
16	0.691	0.009	0.006	1.352	1.001	105.499	0.999	2.330	19.050
17	0.673	0.027	0.008	3.184	1.001	98.354	1.000	2.330	0.000
18	0.672	0.026	0.008	3.171	1.001	98.217	0.999	2.330	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 8 reports the results from the simulated trades based on the Stop Long order strategy on good news releases. The Sharpe ratio column obviously shows that using the Stop Long order in a good news scenario gives the average negative return in every trading case. This is obvious evidence that using Stop Long order for good news is not a suitable choice as suggested from the previous section.

Table 8 also shows that using the stop order gives a high probability for sent order to be matched from the *%Matched* column. The lot size is small as expected. The reason is that the stop order has a high chance of order being matched, and preventing the position from being forced closed is a top priority to consider during the trade. Therefore, the sent order will have a small lot size as a trade-off with the frequently matched order. As shown in Table 8 that there is no case of position forced close from the sent order which is consistent with the small lot size strategy.

By comparing Table 8 with Table 9, which reports the results from the simulated trades based on the Stop Long order strategy on good news releases, it obviously shows that the Sharpe ratio obtained from using the short strategy in a good news scenario is improved in most of the trading cases. This is obvious evidence that using the short order during good news is a better choice than the long strategy. Table 9 also shows that the probability for sent order to be matched is high and the lot size is small as expected in many cases. Like what we see from Table 8, there is a very low chance of position forced close from the sent order. This can be concluded that using the stop order with an appropriate lot size during a good news announcement can prevent the forced close position.

Table 10 reports the results from the simulated trades based on the Limit Long order strategy on good news releases. The Sharpe ratio from using the Limit Long order in a good news scenario is significantly higher than the Sharpe ratio from the stop order though the model suggests using the short strategy. The reason is that the optimal open with a large lot size is set at the place which is far from the current exchange rate which gives a better entry price if the order is matched. With this reason, the Limit Long strategy has a better entry price, if the order is matched, comparing with the stop order strategy but it has a lower chance of order to be matched as a trade-off. Using the Limit Long order is more likely to cause a position forced close as we can observe from the *%Forced* column which is absolutely higher than the stop order strategy.

By comparing Table 10 with Table 11, which reports the results from the simulated trades based on the Limit Short order strategy on good news releases, The Limit Short order strategy also has a significantly lower chance of position forced close than the Limit Long order strategy on average. Also, the average Sharpe ratio from the Limit Short order in a good news scenario is better than the average Sharpe ratio from the Limit Long order (See Table 16 for detail). This is consistent with the suggestion from the model that using short order may give a better Sharpe ratio. Using a limit order gives a significantly better Sharpe ratio than the stop order with a trade-off that the position forced close is more likely and a lower chance of sent order to be matched.

By comparing among all of the trading strategies, we find that using the Limit Short strategy, when a good news is released, gives the best result in terms of the Sharpe ratio with a low chance of order matching as a trade-off.

Table 12: In sample result for each bad news scenario using Stop Long order strategy trading on simulation

Stop Long Strategy									
Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	0.001	0.053	0.019	1.000	62.686	1.001	99.890	2.110
2	0.958	0.001	0.051	0.023	1.000	56.438	1.001	98.330	4.750
3	0.948	0.002	0.052	0.047	0.999	47.237	1.001	100.000	0.000
4	0.909	0.000	0.050	0.010	1.000	62.580	1.001	99.780	1.890
5	0.883	0.002	0.049	0.048	0.999	43.320	1.001	100.000	0.000
6	0.862	0.002	0.047	0.036	1.000	40.069	1.001	98.670	0.000
7	0.855	0.001	0.047	0.023	1.000	41.477	1.001	98.780	0.000
8	0.839	0.002	0.046	0.035	1.000	36.730	1.001	98.890	0.000
9	0.838	0.001	0.046	0.032	1.000	36.736	1.001	98.780	0.000
10	0.811	-0.001	0.044	-0.013	1.000	0.010	1.001	96.330	0.000
11	0.807	0.001	0.046	0.029	0.999	47.340	1.001	100.000	0.000
12	0.806	0.002	0.044	0.054	0.999	30.932	1.001	100.000	0.000
13	0.795	0.002	0.044	0.056	0.999	57.470	1.001	100.000	0.560
14	0.792	0.000	0.042	-0.003	1.000	0.010	1.001	96.330	0.000
15	0.777	0.000	0.045	0.003	1.000	64.671	1.001	99.890	1.330
16	0.776	0.001	0.043	0.012	1.000	31.762	1.001	98.780	0.000
17	0.775	0.000	0.045	0.004	1.000	61.746	1.001	99.890	0.670
18	0.756	0.001	0.044	0.013	1.000	62.874	1.001	100.000	0.560
19	0.739	0.000	0.042	0.012	1.000	37.578	1.001	98.560	0.000
20	0.720	0.002	0.042	0.041	1.000	37.666	1.001	100.000	0.000
21	0.689	0.000	0.040	0.010	1.000	34.804	1.001	99.220	0.000
22	0.684	0.001	0.042	0.035	0.999	51.924	1.001	100.000	0.000
23	0.670	0.002	0.040	0.041	1.000	44.188	1.001	99.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 13: In sample result for each bad news scenario using Stop Short order strategy trading on simulation

Stop Short Strategy									
Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	-0.012	0.050	-0.231	1.000	0.010	0.999	96.780	0.000
2	0.958	-0.012	0.049	-0.241	1.000	0.010	0.999	97.670	0.000
3	0.948	-0.007	0.038	-0.184	1.000	0.020	1.000	97.440	0.000
4	0.909	-0.011	0.049	-0.231	1.000	0.010	0.999	100.000	0.000
5	0.883	-0.010	0.022	-0.478	1.000	0.021	1.000	78.560	0.000
6	0.862	-0.012	0.046	-0.266	1.000	0.010	0.999	97.670	0.000
7	0.855	-0.012	0.045	-0.266	1.000	0.010	0.999	97.330	0.000
8	0.839	-0.012	0.045	-0.262	1.000	0.010	0.999	97.670	0.000
9	0.838	-0.011	0.003	-3.853	0.999	0.010	0.999	5.110	0.000
10	0.811	-0.011	0.044	-0.262	1.000	0.010	0.999	97.780	0.000
11	0.807	-0.009	0.034	-0.258	1.000	0.020	1.000	92.220	0.000
12	0.806	-0.010	0.038	-0.274	1.000	0.018	1.000	91.780	0.000
13	0.795	-0.022	0.024	-0.915	1.000	0.010	0.999	23.780	0.000
14	0.792	-0.014	0.043	-0.318	1.000	0.010	0.999	93.330	0.000
15	0.777	-0.011	0.044	-0.249	1.000	0.010	0.999	99.000	0.000
16	0.776	-0.017	0.037	-0.452	1.000	0.010	0.999	70.780	0.000
17	0.775	-0.011	0.043	-0.259	1.000	0.010	0.999	98.780	0.000
18	0.756	-0.008	0.034	-0.232	1.000	0.018	1.000	95.000	0.000
19	0.739	-0.010	0.042	-0.244	1.000	0.010	0.999	99.440	0.000
20	0.720	-0.014	0.005	-2.753	0.999	0.010	0.999	3.560	0.000
21	0.689	-0.006	0.025	-0.252	1.000	0.036	1.000	99.220	0.000
22	0.684	-0.010	0.040	-0.259	1.000	0.010	0.999	99.000	0.000
23	0.670	-0.011	0.039	-0.281	1.000	0.010	0.999	97.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 14: In sample result for each bad news scenario using Limit Long order strategy trading on simulation

Limit Long Strategy									
Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	0.019	0.015	1.296	0.999	96.704	1.001	6.330	19.300
2	0.958	0.021	0.014	1.493	0.999	93.868	1.001	5.890	11.320
3	0.948	0.022	0.014	1.596	0.999	94.628	1.001	5.780	9.620
4	0.909	0.012	0.009	1.453	0.999	89.336	1.001	5.000	0.000
5	0.883	0.014	0.009	1.672	0.999	93.504	1.001	4.890	2.270
6	0.862	0.021	0.009	2.179	0.999	93.966	1.001	4.000	2.780
7	0.855	0.013	0.010	1.318	0.999	97.817	1.001	4.110	18.920
8	0.839	0.014	0.011	1.256	0.999	107.242	1.001	4.000	27.780
9	0.838	0.014	0.011	1.257	0.999	107.258	1.001	4.000	27.780
10	0.811	0.018	0.010	1.733	0.999	90.478	1.001	3.890	2.860
11	0.807	0.021	0.010	2.116	0.999	89.655	1.001	3.670	0.000
12	0.806	0.017	0.010	1.636	0.999	101.529	1.001	3.670	21.210
13	0.795	0.021	0.010	2.119	0.999	104.848	1.001	3.440	12.900
14	0.792	0.022	0.010	2.173	0.999	105.068	1.001	3.440	12.900
15	0.777	0.021	0.010	1.996	0.999	105.703	1.001	3.670	15.150
16	0.776	0.021	0.010	1.997	0.999	105.729	1.001	3.670	15.150
17	0.775	0.021	0.010	2.080	0.999	105.715	1.001	3.670	12.120
18	0.756	0.018	0.010	1.751	0.999	117.161	1.001	3.220	41.380
19	0.739	0.031	0.011	2.896	0.999	106.484	1.001	2.890	3.850
20	0.720	0.024	0.008	2.902	0.999	110.301	1.001	2.330	4.760
21	0.689	0.038	0.008	4.707	0.999	110.323	1.001	1.890	0.000
22	0.684	0.031	0.008	4.045	0.999	111.165	1.001	1.890	0.000
23	0.670	0.039	0.007	5.917	0.999	120.000	1.000	1.220	9.090

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 15: In sample result for each bad news scenario using Limit Short order strategy trading on simulation

Limit Short Strategy									
Scenario	*Scale Ratio*	*Currency Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	0.020	0.016	1.232	1.001	80.502	0.999	8.000	1.390
2	0.958	0.021	0.018	1.187	1.001	86.571	0.999	7.780	7.140
3	0.948	0.023	0.019	1.253	1.001	90.181	0.999	7.780	11.430
4	0.909	0.025	0.018	1.416	1.001	86.689	0.999	7.000	3.170
5	0.883	0.026	0.017	1.527	1.001	85.093	0.999	7.110	0.000
6	0.862	0.025	0.017	1.429	1.001	91.977	0.999	6.670	10.000
7	0.855	0.026	0.017	1.548	1.001	92.158	0.999	6.220	8.930
8	0.839	0.010	0.013	0.732	1.001	116.010	0.999	5.560	48.000
9	0.838	0.010	0.013	0.730	1.001	115.960	0.999	5.560	48.000
10	0.811	0.014	0.013	1.061	1.001	111.033	0.999	4.780	30.230
11	0.807	0.009	0.013	0.720	1.001	120.000	0.999	4.780	51.160
12	0.806	0.028	0.014	1.933	1.001	90.098	0.999	4.670	2.380
13	0.795	0.015	0.012	1.196	1.001	102.709	0.999	4.780	16.280
14	0.792	0.015	0.012	1.270	1.001	100.312	0.999	4.780	9.300
15	0.777	0.021	0.011	1.805	1.001	100.651	0.999	4.110	8.110
16	0.776	0.024	0.012	2.061	1.001	100.619	0.999	4.110	5.410
17	0.775	0.023	0.012	1.988	1.001	100.725	0.999	4.220	5.260
18	0.756	0.019	0.011	1.670	1.001	112.883	0.999	4.000	27.780
19	0.739	0.027	0.012	2.185	1.001	110.127	0.999	3.560	21.880
20	0.720	0.023	0.010	2.235	1.001	97.217	0.999	3.560	3.130
21	0.689	0.029	0.011	2.687	1.001	111.036	0.999	3.440	16.130
22	0.684	0.026	0.010	2.485	1.001	111.723	0.999	3.670	12.120
23	0.670	0.029	0.010	2.764	1.001	111.224	0.999	3.440	9.680

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return, which is calculated by averaging the currency return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 12 reports the results from the simulated trades based on the Stop Long order strategy on bad news releases. It obviously shows from the Sharpe ratio column that using the Stop Long order in a bad news scenario gives the average positive Sharpe ratio in most of the trading simulation cases with a high probability for sent order to be matched as can be seen from the *%Matched* column. This is expected from the parameter values of 2-MRJ model reported in Table 7. The lot size is reasonable for the initial wealth of $10,000 because the maximum lot size in this case, with 2,000 times leverage, is 200 lots.

As expected, with the higher lot size, the chance of position forced closure is increased. This can be observed from the *%Forced* column in Table 8 and Table 9 as a comparison.

As shown in some trading cases from Table 12, we find that using the optimal lot size from optimization has a low chance of position forced closure. Although, there is a chance that the position forced closure condition is activated but we still have a positive return on average. This is because most of the trades are profitable, except the cases that the positions are forced closed, and the profits are able to fully cover the losses from the position forced close.

Table 13 reports the results from the simulated trades based on the Stop Short order strategy on bad news releases. We can see that most trading simulation cases have negative returns. This suggests that it is better not-to-trade for those cases, or equivalently, to hold the cash in those scenarios. However, using a small lot size can prevent a huge loss occurred from the position forced closure.

By comparing from both strategies (the Stop Long order strategy and the Stop Short order strategy), the Stop Long order give a better result in trading (in terms of the Sharpe ratio) than the Stop Short order strategy when the bad news is released to the market but there still are some scenarios that the position forced closure can be triggered with a small chance. The main reason is that the optimal lot size from the Stop Short order strategy is significantly lower than that of the Stop Long order strategy.

Table 14 reports the results from the simulated trades based on the Limit Long order strategy on bad news releases. We can see from the Sharpe ratio column that using the Limit Long order in the bad news scenarios give the higher positive Sharpe ratio in every trading case with a lower probability of sent order to be matched than the Stop Long strategy from Table 12. However, the chance of position forced closure is significantly increased. We can observe from the *%Forced* column in Table 12 as a comparison.

By comparing Table 14 with Table 15 which reports the results from the simulated trades based on the Limit Short order strategy on bad news releases, the Limit Long order strategy also has a significantly lower chance of position forced close than the Limit Short order strategy on average. Also, the average Sharpe ratio from the Limit Long order in a bad news scenario is better than the average Sharpe ratio from the Limit Short order (See Table 17 for detail). This is consistent with the suggestion from the parameter values of the model from Table 7 that using a long order may give a better Sharpe ratio than a short order.

By Comparing Table 12 with Table 15, we can observe that using the Stop Long order (shown in Table 12) gives a worse Sharpe ratio than the Limit Short order which conflicts to the suggestion from the parameter values in the 2-MRJ model discussed from Table 7 in the previous section that using a long strategy may give a better Sharpe ratio. The reasons are that the entry price of the limit order is better, if the order is matched, and that the lot size is larger than the stop order. Although the percentage of orders being matched is lower, it results in a higher Sharpe ratio for the limit order.

Table 8 : Summary trading simulation statistic for In-sample trading case using 2-regimes pure diffusion with jump (2-PDJ) with a good news announcement

Good News	2-Regimes Pure Diffusion			
Strategy	Stop Long	Stop Short	Limit Long	Limit Short
Total Scenario	18	18	18	18
Scenarios with Loss Trade in simulation	18	6	0	1
Simulation paths for each scenario	5,000	5,000	5,000	5,000
Average chance of order matched (%)	91.99%	97.83%	4.49%	7.17%
Average currency return given trade (%)	-0.00827	0.00031	0.00972	0.01224
STD currency return given trade (%)	0.03486	0.04508	0.01009	0.00985
Average Sharpe Ratio	-0.23723	0.00688	0.96333	1.24264
Forced Closure given trade (%)	0.00%	0.03%	7.78%	2.01%
Average Lot Size	0.016	20.747	89.237	84.480

Table 16 shows the summary result of trading simulation for good news with 2-PDJ model for each strategy using the optimal open prices, target prices and lot sizes. It can be seen that the Sharpe ratio (given there is a trade) of both limit order strategies give the better result than the stop order strategies in the trading simulation. However, the chance for the sent order to be matched is very low. With a low chance of order matching, the limit order strategy is not a good choice for the out-of-sample trading due to the limitation of the numbers of news. Also, we find that using the limit order is more likely to face with the higher chances for forced closure than the stop order. Therefore, using the Stop Short order is preferred with these reasons.

Table 9 Summary trading simulation statistic for In-sample trading case using 2-Regimes Mean-Reverting with Jump (2-MRJ) with a bad news announcement

Bad News	2-Regimes Mean-Reverting with Jump Diffusion			
Strategy	Stop Long	Stop Short	Limit Long	Limit Short
Total Scenario	23	23	23	23
Scenarios with Loss Trade in simulation	2	23	0	0
Simulation paths for each scenario	5,000	5,000	5,000	5,000
Average chance of order matched (%)	99.22%	83.90%	3.76%	5.20%
Average currency return given trade (%)	0.00112	-0.01143	0.02152	0.02116
STD currency return given trade (%)	0.04524	0.03646	0.01023	0.01360
Average Sharpe Ratio	0.02483	-0.31364	2.10320	1.55522
Forced Closure given trade (%)	0.52%	0.00%	11.79%	15.52%
Average Lot Size	43.05	0.01	102.54	101.11

Table 17 shows the summary result of trading simulation for bad news with 2-MRJ model for each strategy using the optimal open prices, target prices and lot sizes. Similar to the good news case, the Sharpe ratios (given there is a trade) of both limit order strategies are significantly better than those of the stop order strategies in the trading simulation. However, the chance for the sent order to be matched is also very low. Therefore, the Stop Long order is preferred with the same reasons.

5.3 Out-of-Sample Trading Result

This section shows the out-of-sample trading results of the good and bad news for each scenario with benchmark returns. These results are based on actual bid-ask quotes from the market.

In each scenario, the bid-ask quote of the exchange rate will be normalized using the mid-price of EUR/USD at the news announcement time by:

$$Normalized\ Bid_t = \frac{Bid\ price_t}{Mid\ price_0}$$
$$Normalized\ Ask_t = \frac{Ask\ price_t}{Mid\ price_0}$$

Each table reports the detailed out-of-sample trading results for the chosen strategy compared with 3 open-and-hold benchmarks. It provides the volatility scale ratio (Scale Ratio), the currency return for each strategy (Currency Return), optimal open prices (Optimal Open), optimal target prices (Optimal Target) , actual open price for each strategy (Opened Price), optimal lot sizes (Optimal Lot) and strategy that the Mixed benchmark use to trade in each scenario (Strategy). In some cases, the value will be assigned as “N/A” which indicates that the order is not submitted due to a negative mean return from the trading simulation.

There are 18 scenarios for the good news announcement and 23 scenarios for the bad news announcement. Each scenario will be sorted in descending order by the volatility scale ratio in *Scale Ratio* column.

Each scenario is sorted in descending order by the volatility scale ratio in *Scale Ratio* column. Table 18 reports the results from the out-of-sample trades using the Stop Short order strategy with 3 open-and-hold benchmarks for good news scenarios. Table 20 reports the results from the out-of-sample trades using the Stop Long order strategy with 3 open-and-hold benchmarks for bad news scenarios. There are also Table 19 and Table 21 which show the summary result which compare the suggested strategy with the benchmark in out-of-sample trading. (summarized from Tables 18 and 20 respectively)

Table 18: Detailed out-of-sample trading for good news using 2-PDJ model with Stop Short order and three benchmark strategies

Good						2-Regimes Pure Diffusion with Jump (2-PDJ)				Benchmark Strategy			
News		*Currency Return (%)*	*Currency Return (%)*	*Currency Return (%)*	*Currency Return (%)*					*Always Short*	*Always Long*	*Mixed*	
	Scale Ratio	*Stop Short order Strategy*	*Always Short*	*Always Long*	*Mixed*	*Optimal Open*	*Optimal Target*	*Opened Price*	*Optimal Lot Sizes*	*Opened Price*	*Opened Price*	*Opened Price*	*Strategy*
1	1.1116	N/A	0.077	-0.085	0.077	N/A	N/A	N/A	N/A	0.99998	1.00002	0.99998	SHORT
2	0.9775	N/A	-0.101	0.096	0.096	N/A	N/A	N/A	N/A	0.99999	1.00001	1.00001	LONG
3	0.9182	N/A	0.049	-0.057	0.049	N/A	N/A	N/A	N/A	0.99998	1.00002	0.99998	SHORT
4	0.8809	-0.041	-0.041	0.035	0.035	0.99998	0.99900	0.99998	24.33	0.99998	1.00002	1.00002	LONG
5	0.8602	-0.023	-0.023	0.011	0.011	0.99999	0.99900	0.99998	32.99	0.99998	1.00002	1.00002	LONG
6	0.8424	-0.005	-0.005	-0.003	-0.003	0.99995	0.99900	0.99998	57.31	0.99998	1.00002	1.00002	LONG
7	0.8419	-0.015	-0.015	0.010	0.010	1.00002	0.99900	0.99998	58.70	0.99998	1.00002	1.00002	LONG
8	0.8398	0.010	0.010	-0.015	0.010	1.00004	0.99900	0.99999	33.03	0.99999	1.00001	0.99999	SHORT
10	0.8142	-0.019	-0.019	0.012	0.012	1.00001	0.99900	0.99998	32.81	0.99998	1.00002	1.00002	LONG
11	0.7946	-0.002	-0.002	-0.002	-0.002	1.00005	0.99900	1.00000	32.85	1.00000	1.00000	1.00000	LONG
12	0.7830	-0.059	-0.059	0.008	0.008	1.00005	0.99900	0.99985	32.30	0.99985	1.00015	1.00015	LONG
13	0.7569	0.004	0.004	-0.055	0.004	1.00000	0.99900	0.99991	32.10	0.99991	1.00009	0.99991	SHORT
14	0.7266	0.013	0.013	-0.020	0.013	1.00000	0.99906	0.99998	0.85	0.99998	1.00002	0.99998	SHORT
15	0.7130	N/A	-0.006	-0.005	-0.005	N/A	N/A	N/A	N/A	0.99997	1.00003	1.00003	LONG
16	0.6910	0.007	0.007	-0.013	0.007	1.00001	0.99906	0.99999	1.56	0.99999	1.00001	0.99999	SHORT
17	0.6734	N/A	-0.018	0.013	0.013	N/A	N/A	N/A	N/A	0.99999	1.00001	1.00001	LONG
18	0.6719	N/A	-0.038	0.035	0.035	N/A	N/A	N/A	N/A	0.99999	1.00001	1.00001	LONG

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return from using the Stop Short order strategy. *Currency Return (%) Always Short* shows the currency return from using the Always Short Strategy. *Currency Return (%) Always Long* shows the currency return from using the Always Long Strategy. *Currency Return (%) Mixed* shows the currency return from using the Mixed Long and Short Strategy. *Optimal Open* shows the optimal Open price. *Optimal Target* shows the optimal Closing price. *Opened Price* shows the actual normalized bid-ask quote price when news is announced. This depends on strategy: we use bid price for short and ask price for long positions. *Optimal Lot Sizes* shows the optimal lot sizes. *Strategy* shows the strategy that the Mixed Benchmark uses in each scenario. "N/A" in *Currency Return (%)* column means we do not submit an order because the chosen strategy suggests a negative mean return from the trading simulation.

Table 18 shows the details of out-of-sample trades for each good-news scenario. The volatility scale ratio (Δ) estimated by GARCH for each scenario is used to adjust the parameter from estimation step.

We can observe from Table 18 that the optimization suggests trading using the Stop Short order strategy for 12 out of 18 scenarios. This is because the in-sample trading simulation does not generate positive mean returns for the chosen strategy in those other 6 scenarios. The result from Table 18 also shows that 4 out of 6 scenarios that the optimization suggests to hold cash on hand give negative currency return for Always Short strategy. This shows the effectiveness of the forecasting performance from in-sample optimization that following the suggestion to hold cash in some trades are effective.

The Mixed Strategy column shows that it is more efficient if the benchmark uses both Long and Short Strategy depending on how the movement of the market price is. However, this strategy cannot be implemented as we did not know ahead of time what the direction of the price movement is.

Note that open prices at the first 10 seconds in all scenarios deviate a lot from the prices at the announcement time 0, and this makes the trades from the suggested strategy open at those open prices, which are the same as the open prices of Always Short strategy, instead of their proposed optimal entry prices.

Returns reported in Table 18 are currency returns and thus do not account for the leverage ratio and the optimal lot size. When the optimal lot size and leverage are used and the returns are calculated based on the trader's wealth, the magnitude of returns can be much larger. The results in terms of the returns of trader's wealth are reported in Table 32 in Appendix 5. With leverage, using the optimal lot size can reduce the loss from forced closure, but the overall risk from leverage can be huge.

Table 19: Out-of-sample trading summary result using Stop Short order with 2-PDJ model for good news announced and three benchmark strategies

Good News	Out-Of-Sample			
Strategy : Stop Short order	Stop Short	Always Short	Always Long	Mixed
Total News	18	18	18	18
Number of Trading Scenarios	12	18	18	18
Currency Average return (%)	-0.00825	-0.00756	-0.00628	0.02228
STD Currency return (%)	0.02378	0.03866	0.04214	0.02687
Sharpe Ratio	-0.34697	-0.19542	-0.14896	0.82917
Percent Win given trade (%)	41.67	38.89	44.44	83.33
Force Close Occurrence given trade (%)	0.00%	5.6%	0.00%	0.00%
Number of Out-of-sample loss Trade	7	11	10	3
Forecasting Performance (%)	67%	N/A	N/A	N/A

Table 19 shows the summary results of the out-of-sample trades using the Stop Short order with 2-PDJ model, and the benchmark strategies.

The Stop Short order strategy column shows the trading performance of using the Stop Short order strategy. We can see that given a trade occurs, this strategy has 41.67% of winning rate without any chance of the position forced closure.

To measure the forecasting performance when the in-sample trading results predict losses in a given scenario, we introduce the forecasted performance row. This row reports the percentage of loss scenarios if the Always Short strategy is used given the in-sample trading simulation suggests no trade (Table 9). In this case, we have 67% accuracy so the suggestion from the in-sample result to hold cash in those scenarios is effective.

Comparing all strategies, we can observe that the Sharpe ratio of using the Stop Short order strategy underperform all of the benchmarks. Table 34 in Appendix 5 reports the summary results with 2,000-time leverage and optimal lot size. As expected, the return and standard deviation are much larger in magnitude. It turned out that the Sharpe ratio with the leverage ratio is worse (more negative).

Table 20: Detailed out-of-sample trading for bad news using 2-MRJ model with Stop Long order and three benchmark strategies

Bad News	Scale Ratio	Currency Return (%) Stop Long order Strategy	Currency Return (%) Always Short	Currency Return (%) Always Long	Currency Return (%) Mixed	2-Regimes Mean-Reverting with Jump (2-MRJ) Optimal Open	Optimal Target	Opened Price	Optimal Lot Sizes	Benchmark Strategy Always Short Opened Price	Always Long Opened Price	Mixed Opened Price	Mixed Strategy
1	1.0213	0.011	-0.019	0.011	0.011	0.99971	1.00100	1.00002	62.68	0.99998	1.00002	1.00002	LONG
2	0.9577	0.034	-0.040	0.034	0.034	1.00000	1.00100	1.00001	56.43	0.99999	1.00001	1.00001	LONG
3	0.9484	-0.033	0.028	-0.033	0.028	0.99936	1.00100	1.00002	47.23	0.99998	1.00002	0.99998	SHORT
4	0.9090	0.000	-0.008	0.000	0.000	0.99994	1.00100	1.00002	62.57	0.99998	1.00002	1.00002	LONG
5	0.8834	0.012	-0.063	0.012	0.012	0.99900	1.00100	1.00002	43.32	0.99998	1.00002	1.00002	LONG
6	0.8615	-0.025	0.024	-0.025	0.024	1.00000	1.00100	1.00002	40.06	0.99998	1.00002	0.99998	SHORT
7	0.8546	-0.026	0.019	-0.026	0.019	1.00000	1.00100	1.00002	41.47	0.99998	1.00002	0.99998	SHORT
8	0.8388	0.002	-0.011	0.002	0.002	0.99999	1.00100	1.00001	36.72	0.99999	1.00001	1.00001	LONG
9	0.8380	0.002	-0.011	0.002	0.002	1.00000	1.00100	1.00002	36.73	0.99998	1.00002	1.00002	LONG
10	0.8111	N/A	-0.044	-0.002	-0.002	N/A	N/A	N/A	N/A	0.99991	1.00009	1.00009	LONG
11	0.8066	0.008	-0.016	0.008	0.008	0.99924	1.00100	1.00002	47.34	0.99998	1.00002	1.00002	LONG
12	0.8058	-0.024	0.016	-0.024	0.016	0.99900	1.00100	1.00003	30.93	0.99997	1.00003	0.99997	SHORT
13	0.7947	-0.003	-0.014	-0.003	-0.003	0.99900	1.00067	1.00001	57.46	0.99999	1.00001	1.00001	LONG
14	0.7922	N/A	-0.056	0.028	0.028	N/A	N/A	N/A	N/A	0.99985	1.00015	1.00015	LONG
15	0.7770	0.012	-0.022	0.012	0.012	0.99983	1.00100	1.00015	64.67	0.99985	1.00015	1.00015	LONG
16	0.7764	-0.034	0.028	-0.034	0.028	1.00000	1.00100	1.00000	31.76	1.00000	1.00000	1.00000	SHORT
17	0.7752	-0.007	-0.003	-0.007	-0.003	0.99990	1.00100	1.00000	61.74	1.00000	1.00000	1.00000	SHORT
18	0.7564	-0.004	-0.003	-0.004	-0.003	0.99969	1.00100	0.99999	62.87	1.00001	0.99999	1.00001	SHORT
19	0.7389	-0.009	-0.008	-0.009	-0.008	1.00000	1.00100	1.00001	37.57	0.99999	1.00001	0.99999	SHORT
20	0.7195	0.000	-0.007	0.000	0.000	0.99966	1.00100	1.00001	37.66	0.99999	1.00001	1.00001	LONG
21	0.6885	-0.030	0.027	-0.030	0.027	0.99999	1.00100	0.99999	34.80	1.00002	0.99998	1.00002	SHORT
22	0.6838	-0.010	-0.001	-0.010	-0.001	0.99901	1.00077	0.99999	51.92	1.00001	0.99999	1.00001	SHORT
23	0.6695	-0.006	0.002	-0.006	0.002	0.99979	1.00100	0.99997	44.18	1.00003	0.99997	1.00003	SHORT

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Currency Return (%)* shows the currency return from using the Stop Short order strategy. *Currency Return (%) Always Short* shows the currency return from using the Always Short Strategy. *Currency Return (%) Always Long* shows the currency return from using the Always Long Strategy. *Currency Return (%) Mixed* shows the currency return from using the Mixed Long and Short Strategy. *Optimal Open* shows the optimal Open price. *Optimal Target* shows the optimal Closing price. *Opened Price* shows the actual normalized bid-ask quote price when news is announced. This depends on strategy: we use bid price for short and ask price for long positions. *Optimal Lot Sizes* shows the optimal lot sizes. *Strategy* shows the strategy that the Mixed Benchmark uses in each scenario. "N/A" in *Currency Return (%)* column means we do not submit an order because the chosen strategy suggests a negative mean return from the trading simulation.

Table 20 shows the details for each bad-news scenario. There are 21 out of 23 scenarios that the optimization result suggests to trade. For most of the trading scenarios, the Long order can gain an advantage on average in 15-minute trading as we can observe from the Always Long strategy that the *Currency Return* column are positive in most scenarios.

The *Optimal Open* for the Stop Long strategy does not provide any advantage for these scenarios as shown from the *Opened price* column that in all the scenarios, the best ask price jumps across the optimal open price when news is announced. Therefore, the Stop Long order is executed immediately at the best ask price which is not the optimal price from optimization.

In the bad news scenarios in Table 20, we can also see from the *Currency Return* column of the *Mixed* strategy that some scenarios give negative return although we can look ahead of time and choose the best strategy to open the position. This indicates that it is better to hold cash than to trade in some scenarios because neither Long order nor Short order is profitable.

Table 33 in Appendix 5 reports the results of the Stop Long strategy with optimal lot size and 2,000-time leverage. The returns are much larger in the magnitude as expected. With the optimal lot size, all trades can avoid the force closure at the large leverage ratio.

Table 21: Out-of-sample trading summary result using Stop Short order with 2-MRJ model for bad news announced and three benchmark strategies

Bad News	Out-Of-Sample			
Strategy : Stop Long order	Stop Long	Always Short	Always Long	Mixed
Total News	23	23	23	23
Number of Trading Scenarios	21	23	23	23
Currency Average return (%)	-0.00619	-0.00791	-0.00452	0.01013
STD Currency return (%)	0.01711	0.02505	0.01778	0.01260
Sharpe Ratio	-0.36188	-0.31592	-0.25437	0.80410
Percent Win given trade (%)	33.33%	30.43%	34.78%	65.22%
Force Close Occurrence given trade (%)	0.00%	0.00%	0.00%	0.00%
Number of Out-of-sample loss Trade	12	13	15	6
Forecasting Performance (%)	50%	N/A	N/A	N/A

Table 21 shows the summary results for the out-of-sample trades using Stop Long order strategy with 2-MRJ model and the benchmark strategies.

We can see that the Sharpe ratio of using the Stop Long order strategy is the worst comparing with all other strategies. However, all strategies (except the Mixed Strategy) give the negative Sharpe ratios. The reason that the Sharpe ratio of the Stop Long order strategy is lower than the Always Long strategy mainly comes from avoiding the trade in scenario 14 in Table 20 which is quite a high profit scenario comparing with others. In the bad news announcement, the forecasting performance for holding cash is 50% which is equal to tossing a fair coin. Comparing with the good news case, this suggestion is worse.
We report the results of the Stop Long strategy with 2,000-time leverage and optimal lot size case by case in Table 33 and report the summary result in Table 35 in Appendix 5. As

expected, the returns and standard deviations are much larger in magnitude, but the Sharpe ratio improves a little (-0.250).

Chapter 6 Conclusion and Further Study

The objective of this thesis is to find the suitable trading strategy with the optimal solution for trading the EUR/USD currency in the Unemployment claims announcement period.

This thesis presents the EUR/USD price dynamic, and an estimation method based on the EM algorithm. The models that best describe the dynamics of EUR/USD data during the announcement periods for the good news and bad news based on AIC and BIC are the two-regime pure-diffusion with jumps (2-PDJ) and two-regime mean-reversion with jumps (2-MRJ) respectively.

The parameter values from the model estimation show that the result of the news announcement and the dynamic of the EUR/USD prices are inconsistent. More precisely, the model suggests us to open a short position after a release of good news, and to open a long position when the news result turns out to be bad.

For the good news case, the estimated model suggests that prices are trending downward with no mean reversion. Although, the best strategy in our scope is the Limit Short order strategy judged by the Sharpe ratio, there is a trade-off for using this strategy which is the chance of order matching is significantly low compared with the stop order strategy. Therefore, the Stop Short order strategy is preferred for this situation due to the data limitation. Also, the optimal lot size from this strategy can help prevent the traders' portfolio from the forced closure.

In the out-of-sample cases, our strategy and the benchmarks (except for the Mixed strategy) have the negative Sharpe ratio and our strategy underperform all of the benchmarks. When the large leverage ratio is applied to our strategy, the magnitude of the risk to trader's wealth is increased. The negative Sharpe ratio is also worse after the large leverage ratio is applied. In terms of preventing the position forced closure, the optimal lot size for the Stop Short order strategy does a great job, since this strategy can perfectly prevent the position forced closure for the out-of-sample data.

For the bad news case, the estimated parameter values suggest that there is a mean reversion effect with a fast reverting speed to a mean value higher than the announced currency price. Although, the Limit Buy order strategy is obviously a good strategy to trade during this situation if the price goes down before it reverts back up to its mean, the trade-off of using this strategy is that the chance of order matching is low. With this reason, the Stop Long order comes as an alternative choice. Due to the fast converging to the mean, there is a small trend from the entry price toward the mean if the current price is significantly lower than the mean, so using the Stop Long order will gain a small trending benefit while the currency price is converging to its mean value.

In the out-of-sample trading, the result turns out to be a loss on average and our strategy underperforms all of the benchmarks. When the large leverage ratio is applied to our strategy, it does a great adjustment in the lot size applied in trading since the Sharpe ratio is improved, and the forced closure condition is not activated. With these reasons, we can conclude that using optimal lot size in real trading with the large leverage ratio can help traders to improve their trading performance and prevent their loss from forced closure condition. However, the current strategy still yields a negative mean return.

Since, the Mixed strategy has a look-ahead bias and cannot be implemented in reality. We use it to provide a high-performance benchmark for our comparison. In some scenarios, using the Mixed strategy gives a negative currency return. This shows that in some cases using neither long nor short strategy is the best. So the trading performance can be improved if we avoid some trades. Also, instead of trading using one strategy in all the cases, it is better to allow strategy to be adaptive to the current situation to improve the trading performance. We suggest to test on the combination with other strategies or to apply a longer trading period for a further study.

REFERENCES

Acquah, H. d.-G. (2010). "Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in Selection of an Asymmetric Price Relationship." Journal of Development and Agricultural Economics **2**(1): 1-6.

Allen, H. and M. P. Taylor (1990). "Charts, Noise and Fundamentals in the London Foreign Exchange Market." The Economic Journal **100**(400): 49-59.

Ball, C. A. and W. N. Torous (1985). "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing." Jounal of Finance **40**(1): 155-173.

Beckers, S. (1981). "A Note on Estimating the Parameters of the Diffusion-Jump Model of Stock Returns." Jounal of Financial and Quantitative Analysis **16**(1): 127-140.

Biais, B. (1993). "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets." Journal of Finance **48**(1): 157-185.

Bollerslev, T. and M. Melvin (1994). "Bid—Ask Spreads and Volatility in the Foreign Exchange Market: An Empirical Analysis." Journal of International Economics **36**(3): 355-372.

Clewlow, L. and C. Strickland (2000). Energy Derivatives: Pricing and Risk Management, Lacima Publ.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977). "Maximum Likelihood from Incomplete Data via the EM Algorithm." Journal of the Royal Statistical Society. Series B (Methodological) **39**(1): 1-38.

Engel, C. and J. D. Hamilton (1990). "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" The American Economic Review **80**(4): 689-713.

Evans, M. D. and R. K. Lyons (2002). "Order Flow and Exchange Rate Dynamics." Journal of Political Economy **110**(1): 170-180.

Frankel, J. A. and K. A. Froot (1985). "Using Survey Data to Test some Standard Propositions Regarding Exchange Rate Expectations." The American Economic Review **77**(1): 133-153.

Froot, K. A., D. S. Scharfstein and J. C. Stein (1990). "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation." Journal of Finance **47**(4).

Galati, G. and C. Ho (2001). Macroeconomic News and the Euro/Dollar Exchange Rate. BIS website (www.bis.org).

Glosten, L. R. and P. R. Milgrom (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." Journal of Financial Economics **14**(1): 71-100.

Goodman, S. H. (1979). "Foreign Exchange Rate Forecasting Techniques: Implications for Business and Policy." Journal of Finance **34**(2): 415-427.

Johnson, G. and T. Schneeweis (1994). "Jump-Diffusion Processes in the Foreign Exchange Markets and the Release of Macroeconomic News." Journal of Computational Economics **7**(4): 309-329.

Levich, R. M. and L. R. Thomas (1993). "The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach." Journal of International Money and Finance **12**(5): 451-474.

Marsh, I. W. (2000). "High-Frequency Markov Switching Models in the Foreign Exchange Market." Journal of Forecasting **19**(2): 123-134.

Schulmeister, S. (1988). "Currency Speculation and Dollar Fluctuations." The Banca Nazionale del Lavoro Quarterly Review **41**(167): 343-365.

Springer, T. and K. Urban (2014). "Comparison of the EM Algorithm and Alternatives." The Numerical Algorithms Journal **67**(2): 335-364.

Stephane, G. and B. Zou (2011). Foreign exchange rates under Markov Regime switching model.

APPENDIX

Appendix 1: News Data Table

Table 22: Good news announced date

News Date	Announcement time(+7 GMT)	Result ($x10^3$)	Forecast ($x10^3$)
2010/01/21	20:30	482	441
2010/02/04	20:30	480	461
2010/04/08	19:30	460	434
2010/04/29	19:30	448	442
2010/05/20	19:30	471	439
2010/07/01	19:30	472	454
2010/07/22	19:30	464	449
2010/08/05	19:30	479	456
2010/08/12	19:30	484	465
2010/08/19	19:30	500	478
2010/09/23	19:30	465	451
2010/11/04	19:30	457	437
2010/12/02	20:30	436	425
2011/01/06	20:30	409	400
2011/03/31	19:30	388	379
2011/04/21	19:30	403	394
2011/05/05	19:30	474	415
2011/06/02	19:30	422	416
2011/06/23	19:30	429	414
2011/12/01	20:30	402	390
2011/12/29	20:30	381	372
2012/03/08	20:30	362	352
2012/03/29	19:30	359	351
2012/04/26	19:30	388	374
2012/06/21	19:30	387	381
2013/01/10	21:30	371	361
2013/01/31	21:30	368	362
2013/02/07	21:30	366	361
2013/03/28	20:30	357	340
2013/04/04	20:30	385	352
2013/06/20	20:30	354	343
2013/07/11	20:30	360	342
2013/08/22	20:30	336	329
2013/10/10	20:30	374	307
2013/12/19	21:30	379	336
2014/01/02	20:30	339	341
2014/01/30	20:30	348	329

Table 23: Bad news announced date

News Date	Announcement time(+7 GMT)	Result (x10^3)	Forecast (x10^3)
2010/01/07	20:30	434	449
2010/02/11	20:30	440	460
2010/03/25	19:30	442	452
2010/07/08	19:30	454	461
2010/08/26	19:30	473	488
2010/09/30	19:30	453	458
2010/10/07	19:30	445	454
2010/10/28	19:30	434	453
2010/12/09	20:30	421	426
2010/12/30	20:30	388	416
2011/01/20	20:30	404	422
2011/02/03	20:30	415	420
2011/02/10	20:30	383	411
2011/03/03	20:30	368	394
2011/05/19	19:30	409	421
2011/06/16	19:30	414	421
2011/07/28	19:30	398	413
2011/09/29	19:30	391	420
2011/10/06	19:30	401	411
2011/11/17	20:30	388	396
2011/12/08	20:30	381	397
2011/12/22	20:30	364	376
2012/02/02	20:30	367	373
2012/02/09	20:30	358	369
2012/03/22	19:30	348	353
2012/04/19	19:30	386	370
2012/05/03	19:30	365	381
2013/01/17	21:30	335	369
2013/01/24	21:30	330	359
2013/02/14	21:30	341	361
2013/03/21	20:30	336	343
2013/04/11	20:30	346	362
2013/04/25	20:30	339	352

Table 23(cont.): Bad news announced date

News Date	Announcement time(+7 GMT)	Result (x10^3)	Forecast (x10^3)
2013/05/09	20:30	323	333
2013/05/23	20:30	340	347
2013/07/18	20:30	334	344
2013/08/01	20:30	326	346
2013/09/12	20:30	292	332
2013/09/19	20:30	309	331
2013/09/26	20:30	305	319
2013/10/03	20:30	308	315
2013/12/26	21:30	338	346

Appendix 2: Deriving log complete likelihood

For deriving the likelihood function, it is obvious that

$$f(C_1, C_2, \dots, C_M) = P(C_1)P(C_2, C_3, \dots C_M|C_1)$$

where,

$$P(C_1) = P\left(y_1^{(1)}, \Delta N_1^{(1)}, \delta^{(1)}_{\Delta N_1^{(1)}}, z_1^{(1)}, y_2^{(1)}, \Delta N_2^{(1)}, \delta^{(1)}_{\Delta N_2^{(1)}}, z_2^{(1)}, \dots, y_T^{(1)}, \Delta N_T^{(1)}, \delta^{(1)}_{\Delta N_T^{(1)}}, z_T^{(1)}\right).$$

By recursively use of the properties of the conditional probability $P(A,B,C) = P(B,C|A)P(A)$
, the equation will be rearranged into

$$P(C_1) = P(y_1^{(1)})\left[\prod_{t=1}^{T} P(\Delta N_t^{(1)}|y_t^{(1)})P\left(\delta^{(1)}_{\Delta N_t^{(1)}}\middle|\Delta N_t^{(1)}, y_t^{(1)}\right)P(z_t^{(1)}|y_t^{(1)})\right]\prod_{t=1}^{T-1} P(y_{t+1}^{(1)}|y_t^{(1)}).$$

By recursively substitute to $f(C_1, C_2, \dots, C_M)$
, the complete likelihood function can be shown as follows:

$$\begin{aligned} f(C_1, C_2, \dots, C_M) &= P(C_1)P(C_2, C_3, \dots C_M|C_1) \\ &= P(C_1)P(C_2|C_1)P(C_3|C_1, C_2)\dots P(C_M|C_1, C_2, \dots, C_{M-1}) \end{aligned}$$

Because of the Markov properties,

The equation $f(C_1, C_2, \dots, C_M) = P(C_1)P(C_2|C_1)P(C_3|C_1, C_2)\dots P(C_M|C_1, C_2, \dots, C_{M-1})$ can be written as

$$\begin{aligned} f(C_1, C_2, \dots, C_M) &= P(C_1)P(C_2|C_1)P(C_3|C_2)\dots P(C_M|C_{M-1}) \\ &= \prod_{m=1}^{M}(P(y_1)^{(m)})\left\{\left[\prod_{t=1}^{T} P(\Delta N_t^{(m)}|y_t^{(m)})P\left(\delta^{(m)}_{\Delta N_t^{(m)}}\middle|\Delta N_t^{(m)}, y_t^{(m)}\right)P(z_t^{(m)}|y_t^{(m)})\right]\prod_{t=1}^{T-1} P(y_{t+1}^{(m)}|y_t^{(m)})\right\} \end{aligned}$$

Therefore, the log complete likelihood can be written as:

$$\begin{aligned} ln(f(C_1, C_2, \dots, C_M)|\Theta_r) = &\sum_{m=1}^{M} ln\left(P(y_1^{(m)})\right) + \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(\Delta N_t^{(m)}|y_t^{(m)})\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta^{(m)}_{\Delta N_t^{(m)}}\middle|\Delta N_t^{(m)}, y_t^{(m)}\right)\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(z_t^{(m)}|y_t^{(m)}, \Delta N_t^{(m)})\right) \\ &+ \sum_{m=1}^{M}\sum_{t=1}^{T-1} ln(P(y_{t+1}^{(m)}|y_t^{(m)})) \ , m = 1,2,3, \dots, M \end{aligned}$$

Appendix 3: Deriving M-step parameter

Derive the Initial State Probability ($\hat{\pi}$)

The parameter $\hat{\pi}_k$ can be derived by using the lagrange multiplier method with the constraint $\sum_{k=1}^{2} \pi_k = 1$ for the expected log complete likelihood $\mathbb{E}(ln(f(C_1, C_2, \dots, C_M)|\Theta_r)\,|\mathbb{X}, \Theta_{r-1})$.
Let $P(y_1^{(m)} = k) = \pi_k$. Therefore,

$$\mathbb{E}[\sum_{m=1}^{M} ln\left(P(y_1^{(m)})\right)|\mathbb{X},\Theta_{r-1}]$$
$$=\sum_{m=1}^{M}\sum_{k=1}^{2}\mathbb{E}\left[ln\left(P(y_1^{(m)}=k)\right)\middle|\mathbb{X},\Theta_{r-1},y_1^{(m)}=k\right]P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})$$
$$=\sum_{m=1}^{M}\sum_{k=1}^{2}ln\left(P(y_1^{(m)}=k)\right)P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})$$
$$=\sum_{m=1}^{M}\sum_{k=1}^{2}ln(\pi_k)\,P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})$$

$$\frac{\partial}{\partial\pi_k}\left[\mathbb{E}[\sum_{m=1}^{M} ln\left(P(y_1^{(m)})\right)|\mathbb{X},\Theta_{r-1}]-\left(\left(\sum_{k=1}^{2}\pi_k\right)-1\right)\mu\right] = 0$$

where, μ is a lagrange multiplier variable.

$$\sum_{m=1}^{M}\frac{P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})}{\pi_k}-\mu=0$$
$$\sum_{m=1}^{M}\frac{P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})}{\mu}=\hat{\pi}_k$$
$$\sum_{m=1}^{M}\sum_{k=1}^{2}\frac{P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})}{\mu}=\sum_{k=1}^{2}\pi_k=1$$
$$\sum_{m=1}^{M}\sum_{k=1}^{2}P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})=\mu=M$$
$$\sum_{m=1}^{M}\frac{P(y_1^{(m)}=k|\mathbb{X},\Theta_{r-1})}{M}=\hat{\pi}_k$$

Derive the Jump Intensity ($\hat{\lambda}$)

The parameter $\hat{\lambda}_k$ can be derived by using the first order derivative method with respect to λ_k and setting it equal to zero for the expected log complete likelihood $\mathbb{E}(ln(f(C_1,C_2,\dots,C_M)|\Theta_r)\,|\mathbb{X},\Theta_{r-1})$. First, we will consider the term $\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(\Delta N_t^{(m)}|y_t^{(m)})\right)|\mathbb{X},\Theta_{r-1}]$ which is the only term where the parameter λ_k is concealed in.

$$\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(\Delta N_t^{(m)}|y_t^{(m)})\right)|\mathbb{X},\Theta_{r-1}] = \sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}[ln\left(P(\Delta N_t^{(m)}|y_t^{(m)})\right)|\mathbb{X},\Theta_{r-1}]$$

where, $P\left(\Delta N_t^{(m)}\middle|y_t^{(m)}=k\right)$ is the Poisson distribution with the parameter λ_k. Therefore,

$$\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(\Delta N_t^{(m)}|y_t^{(m)})\right)|\mathbb{X},\Theta_{r-1}]$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}[ln\left(P(\Delta N_t^{(m)}|y_t^{(m)}=k)\right)|\mathbb{X},\Theta_{r-1}]$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}\left[ln\left(\prod_{k=1}^{2}\frac{\lambda_k^{\Delta N_t^{(m)}\mathbb{I}\{y_t^{(m)}=k\}}e^{-\lambda_k\mathbb{I}\{y_t^{(m)}=k\}}}{\Delta N_t^{(m)}!}\right)\middle|\mathbb{X},\Theta_{r-1}\right]$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{2}\mathbb{E}\,[\,\Delta N_t^{(m)}\mathbb{I}\{y_t^{(m)}=k\}ln(\lambda_k)-\lambda_k\mathbb{I}\{y_t^{(m)}=k\}-ln(\Delta N_t^{(m)}!)|\;\mathbb{X},\Theta_{r-1}]$$

Using the first order derivative method with respect to λ_k and setting it equal to zero, we can rearrange it into

$$\frac{\partial}{\partial\lambda_k}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\Delta N_t^{(m)}|y_t^{(m)}\right)\right)|\mathbb{X},\Theta_{r-1}]\right]=0$$
$$\sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}\left[\frac{\Delta N_t^{(m)}}{\lambda_k}\mathbb{I}\{y_t^{(m)}=k\}-\mathbb{I}\{y_t^{(m)}=k\}\middle|\mathbb{X},\Theta_{r-1}\right]=0$$
$$\hat{\lambda}_k=\frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\mathbb{E}[\Delta N_t^{(m)}|\mathbb{X},\Theta_{r-1},y_t^{(m)}=k]P(y_t^{(m)}=k|\mathbb{X},\Theta_{r-1})}{\sum_{m=1}^{M}\sum_{t=1}^{T}P(y_t^{(m)}=k|\mathbb{X},\Theta_{r-1})}$$

where, $\mathbb{E}[\mathbb{I}\{y_t^{(m)}=k\}|\mathbb{X},\Theta_{r-1}]=P(y_t^{(m)}=k|\mathbb{X},\Theta_{r-1})$

Derive the Mean of jump size ($\hat{\eta}$) and the Volatility of jump size ($\hat{\omega}$)

The parameter $\hat{\eta}_k$ and $\hat{\omega}_k^2$ can be derived by using the first order derivative method for the expected log complete likelihood $\mathbb{E}(ln(f(C_1,C_2,\dots,C_M)|\Theta_r)|\mathbb{X},\Theta_{r-1})$ with respect to η_k and ω_k^2 respectively. We will set each differentiated equation equal to zero.

We will consider the term $\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta_{\Delta N_t^{(m)}}^{(m)}\middle|\Delta N_t^{(m)},y_t^{(m)}\right)\right)|\mathbb{X},\Theta_{r-1}]$ which is the only term where the parameters η_k and ω_k^2 are concealed in.

$$\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta_{\Delta N_t^{(m)}}^{(m)}\middle|\Delta N_t^{(m)},y_t^{(m)}\right)\right)|\mathbb{X},\Theta_{r-1}]$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{2}\sum_{j=0}^{\infty}\mathbb{E}\left[ln\left(P\left(\delta_{t,i}^{(m)}\middle|\Delta N_t^{(m)}=j,y_t^{(m)}=k\right)\right)\middle|\mathbb{X},\Theta_{r-1}\right]P\left(\Delta N_t^{(m)}=j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k\middle|\mathbb{X},\Theta_{r-1}\right)$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{2}\sum_{j=0}^{\infty}\mathbb{E}\left[j\left(ln(2\pi)^{-\frac{1}{2}}-\frac{1}{2}ln(\omega_k^2)-\frac{1}{2\omega_k^2}\left(\delta_{t,i}^{(m)}-\eta_k\right)^2\right)\middle|\mathbb{X},\Theta_{r-1}\right]P(\Delta N_t^{(m)}=j|\mathbb{X},\Theta_{r-1},y_t^{(m)}=k)P\left(y_t^{(m)}=k\middle|\mathbb{X},\Theta_{r-1}\right)$$

where, $\delta_{t,i}^{(m)}|\Delta N_t^{(m)}=j,y_t^{(m)}=k\ \sim^{iid}\ Nor(\eta_k,\omega_k^2)$ for $i=1,2,3,\dots,j$
To find $\hat{\eta}_k$, we will apply the first order derivative method with respect to η_k and set it equal to zero, we can rearrange it into

$$\frac{\partial}{\partial\eta_k}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta_{\Delta N_t^{(m)}}^{(m)}\middle|\Delta N_t^{(m)},y_t^{(m)}\right)\right)|\mathbb{X},\Theta_{r-1}]\right]=0$$
$$\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{2}\sum_{j=0}^{\infty}\mathbb{E}\left[\frac{j}{\omega_k^2}(\delta_{t,i}^{(m)}-\eta_k)\middle|\mathbb{X},\Theta_{r-1}\right]P(\Delta N_t^{(m)}=j|\mathbb{X},\Theta_{r-1},y_t^{(m)}=k)P(y_t^{(m)}=k|\mathbb{X},\Theta_{r-1})=0$$
$$\hat{\eta}_k=\frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}j\mathbb{E}[\delta_{t,i}^{(m)}|\mathbb{X},\Theta,\Delta N_t^{(m)}=j,y_t^{(m)}=k]P(\Delta N_t^{(m)}=j|\mathbb{X},\Theta,y_t^{(m)}=k)P(y_t^{(m)}=k|\mathbb{X},\Theta)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}jP(\Delta N_t^{(m)}=j|\mathbb{X},\Theta,y_t^{(m)}=k)P(y_t^{(m)}=k|\mathbb{X},\Theta)}$$

To find $\hat{\omega}_k^2$, we will apply the first order derivative method with respect to ω_k and set it equal to zero. We can rearrange it into

$$\frac{\partial}{\partial\omega_k^2}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(\delta_{\Delta N_t^{(m)}}^{(m)}\middle|\Delta N_t^{(m)},y_t^{(m)}\right)\right)|\mathbb{X},\Theta_{r-1}]\right]=0$$
$$\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{2}\sum_{j=0}^{\infty}\mathbb{E}\left[j(-\frac{1}{2\omega_k^2}+\frac{1}{2(\omega_k^2)^2}\left(\delta_{t,i}^{(m)}-\eta_k\right)^2\middle|\mathbb{X},\Theta_{r-1}\right]P\left(\Delta N_t^{(m)}=j\middle|\mathbb{X},\Theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k\middle|\mathbb{X},\Theta_{r-1}\right)=0$$
$$\hat{\omega}_k^2=\frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}j\mathbb{E}\left[\left(\delta_{t,i}^{(m)}-\eta_k\right)^2\middle|\mathbb{X},\Theta,\Delta N_t^{(m)}=j,y_t^{(m)}=k\right]P\left(\Delta N_t^{(m)}=j\middle|\mathbb{X},\Theta,y_t^{(m)}=k\right)P\left(y_t^{(m)}=k\middle|\mathbb{X},\Theta\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}jP\left(\Delta N_t^{(m)}=j\middle|\mathbb{X},\Theta,y_t^{(m)}=k\right)P\left(y_t^{(m)}=k\middle|\mathbb{X},\Theta\right)}$$

Derive the Long-Term Mean ($\hat{\bar{\alpha}}$), the Mean Converging rate ($\hat{\kappa}$) and the Volatility ($\hat{\sigma}$)

The parameter $\hat{\bar{\alpha}}_k$, $\hat{\kappa}_k$ and $\hat{\sigma}_k^2$ can be derived by using the first order derivative method for the expected log complete likelihood $\mathbb{E}\left(ln\left(f\left(C_1, C_2, \dots, C_M\right)|\Theta_r\right) |\mathbb{X}, \Theta_{r-1}\right)$ with respect to $\bar{\alpha}_k$, κ_k and σ_k^2 respectively. We will set each differentiated equation equal to zero.

We will consider the term $\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(z_t^{(m)}|y_t^{(m)}\right)\right)|\mathbb{X}, \Theta_{r-1}]$ which is the only term where the parameters $\bar{\alpha}_k, \kappa_k$ and σ_k^2 are concealed in.

$$\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(z_t^{(m)}|y_t^{(m)},\Delta N_t^{(m)}\right)\right)|\mathbb{X},\theta_{r-1}]$$
$$=\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{k=1}^{Z}\sum_{j=0}^{\infty}\left\{\mathbb{E}\left[-\frac{1}{2}ln(2\pi)-\frac{1}{2}ln(\sigma_k^2)-\frac{1}{2\sigma_k^2}\left(z_t^{(m)}-\kappa_k\overline{\alpha_k}+\kappa_k x_t^{(m)}\right)^2\middle|\mathbb{X},\theta_{r-1},y_t^{(m)}=k,\Delta N_t^{(m)}=j\right]P\left(\Delta N_t^{(m)}=j|\mathbb{X},\theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k|\mathbb{X},\theta_{r-1}\right)\right\}$$

where, $z_t^{(m)}|\Delta N_t^{(m)}=j, y_t^{(m)}=k, \mathbb{X} \sim^{iid} Nor\left(\kappa_k\overline{\alpha_k}-\kappa_k x_t^{(m)},\sigma_k^2\right)$ for $j=1,2,3,\ldots$

To find $\hat{\kappa}_k$, we will apply the first order derivative method with respect to κ_k and set it equal to zero. We can rearrange it into

$$\frac{\partial}{\partial\kappa_k}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(z_t^{(m)}|y_t^{(m)},\Delta N_t^{(m)}\right)\right)|\mathbb{X},\theta_{r-1}]\right]=0$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}\left\{\mathbb{E}\left[-\frac{1}{\sigma_k^2}\left(z_t^{(m)}-\kappa_k\overline{\alpha_k}+\kappa_k x_t^{(m)}\right)x_t^{(m)}\middle|\mathbb{X},\theta_{r-1},y_t^{(m)}=k,\Delta N_t^{(m)}=j\right]P\left(\Delta N_t^{(m)}=j|\mathbb{X},\theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k|\mathbb{X},\theta_{r-1}\right)\right\}=0$$

$$\hat{\kappa}_k=\frac{\hat{\overline{\alpha}}_k\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}x_t^{(m)}P_{(j,t,k)}^{(m)}\right)-\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}x_t^{(m)}Z_t^{(m)}P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}x_t^{2(m)}P_{(j,t,k)}^{(m)}\right)}$$

where,

$$Z_t^{(m)}=\mathbb{E}[z_t^{(m)}|\mathbb{X},\theta_{r-1},\Delta N_t^{(m)}=j,y_t^{(m)}=k]$$
$$P_{(j,t,k)}^{(m)}=P\left(\Delta N_t^{(m)}=j|\mathbb{X},\theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k|\mathbb{X},\theta_{r-1}\right)$$

To find $\hat{\alpha}_k$, we will apply the first order derivative method with respect to $\overline{\alpha_k}$ and set it equal to zero. We can rearrange it into

$$\frac{\partial}{\partial\,\overline{\alpha_k}}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P\left(z_t^{(m)}|y_t^{(m)},\Delta N_t^{(m)}\right)\right)|\mathbb{X},\theta_{r-1}]\right]=0$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}\left\{\mathbb{E}\left[\frac{1}{\sigma_k^2}\left(z_t^{(m)}-\kappa_k\overline{\alpha_k}+\kappa_k x_t^{(m)}\right)\middle|\mathbb{X},\theta_{r-1},y_t^{(m)}=k,\Delta N_t^{(m)}=j\right]P\left(\Delta N_t^{(m)}=j|\mathbb{X},\theta_{r-1},y_t^{(m)}=k\right)P\left(y_t^{(m)}=k|\mathbb{X},\theta_{r-1}\right)\right\}=0$$

$$\hat{\alpha}_k = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} \mathbb{E}\left[(z_t^{(m)} + \hat{\kappa}_k x_t^{(m)}) \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] P(\Delta N_t^{(m)} = j | \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k) P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1})}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} P(\Delta N_t^{(m)} = j | \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k) P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1})}$$

By substituting, $\hat{\kappa}_k = \frac{\hat{\alpha}_k\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2(m)} P_{(j,t,k)}^{(m)}\right)}$ and $P_{(j,t,k)}^{(m)} = P\left(\Delta N_t^{(m)} = j \middle| \mathbb{X}, \theta_{r-1}, y_t^{(m)} = k\right) P\left(y_t^{(m)} = k \middle| \mathbb{X}, \theta_{r-1}\right)$, the equation can be written as:

$$\hat{\alpha}_k = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} \mathbb{E}\left[\left(z_t^{(m)} + \left(\frac{\hat{\alpha}_k\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2(m)} P_{(j,t,k)}^{(m)}\right)}\right) x_t^{(m)}\right) \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] P_{(j,t,k)}^{(m)}}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{n=0}^{\infty} P_{(j,t,k)}^{(m)}}$$

By isolating the parameter $\hat{\alpha}_k$ and rearranging the equation, the parameter $\hat{\alpha}_k$ can be written as follows:

$$\hat{\alpha}_k = \frac{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2(m)} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} Z_t^{(m)} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right)}{\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} P_{(j,t,k)}^{(m)}\right)\left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{2(m)} P_{(j,t,k)}^{(m)}\right) - \left(\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} x_t^{(m)} P_{(j,t,k)}^{(m)}\right)^2}$$

where,

$$Z_t^{(m)} = \mathbb{E}\left[z_t^{(m)} \middle| \mathbb{X}, \theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$$
$$P_{(j,t,k)}^{(m)} = P\left(\Delta N_t^{(m)} = j \middle| \mathbb{X}, \theta_{r-1}, y_t^{(m)} = k\right) P\left(y_t^{(m)} = k \middle| \mathbb{X}, \theta_{r-1}\right)$$

To find $\hat{\sigma}_k^2$, we will apply the first order derivative method with respect to σ_k^2 and set it equal to zero. We can rearrange it into

$$\frac{\partial}{\partial \overline{\alpha_k}}\left|\mathbb{E}\left[\sum_{m=1}^{M}\sum_{t=1}^{T} ln\left(P(z_t^{(m)} | y_t^{(m)}, \Delta N_t^{(m)})\right) | \mathbb{X}, \theta_{r-1}\right]\right| = 0$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}\left\{\mathbb{E}\left[-\frac{1}{2\sigma_k^2} + \frac{1}{2(\sigma_k^2)^2}\left(z_t^{(m)} - \kappa_k\overline{\alpha_k} + \kappa_k x_t^{(m)}\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] P(\Delta N_t^{(m)} = j | \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k) P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1})\right\} = 0$$

$$\hat{\sigma}_k^2 = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty}\left\{\mathbb{E}\left[\left(z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)})\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] P(\Delta N_t^{(m)} = j | \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k) P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1})\right\}}{\sum_{m=1}^{M}\sum_{t=1}^{T}\sum_{j=0}^{\infty} P(\Delta N_t^{(m)} = j | \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k) P(y_t^{(m)} = k | \mathbb{X}, \Theta_{r-1})}$$

By considering the term $\mathbb{E}\left[\left(z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)})\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right]$, we can use the algebra technique to rearrange it as follows:

$$\mathbb{E}\left[\left(z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)})\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] = \mathbb{E}\left[\left((z_t^{(m)} - Z_t^{(m)}) + (Z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)}))\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right]$$
$$= \mathbb{E}\left[\left((z_t^{(m)} - Z_t^{(m)})\right)^2 \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right]$$
$$+ 2(Z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)}))\mathbb{E}\left[(z_t^{(m)} - Z_t^{(m)}) \middle| \mathbb{X}, \Theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right]$$
$$+ \left((Z_t^{(m)} - (\hat{\kappa}_k\hat{\alpha}_k - \hat{\kappa}_k x_t^{(m)}))\right)^2$$

$$= \mathbb{V}ar\left[z_t^{(m)} \middle| \mathbb{X}, \theta, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] + \left(Z_t^{(m)} - \left(\hat{k}_k \hat{\hat{\alpha}}_k - \hat{k}_k x_t^{(m)}\right)\right)^2$$

where, $Z_t^{(m)} = \mathbb{E}\left[z_t^{(m)} \middle| \mathbb{X}, \theta_{r-1}, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right]$ and $\mathbb{E}\left[\left(z_t^{(m)} - Z_t^{(m)}\right) \middle| \mathbb{X}, \theta_{r-1}, y_t^{(m)} = k, \Delta N_t^{(m)} = j\right] = 0$. Therefore,

$$\hat{\sigma}_k^2 = \frac{\sum_{m=1}^{M} \sum_{t=1}^{T} \sum_{j=0}^{\infty} \left(\mathbb{V}ar\left[z_t^{(m)} \middle| \mathbb{X}, \theta, \Delta N_t^{(m)} = j, y_t^{(m)} = k\right] + \left(Z_t^{(m)} - \left(\hat{k}_k \hat{\alpha}_k - \hat{k}_k x_t^{(m)}\right)\right)^2 \right) P_{(j,t,k)}^{(m)}}{\sum_{m=1}^{M} \sum_{t=1}^{T} P\left(y_t^{(m)} = k \middle| \mathbb{X}, \theta\right)}$$

where, $P_{(j,t,k)}^{(m)} = P\left(\Delta N_t^{(m)} = j \middle| \mathbb{X}, \theta_{r-1}, y_t^{(m)} = k\right) P\left(y_t^{(m)} = k \middle| \mathbb{X}, \theta_{r-1}\right)$ and $\sum_{j=0}^{\infty} P\left(\Delta N_t^{(m)} = j \middle| \mathbb{X}, \theta_{r-1}, y_t^{(m)} = k\right) = 1$

Derive the Transition Probability ($\hat{q}$)

The parameter $\hat{q}_{i,l}$ can be derived by using the lagrange multiplier method with the constraint $\sum_{l=1}^{2} q_{i,l} = 1$ for the expected log complete likelihood $\mathbb{E}(ln(f(C_1, C_2, \dots, C_M)|\Theta_r)\ |\mathbb{X}, \Theta_{r-1})$. We will consider the term $\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T-1} ln(P(y_{t+1}^{(m)}|y_t^{(m)}))\ |\mathbb{X}, \Theta_{r-1}]$ which is the only term where the parameter $\hat{q}_{i,l}$is concealed in.

$$\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T-1} ln\left(P\left(y_{t+1}^{(m)}|y_t^{(m)}\right)\right) |\mathbb{X}, \Theta_{r-1}] = \sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{i=1}^{2}\sum_{l=1}^{2} \mathbb{E}[\mathbb{I}\left\{y_{t+1}^{(m)} = l, y_t^{(m)} = i\right\} lnq_{i,l}|\mathbb{X}, \Theta_{r-1}]$$

$$= \sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{i=1}^{2}\sum_{l=1}^{2} P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})\ lnq_{i,l}$$

$$\frac{\partial}{\partial q_{i,l}}\left[\mathbb{E}[\sum_{m=1}^{M}\sum_{t=1}^{T-1} ln\left(P\left(y_{t+1}^{(m)}|y_t^{(m)}\right)\right) |\mathbb{X}, \Theta_{r-1}] - \left(\sum_{l=1}^{2} q_{i,l} - 1\right)\mu\right] = 0$$

where, μ is a lagrange multiplier variable.

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1}\frac{P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})}{q_{i,l}} - \mu = 0$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1}\frac{P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})}{\mu} = q_{i,l}$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{l=1}^{2}\frac{P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})}{\mu} = \sum_{l=1}^{2} q_{i,l} = 1$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{l=1}^{2} P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1}) = \mu$$

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1} P\left(y_t^{(m)} = i\middle|\mathbb{X}, \Theta_{r-1}\right) = \mu$$

where,

$$\sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{l=1}^{2} P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1}) = \sum_{m=1}^{M}\sum_{t=1}^{T-1}\sum_{l=1}^{2} P(y_{t+1}^{(m)} = l|\mathbb{X}, \Theta_{r-1}, y_t^{(m)} = i)P(y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})$$

$$= \sum_{m=1}^{M}\sum_{t=1}^{T-1} P(y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})$$

Therefore,

$$\hat{q}_{i,l} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T-1} P(y_{t+1}^{(m)} = l, y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})}{\sum_{m=1}^{M}\sum_{t=1}^{T-1} P(y_t^{(m)} = i|\mathbb{X}, \Theta_{r-1})} \quad for\ i, l = 1,2$$

Appendix 4: Leveraged returns of in-sample trading

This appendix shows the results from Section 5.2 using leveraged return. The returns which are shown in this appendix is the percentage of leveraged realized profit/loss using the fixed margin (B) at 100,000 US dollar per contract and the leverage ratio (L) at 2,000 times for the simulation trade.

Table 24: In-sample result for each good news scenario using Stop Long order strategy trading on simulation

Stop Long Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	-0.001	0.005	-0.139	1.000	0.010	1.001	97.330	0.000
2	0.978	-0.001	0.005	-0.156	1.000	0.010	1.001	96.440	0.000
3	0.918	-0.001	0.005	-0.176	1.000	0.010	1.001	96.560	0.000
4	0.881	-0.002	0.009	-0.207	1.000	0.023	1.000	96.780	0.000
5	0.860	-0.002	0.007	-0.255	1.000	0.024	1.000	97.000	0.000
6	0.842	-0.002	0.007	-0.270	1.000	0.023	1.000	96.330	0.000
7	0.842	-0.002	0.007	-0.271	1.000	0.023	1.000	96.560	0.000
8	0.840	-0.002	0.007	-0.277	1.000	0.022	1.000	96.560	0.000
9	0.838	-0.002	0.007	-0.278	1.000	0.022	1.000	96.560	0.000
10	0.814	-0.001	0.003	-0.292	1.000	0.013	1.000	97.220	0.000
11	0.795	-0.002	0.006	-0.308	1.000	0.022	1.000	96.560	0.000
12	0.783	-0.002	0.005	-0.338	1.000	0.023	1.000	96.670	0.000
13	0.757	-0.001	0.001	-1.387	1.001	0.010	1.001	6.110	0.000
14	0.727	-0.002	0.008	-0.230	1.000	0.022	1.000	99.330	0.000
15	0.713	-0.001	0.004	-0.214	1.000	0.010	1.001	96.780	0.000
16	0.691	-0.001	0.004	-0.211	1.000	0.010	1.001	99.330	0.000
17	0.673	-0.001	0.004	-0.202	1.000	0.010	1.001	96.890	0.000
18	0.672	-0.001	0.004	-0.204	1.000	0.010	1.001	96.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 25: In-sample result for each good news scenario using Stop Short order strategy trading on simulation

Stop Short Strategy

Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	0.000	0.006	-0.017	1.000	0.010	0.999	99.110	0.000
2	0.978	0.000	0.005	-0.008	1.000	0.010	0.999	94.440	0.000
3	0.918	0.000	0.005	-0.008	1.000	0.010	0.999	98.110	0.000
4	0.881	0.034	11.092	0.003	1.000	24.330	0.999	94.560	0.000
5	0.860	0.172	15.298	0.011	1.000	32.997	0.999	96.670	0.000
6	0.842	0.672	24.253	0.028	1.000	57.316	0.999	91.110	0.240
7	0.842	0.635	27.314	0.023	1.000	58.708	0.999	98.780	0.340
8	0.840	0.475	15.506	0.031	1.000	33.034	0.999	99.670	0.000
9	0.838	0.403	16.173	0.025	1.000	34.328	0.999	99.890	0.000
10	0.814	0.285	14.929	0.019	1.000	32.811	0.999	98.330	0.000
11	0.795	0.282	15.238	0.018	1.000	32.855	0.999	99.890	0.000
12	0.783	0.159	14.771	0.011	1.000	32.300	0.999	99.890	0.000
13	0.757	0.021	13.919	0.002	1.000	32.100	0.999	97.560	0.000
14	0.727	0.001	0.361	0.002	1.000	0.856	0.999	97.890	0.000
15	0.713	0.000	0.004	-0.008	1.000	0.010	0.999	98.440	0.000
16	0.691	0.000	0.644	0.001	1.000	1.566	0.999	98.440	0.000
17	0.673	0.000	0.043	-0.003	1.000	0.110	0.999	98.440	0.000
18	0.672	0.000	0.038	-0.004	1.000	0.096	0.999	99.670	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 26: In-sample result for each good news scenario using Limit Long order strategy trading on simulation

Limit Long Strategy

Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	7.922	10.128	0.782	0.999	67.201	1.001	8.000	0.000
2	0.978	16.717	11.426	1.463	0.999	79.471	1.001	5.330	0.000
3	0.918	7.352	9.632	0.763	0.999	82.822	1.001	4.560	2.440
4	0.881	0.008	2.372	0.003	0.999	26.810	1.001	4.780	0.000
5	0.860	6.697	10.065	0.665	0.999	86.833	1.001	4.670	2.380
6	0.842	9.550	12.021	0.795	0.999	92.839	1.001	5.110	8.700
7	0.842	8.418	12.024	0.700	0.999	92.243	1.001	5.220	6.380
8	0.840	8.798	12.057	0.730	0.999	92.911	1.001	5.000	13.330
9	0.838	4.306	13.861	0.311	0.999	110.669	1.001	4.780	48.840
10	0.814	4.239	9.180	0.462	0.999	92.951	1.001	4.890	6.820
11	0.795	4.955	11.447	0.433	0.999	101.973	1.001	5.110	26.090
12	0.783	18.297	10.078	1.816	0.999	91.042	1.000	4.670	2.380
13	0.757	4.677	6.873	0.680	0.999	89.850	1.001	4.330	2.560
14	0.727	0.685	4.409	0.155	0.999	88.231	1.001	3.330	0.000
15	0.713	5.422	4.835	1.121	0.999	97.669	1.001	3.110	0.000
16	0.691	15.424	6.824	2.260	0.999	104.142	1.001	2.890	11.540
17	0.673	19.254	6.371	3.022	0.999	104.331	1.001	2.560	4.350
18	0.672	19.177	6.360	3.015	0.999	104.273	1.001	2.560	4.350

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 27: In-sample result for each good news scenario using Limit Short order strategy trading on simulation

Limit Short Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.112	2.558	8.096	0.316	1.001	67.031	0.999	6.670	0.000
2	0.978	2.711	7.174	0.378	1.001	73.814	0.999	5.670	0.000
3	0.918	7.579	8.886	0.853	1.001	83.042	0.999	4.670	2.380
4	0.881	5.006	7.465	0.671	1.001	89.001	0.999	3.890	5.710
5	0.860	5.075	6.279	0.808	1.001	83.736	0.999	3.780	2.940
6	0.842	13.863	8.497	1.632	1.001	86.872	0.999	3.670	3.030
7	0.842	15.770	8.377	1.883	1.001	87.090	0.999	3.560	0.000
8	0.840	14.425	8.499	1.697	1.001	87.108	0.999	3.440	3.230
9	0.838	13.282	8.554	1.553	1.001	87.225	0.999	3.560	0.000
10	0.814	10.759	6.934	1.552	1.001	88.256	0.999	3.110	0.000
11	0.795	6.162	6.894	0.894	1.001	88.571	0.999	3.330	0.000
12	0.783	-1.578	7.931	-0.199	1.000	30.000	1.000	67.890	0.000
13	0.757	14.239	7.103	2.005	1.001	87.926	0.999	2.890	0.000
14	0.727	15.424	7.274	2.120	1.001	86.707	0.999	3.000	0.000
15	0.713	15.541	7.771	2.000	1.001	92.335	0.999	2.890	0.000
16	0.691	9.239	6.835	1.352	1.001	105.499	0.999	2.330	19.050
17	0.673	26.483	8.317	3.184	1.001	98.354	1.000	2.330	0.000
18	0.672	25.305	7.979	3.171	1.001	98.217	0.999	2.330	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 28: In sample result for each bad news scenario using Stop Long order strategy trading on simulation

Stop Long Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	0.619	33.093	0.019	1.000	62.686	1.001	99.890	2.110
2	0.958	0.667	28.538	0.023	1.000	56.438	1.001	98.330	4.750
3	0.948	1.154	24.510	0.047	0.999	47.237	1.001	100.000	0.000
4	0.909	0.303	31.127	0.010	1.000	62.580	1.001	99.780	1.890
5	0.883	1.025	21.150	0.048	0.999	43.320	1.001	100.000	0.000
6	0.862	0.665	18.693	0.036	1.000	40.069	1.001	98.670	0.000
7	0.855	0.454	19.399	0.023	1.000	41.477	1.001	98.780	0.000
8	0.839	0.589	16.729	0.035	1.000	36.730	1.001	98.890	0.000
9	0.838	0.527	16.714	0.032	1.000	36.736	1.001	98.780	0.000
10	0.811	0.000	0.004	-0.013	1.000	0.010	1.001	96.330	0.000
11	0.807	0.642	21.850	0.029	0.999	47.340	1.001	100.000	0.000
12	0.806	0.739	13.702	0.054	0.999	30.932	1.001	100.000	0.000
13	0.795	1.407	25.268	0.056	0.999	57.470	1.001	100.000	0.560
14	0.792	0.000	0.004	-0.003	1.000	0.010	1.001	96.330	0.000
15	0.777	0.080	28.915	0.003	1.000	64.671	1.001	99.890	1.330
16	0.776	0.161	13.710	0.012	1.000	31.762	1.001	98.780	0.000
17	0.775	0.117	27.529	0.004	1.000	61.746	1.001	99.890	0.670
18	0.756	0.366	27.672	0.013	1.000	62.874	1.001	100.000	0.560
19	0.739	0.181	15.763	0.012	1.000	37.578	1.001	98.560	0.000
20	0.720	0.649	15.733	0.041	1.000	37.666	1.001	100.000	0.000
21	0.689	0.139	14.020	0.010	1.000	34.804	1.001	99.220	0.000
22	0.684	0.755	21.598	0.035	0.999	51.924	1.001	100.000	0.000
23	0.670	0.716	17.673	0.041	1.000	44.188	1.001	99.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 29: In sample result for each bad news scenario using Stop Short order strategy trading on simulation

Stop Short Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	-0.001	0.005	-0.231	1.000	0.010	0.999	96.780	0.000
2	0.958	-0.001	0.005	-0.241	1.000	0.010	0.999	97.670	0.000
3	0.948	-0.001	0.007	-0.184	1.000	0.020	1.000	97.440	0.000
4	0.909	-0.001	0.005	-0.231	1.000	0.010	0.999	100.000	0.000
5	0.883	-0.002	0.004	-0.478	1.000	0.021	1.000	78.560	0.000
6	0.862	-0.001	0.005	-0.266	1.000	0.010	0.999	97.670	0.000
7	0.855	-0.001	0.005	-0.266	1.000	0.010	0.999	97.330	0.000
8	0.839	-0.001	0.005	-0.262	1.000	0.010	0.999	97.670	0.000
9	0.838	-0.001	0.000	-3.853	0.999	0.010	0.999	5.110	0.000
10	0.811	-0.001	0.004	-0.262	1.000	0.010	0.999	97.780	0.000
11	0.807	-0.002	0.007	-0.258	1.000	0.020	1.000	92.220	0.000
12	0.806	-0.002	0.007	-0.274	1.000	0.018	1.000	91.780	0.000
13	0.795	-0.002	0.002	-0.915	1.000	0.010	0.999	23.780	0.000
14	0.792	-0.001	0.004	-0.318	1.000	0.010	0.999	93.330	0.000
15	0.777	-0.001	0.004	-0.249	1.000	0.010	0.999	99.000	0.000
16	0.776	-0.002	0.004	-0.452	1.000	0.010	0.999	70.780	0.000
17	0.775	-0.001	0.004	-0.259	1.000	0.010	0.999	98.780	0.000
18	0.756	-0.001	0.006	-0.232	1.000	0.018	1.000	95.000	0.000
19	0.739	-0.001	0.004	-0.244	1.000	0.010	0.999	99.440	0.000
20	0.720	-0.001	0.001	-2.753	0.999	0.010	0.999	3.560	0.000
21	0.689	-0.002	0.009	-0.252	1.000	0.036	1.000	99.220	0.000
22	0.684	-0.001	0.004	-0.259	1.000	0.010	0.999	99.000	0.000
23	0.670	-0.001	0.004	-0.281	1.000	0.010	0.999	97.890	0.000

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 30: In sample result for each bad news scenario using Limit Long order strategy trading on simulation

Limit Long Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	18.689	14.420	1.296	0.999	96.704	1.001	6.330	19.300
2	0.958	19.481	13.046	1.493	0.999	93.868	1.001	5.890	11.320
3	0.948	20.970	13.136	1.596	0.999	94.628	1.001	5.780	9.620
4	0.909	11.132	7.663	1.453	0.999	89.336	1.001	5.000	0.000
5	0.883	13.524	8.088	1.672	0.999	93.504	1.001	4.890	2.270
6	0.862	19.262	8.838	2.179	0.999	93.966	1.001	4.000	2.780
7	0.855	12.653	9.602	1.318	0.999	97.817	1.001	4.110	18.920
8	0.839	15.274	12.166	1.256	0.999	107.242	1.001	4.000	27.780
9	0.838	15.294	12.167	1.257	0.999	107.258	1.001	4.000	27.780
10	0.811	15.931	9.192	1.733	0.999	90.478	1.001	3.890	2.860
11	0.807	19.006	8.981	2.116	0.999	89.655	1.001	3.670	0.000
12	0.806	17.324	10.593	1.636	0.999	101.529	1.001	3.670	21.210
13	0.795	22.382	10.564	2.119	0.999	104.848	1.001	3.440	12.900
14	0.792	22.916	10.545	2.173	0.999	105.068	1.001	3.440	12.900
15	0.777	21.697	10.872	1.996	0.999	105.703	1.001	3.670	15.150
16	0.776	21.712	10.874	1.997	0.999	105.729	1.001	3.670	15.150
17	0.775	22.448	10.792	2.080	0.999	105.715	1.001	3.670	12.120
18	0.756	21.207	12.111	1.751	0.999	117.161	1.001	3.220	41.380
19	0.739	33.282	11.493	2.896	0.999	106.484	1.001	2.890	3.850
20	0.720	26.909	9.274	2.902	0.999	110.301	1.001	2.330	4.760
21	0.689	42.319	8.991	4.707	0.999	110.323	1.001	1.890	0.000
22	0.684	34.397	8.503	4.045	0.999	111.165	1.001	1.890	0.000
23	0.670	47.387	8.008	5.917	0.999	120.000	1.000	1.220	9.090

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Table 31: In sample result for each bad news scenario using Limit Short order strategy trading on simulation

Limit Short Strategy									
Scenario	*Scale Ratio*	*Return (%)*	*Std. (%)*	*Sharpe Ratio*	*Open*	*Lot*	*Target*	*%Matched*	*%Forced*
1	1.021	16.305	13.231	1.232	1.001	80.502	0.999	8.000	1.390
2	0.958	18.109	15.257	1.187	1.001	86.571	0.999	7.780	7.140
3	0.948	20.980	16.743	1.253	1.001	90.181	0.999	7.780	11.430
4	0.909	21.744	15.351	1.416	1.001	86.689	0.999	7.000	3.170
5	0.883	21.793	14.275	1.527	1.001	85.093	0.999	7.110	0.000
6	0.862	22.680	15.872	1.429	1.001	91.977	0.999	6.670	10.000
7	0.855	24.130	15.590	1.548	1.001	92.158	0.999	6.220	8.930
8	0.839	11.310	15.459	0.732	1.001	116.010	0.999	5.560	48.000
9	0.838	11.268	15.443	0.730	1.001	115.960	0.999	5.560	48.000
10	0.811	15.107	14.233	1.061	1.001	111.033	0.999	4.780	30.230
11	0.807	11.336	15.739	0.720	1.001	120.000	0.999	4.780	51.160
12	0.806	24.932	12.899	1.933	1.001	90.098	0.999	4.670	2.380
13	0.795	15.065	12.592	1.196	1.001	102.709	0.999	4.780	16.280
14	0.792	15.333	12.069	1.270	1.001	100.312	0.999	4.780	9.300
15	0.777	20.666	11.451	1.805	1.001	100.651	0.999	4.110	8.110
16	0.776	24.294	11.786	2.061	1.001	100.619	0.999	4.110	5.410
17	0.775	23.588	11.863	1.988	1.001	100.725	0.999	4.220	5.260
18	0.756	21.525	12.886	1.670	1.001	112.883	0.999	4.000	27.780
19	0.739	29.290	13.405	2.185	1.001	110.127	0.999	3.560	21.880
20	0.720	22.664	10.142	2.235	1.001	97.217	0.999	3.560	3.130
21	0.689	31.702	11.796	2.687	1.001	111.036	0.999	3.440	16.130
22	0.684	29.083	11.702	2.485	1.001	111.723	0.999	3.670	12.120
23	0.670	32.072	11.602	2.764	1.001	111.224	0.999	3.440	9.680

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the average of leveraged return, which is calculated by averaging the return per trade from simulation paths in each scenario using only the matched order, in percent. *Std. (%)* shows the standard deviation of return only in the case that order is matched using the optimal solution set in percent. *Sharpe Ratio* shows the Sharpe ratio for each trading scenario. *Open* shows the optimal Open price from optimization. *Lot* shows the optimal lot sizes from optimization. *Target* shows the optimal Closing price from optimization. *%Matched* shows the probability that the sent order will be matched from 5,000 trading simulated paths. *%Forced* shows the probability that the matched order will be forced closed.

Appendix 5: Leveraged returns of out-of-sample trading

This appendix shows the results from Section 5.3 using leveraged return. The returns which are shown in this appendix is the percentage of leveraged realized profit/loss using the fixed margin (B) at 100,000 US dollar per contract and the leverage ratio (L) at 2,000 times for the out-of-sample trade.

Table 32: Detailed out-of-sample trading for good news using 2-PDJ model with Stop Short order and three benchmark strategies (2,000 times leverage ratio is applied)

Good News	**2-Regimes Pure Diffusion with Jump (2-PDJ)**					
		Return (%)				
	Scale Ratio	*Stop Short order Strategy*	*Optimal Open*	*Optimal Target*	*Opened Price*	*Optimal Lot Sizes*
1	1.1116	N/A	N/A	N/A	N/A	N/A
2	0.9775	N/A	N/A	N/A	N/A	N/A
3	0.9182	N/A	N/A	N/A	N/A	N/A
4	0.8809	-10.253	0.99998	0.99900	0.99998	24.33028
5	0.8602	-7.466	0.99999	0.99900	0.99998	32.99715
6	0.8424	-8.225	0.99995	0.99900	0.99998	57.31578
7	0.8419	-8.891	1.00002	0.99900	0.99998	58.70820
8	0.8398	3.205	1.00004	0.99900	0.99999	33.03402
9	0.8376	10.744	1.00005	0.99900	0.99985	34.32839
10	0.8142	-6.089	1.00001	0.99900	0.99998	32.81075
11	0.7946	-0.726	1.00005	0.99900	1.00000	32.85535
12	0.7830	-19.053	1.00005	0.99900	0.99985	32.30013
13	0.7569	1.201	1.00000	0.99900	0.99991	32.10029
14	0.7266	0.108	1.00000	0.99906	0.99998	0.85629
15	0.7130	N/A	N/A	N/A	N/A	N/A
16	0.6910	0.115	1.00001	0.99906	0.99999	1.56578
17	0.6734	N/A	N/A	N/A	N/A	N/A
18	0.6719	N/A	N/A	N/A	N/A	N/A

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the leveraged return from using the Stop Short order strategy in out-of-sample trade. *Optimal Open* shows the optimal Open price. *Optimal Target* shows the optimal Closing price. *Opened Price* shows the actual normalized bid-ask quote price when news is announced. This depends on strategy: we use bid price for short and ask price for long positions. *Optimal Lot Sizes* shows the optimal lot sizes. “N/A” in *Return (%)* column means we do not submit an order because the chosen strategy suggests a negative mean return from the trading simulation.

Table 33: Detailed out-of-sample trading for bad news using 2-MRJ model with Stop Long order and three benchmark strategies (2,000 times leverage ratio is applied)

Bad News	**2-Regimes Mean-Reverting with Jump (2-MRJ)**					
		Return (%)				
	Scale Ratio	*Stop Long order Strategy*	*Optimal Open*	*Optimal Target*	*Opened Price*	*Optimal Lot Sizes*
1	1.0213	6.896	0.99971	1.00100	1.00002	62.68645
2	0.9577	19.189	1.00000	1.00100	1.00001	56.43788
3	0.9484	-15.588	0.99936	1.00100	1.00002	47.23739
4	0.9090	0.000	0.99994	1.00100	1.00002	62.57991
5	0.8834	5.198	0.99900	1.00100	1.00002	43.32004
6	0.8615	-10.017	1.00000	1.00100	1.00002	40.06876
7	0.8546	-10.784	1.00000	1.00100	1.00002	41.47650
8	0.8388	0.735	0.99999	1.00100	1.00001	36.72956
9	0.8380	0.735	1.00000	1.00100	1.00002	36.73550
10	0.8111	N/A	N/A	N/A	N/A	N/A
11	0.8066	3.787	0.99924	1.00100	1.00002	47.34020
12	0.8058	-7.424	0.99900	1.00100	1.00003	30.93170
13	0.7947	-1.724	0.99900	1.00067	1.00001	57.46999
14	0.7922	N/A	N/A	N/A	N/A	N/A
15	0.7770	7.760	0.99983	1.00100	1.00015	64.67080
16	0.7764	-10.799	1.00000	1.00100	1.00000	31.76168
17	0.7752	-4.322	0.99990	1.00100	1.00000	61.74571
18	0.7564	-2.515	0.99969	1.00100	0.99999	62.87445
19	0.7389	-3.382	1.00000	1.00100	1.00001	37.57832
20	0.7195	0.000	0.99966	1.00100	1.00001	37.66578
21	0.6885	-10.441	0.99999	1.00100	0.99999	34.80360
22	0.6838	-5.192	0.99901	1.00077	0.99999	51.92372
23	0.6695	-2.651	0.99979	1.00100	0.99997	44.18797

Scale Ratio shows the volatility ratio calculated by a GARCH model. *Return (%)* shows the leveraged return from using the Stop Short order strategy in out-of-sample trade. *Optimal Open* shows the optimal Open price. *Optimal Target* shows the optimal Closing price. *Opened Price* shows the actual normalized bid-ask quote price when news is announced. This depends on strategy: we use bid price for short and ask price for long positions. *Optimal Lot Sizes* shows the optimal lot sizes. “N/A” in *Return (%)* column means we do not submit an order because the chosen strategy suggests a negative mean return from the trading simulation

.

Table 34: Out-of-sample trading summary result using Stop Short order with 2-PDJ model for good news announced and three benchmark strategies (2,000 times leverage ratio is applied)

Good News	Out-Of-Sample
Strategy : Stop Short order	Stop Short
Total News	18
Number of Trading Scenarios	12
Leveraged Average return (%)	-3.778
STD Leveraged return (%)	7.439
Sharpe Ratio	-0.508
Percent Win given trade (%)	41.67
Force Close Occurrence given trade (%)	0.00%
Number of Out-of-sample loss Trade	7
Forecasting Performance (%)	67%

Table 35: Out-of-sample trading summary result using Stop Long order with 2-MRJ model for bad news announced and three benchmark strategies (2,000 times leverage ratio is applied)

Bad News	Out-Of-Sample
Strategy : Stop Long order	Stop Long
Total News	23
Number of Trading Scenarios	21
Currency Average return (%)	-1.930
STD Currency return (%)	7.727
Sharpe Ratio	-0.250
Percent Win given trade (%)	42.85%
Force Close Occurrence given trade (%)	0.00%
Number of Out-of-sample loss Trade	12
Forecasting Performance (%)	50%

VITA

Pavarich Suwanpetai was born in Bangkok, Thailand, on January 25, 1990. He received Bachelor of Engineering major in the Information and Communication from Chulalongkorn University in July 2012. After that, he continued his plan to improve his knowledge in finance, he decided to study at Chulalongkorn University in August 2013 in Master of Science (M.Sc.) major in Financial Engineering Program. If there is any question about this thesis, please contract him via E-mail (pasee4109@gmail.com).