

# FX Spot Trading and Risk Management from A Market Maker's Perspective

by  
Mu Yang

A thesis  
presented to the University of Waterloo  
in fulfilment of the  
thesis requirement for the degree of  
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2011

© Mu Yang 2011

## **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

## Abstract

Due to the rapid development of computing technology and faster growth of financial industry, Foreign Exchange high-frequency trading has become substantially more prominent to today's market players, especially to bankers and market makers. This research aims at introducing today's FX high-frequency trading structure and discussing how a market maker can effectively reduce downside risk when market faces a huge upward or downward stress. An Exponential Moving Average operator is introduced and implemented using a Matlab software for tick-by-tick data analysis. Simulation framework for market high-frequency data and client trading flow is also introduced and implemented using the Matlab software. Real-time P&L calculation is introduced and used to determine the performance of a proposed risk hedging strategy. On the other hand, due to the financial crisis we experienced in 2007, 2008, and 2009, we analyze the tail risk of foreign exchange market. Extreme Value Theory (EVT) has been applied to real EUR/USD data, which contains eight-year daily closing exchange rate. An extension of from EVT to Value-at-Risk (VaR) calculation is introduced. We also consider the volatility clustering issue in asset returns and demonstrate how GARCH model can be applied for VaR calculation. Lastly, we propose a method of using VaR as a high-frequency risk measure for risk hedging strategies during intra-day trading.

## **Acknowledgements**

I am grateful to Professor Tony Wirjanto from the School of Accounting and Finance at the University of Waterloo, who supervised and encouraged me in finishing this thesis by providing a lot of insightful ideas and guidance. I would also like to thank my readers Professor Don McLeish from the Department of Statistics and Actuarial Science and Professor Allen Huang from the School of Accounting and Finance, who helped me polishing this document through their important and helpful suggestions. All your guidance was key to finish this thesis. Finally, I would also like to express my gratitude to my family and friends for their supports and inspirations.

# Contents

<b>Author's Declaration</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>x</b>
<b>1 Introduction and Selected Literature Review</b>	<b>1</b>
1.1 Motivation and Introduction . . . . .	1
1.2 Foreign Exchange Market . . . . .	2
1.3 FX Spot Exchange Rate . . . . .	5
1.4 FX Market Participants . . . . .	6
1.5 Interbank Market and Market Making . . . . .	8
1.6 Electronic Broking System . . . . .	10
1.7 FX High-Frequency Trading . . . . .	13
<b>2 Exponential Moving Average</b>	<b>16</b>
2.1 Market Data . . . . .	16
2.2 Exponential Moving Average . . . . .	20
2.3 Application of the EMA Operator . . . . .	24

<b>3</b>	<b>Simulation Framework</b>	<b>28</b>
3.1	Motivation . . . . .	28
3.2	Poisson Process . . . . .	29
3.3	Geometric Brownian Motion . . . . .	30
3.4	Simulation of Market Data . . . . .	31
3.5	Simulation of Client Trades . . . . .	33
<b>4</b>	<b>Risk Hedging</b>	<b>39</b>
4.1	Hedging Strategy . . . . .	39
4.2	Implementation of the Hedging Strategy . . . . .	41
4.3	Scenario Analysis . . . . .	42
4.3.1	Balanced Client Buying and Selling Flows under Flat Market Condition . . . . .	44
4.3.2	Intensive Client Selling Flow under Downward Market Condition.	47
4.3.3	Intensive Client Buying Flow under Upward Market Condition.	48
<b>5</b>	<b>Tail Risk Analysis</b>	<b>54</b>
5.1	Overview of Extreme Value Theory . . . . .	54
5.2	Maximum Likelihood Methods for EVT . . . . .	58
5.3	Empirical Analysis on EVT . . . . .	60
5.4	Value-at-Risk (VaR) . . . . .	64
5.5	An Extreme Value Approach to VaR . . . . .	65
5.6	Considering Volatility Clustering . . . . .	68
5.6.1	GARCH $(p, q)$ Model . . . . .	69
5.6.2	Conditional Quantile . . . . .	69
5.6.3	Empirical Analysis on GARCH(1,1) . . . . .	70
<b>6</b>	<b>VaR for A Trading Strategy</b>	<b>76</b>
6.1	Ideas . . . . .	76
6.2	Methodology and Example . . . . .	77

<b>7 Conclusions and Future Work</b>	<b>84</b>
<b>Bibliography</b>	<b>86</b>
<b>Appendix</b>	<b>89</b>
<b>A Proof of EMA Iteration Formula for <math>t</math> Starting from <math>-\infty</math></b>	<b>89</b>
<b>B Proof of EMA Iteration Formula for <math>t</math> Starting at Zero</b>	<b>90</b>

# List of Figures

1.1	FX Daily Trading Volume Distribution . . . . .	4
1.2	Reuters Dealing and EBS Trading Screens . . . . .	12
2.1	Number of Quote Updates for Different Inter-arrival Time on 2010/05/31 . . . . .	17
2.2	Average Number of Quote Updates per Minute for Each 30-Minute Period on 2010/05/31 . . . . .	17
2.3	USD/CAD Spot Rate from 00:00 a.m. to 23:59 p.m. on 2010/05/31 . . . . .	19
2.4	USD/CAD Spot Rate from 10:00 a.m. to 12:00 p.m. on 2010/05/31 . . . . .	19
2.5	USD/CAD Spot Rate from 11:00 a.m. to 11:10 a.m. on 2010/05/31 . . . . .	20
2.6	EMA Weight Functions with Different Range Values . . . . .	22
2.7	Time Series of Mid Price and Its EMA Values with Different Values of Range $\lambda$ .	25
2.8	Mean Squared Errors of EMA estimates with Different Values of Range $\lambda$ . . . . .	26
3.1	Histogram of Simulation for USD/CAD Market Data Inter-Arrival Time . . . . .	33
3.2	Sample Paths of USD/CAD Bid and Ask Prices by an Ito Process . . . . .	34
3.3	Sample Paths of Market Maker's Base and Counter Currency Wealth Processes .	37
3.4	Sample Path of Market Maker's P&L . . . . .	38
4.1	Sample Paths of Market Maker's $AW_1(T_{C_k})$ and $AW_2(T_{C_k})$ . . . . .	42
4.2	Sample Path of Market Maker's $P\&L(t)$ and $PL_A(t)$ . . . . .	43
4.3	P&L Without Risk Hedging in Flat Market with Balanced Client Flow . . . . .	45
4.4	P&L With Risk Hedging in Flat Market with Balanced Client Flow . . . . .	45
4.5	P&L Differences in Flat Market with Balanced Client Flow . . . . .	46

4.6	Sharp Ratio Comparison in Flat Market with Balanced Client Flow . . . . .	46
4.7	P&L Without Risk Hedging in Downward Trend Market with Intensive Client Sell	49
4.8	P&L With Risk Hedging in Downward Trend Market with Intensive Client Sell .	49
4.9	P&L Differences in Downward Trend Market with Intensive Client Sell . . . . .	50
4.10	Sharp Ratio Comparison in Downward Trend Market with Intensive Client Sell .	50
4.11	P&L Without Risk Hedging in Upward Trend Market with Intensive Client Buy .	52
4.12	P&L With Risk Hedging in Upward Trend Market with Intensive Client Buy . .	52
4.13	P&L Differences in Upward Trend Market with Intensive Client Buy . . . . .	53
4.14	Sharp Ratio Comparison in Upward Trend Market with Intensive Client Buy . .	53
5.1	Density Functions for Frechet, Weibull, and Gumbel when $\alpha = 1.5$ . . . . .	57
5.2	EUR/USD Day-End Price . . . . .	61
5.3	Q-Q Plot of EUR/USD Daily Loss . . . . .	61
5.4	EUR/USD Maximum Daily Loss for Each Month . . . . .	62
5.5	Empirical and Theoretical CDF Comparison for Maximum Daily Loss in Monthly Block . . . . .	63
5.6	Autocorrelation Function of EUR/USD Daily Loss . . . . .	71
5.7	Inferred Innovations, Standard Deviations, and Original Time Series . . . . .	72
5.8	Autocorrelation Function of Innovations . . . . .	73
6.1	Simulated P&L Sample Path for a 30-hour Trading Period . . . . .	79
6.2	P&L Increments for a 30-hour Trading Period . . . . .	80
6.3	Q-Q Plot for P&L Increments for A 30-hour Trading Period . . . . .	81
6.4	Maximum 2-Minute Loss in Hourly Block for A 30-Hours Period . . . . .	82
6.5	Empirical and Theoretical CDF Comparison for Maximum 2-Minute Loss in Hourly Block . . . . .	83

# List of Tables

1.1	ISO 4217 Code for G10 Currencies . . . . .	5
1.2	Top 10 FX Market Participants by % of Overall Market Volume . . .	9
5.1	MLE Estimates for Generalized Extreme Value Distribution with Dif- ferent Block Size . . . . .	63
5.2	AR(1)-GARCH(1,1)Parameter Estimates given by MLE with Normal Innovations . . . . .	72
5.3	AR(1)-GARCH(1,1)Parameter Estimates given by MLE with Student $t$ Innovations . . . . .	74
5.4	Parameter Estimates for GJR-GARCH(1,1,1) . . . . .	75

# Chapter 1

## Introduction and Selected Literature Review

### 1.1 Motivation and Introduction

The motivation of conducting this research is to provide an overview of today's FX market making business and its risk management.

The first chapter contains introductions to current structure of global FX market. We introduce the market by its segment and participants. We also introduce the FX interbank market, market making business model, and Electronic Broking System for FX trading. In Chapter 2, we introduce the exponential moving average method and its application in FX high-frequency data cleaning. In Chapter 3, we propose a simulation framework for both client trading and FX market rate simulation in Matlab. A Poisson process and Geometric Brownian Motion will respectively be applied in the simulation of an event arrival process and asset price movements. In Chapter 4, we implement a position limit risk hedging strategy based on simulated client trading process. Definitions of market maker's wealth process and P&L measures will be defined and calculated. In order to accomplish this objective, we need to assume that the market maker has the ability to access liquidity, which means

that the market maker can execute trades in the market at a given price at any time. Thanks to the super power of today's computer technology and networks, this assumption is in keeping with the reality for some of the most liquid currency pairs such as USD/CAD, USD/JPY and EUR/USD etc. In Chapter 5, we discuss about the tail risk in foreign exchange market by using Extreme Value Theory (EVT on EUR/USD daily returns. Maximum likelihood estimates of shape, scale, and location parameters are calculated and compared for different block sizes. An extension of calculating Value-At-Risk (VaR) by EVT is promoted. We also discuss about GARCH model in handling stochastic volatilities. In Chapter 6, we apply the EVT on the results produced by the simulation framework and hedging strategies introduced in Chapter 3 and Chapter 4 respectively. Thus, we propose a method of using VaR as a risk measure on a high-frequency level<sup>1</sup> for a risk hedging strategy.

In this research, we assume that we have available to us a powerful computer that is capable of obtaining all the liquidities at the given prices at any time.

## 1.2 Foreign Exchange Market

Foreign Exchange (Forex or FX) market is a twenty-four hour, decentralized over-the-counter (OTC) financial market for currency trading. Its primary purpose is to carry out international trade and investment activities, by allowing business to convert one currency to another one. For example, if a Canadian company wishes to import three million US dollars worth of raw materials from the US, then an exchange between Canadian and US dollars needs to take place, and FX market is the place where the company can carry out this currency exchange. Some other reasons to conduct foreign exchange are to invest in foreign financial assets, to hedge against unfavorable rates of exchange in the future and to profit from those changes. This activity is also known as speculation.

---

<sup>1</sup>VaR can be calculated as many times as needed during intra-day trading.

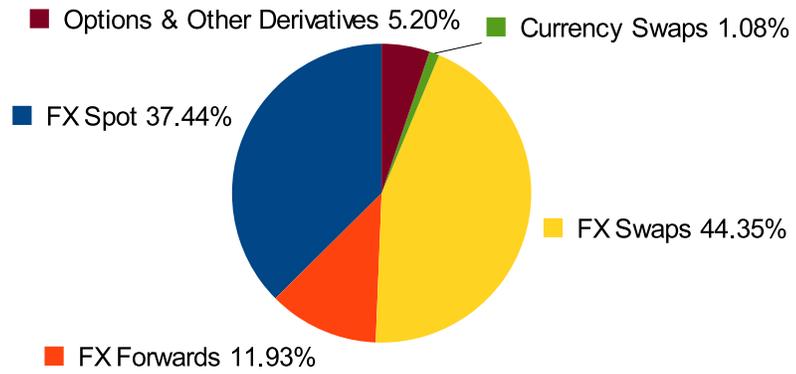
According to [16], FX market is the largest and most liquid financial market in the world in terms of daily trading volume. Based on the numbers published by Bank for International Settlements (BIS)<sup>2</sup>, the FX total trading volume increased by 38% between April 2005 and April 2006 and has increased more than two-folds since 2001. In 2010, the average daily turnover was reported to be \$3.98 trillion, of which \$1.49 trillion was traded in FX spot market, \$475 billion in the forward market, \$1.765 trillion in FX swaps, \$43 billion in currency swaps, and \$207 billion in options and other derivatives. Geographically, 36.7% of the total trading volume was made in London, while 17.9% in New York City, and 6.2% in Tokyo. The dramatic increase in trading volume is mainly due to the growing importance of FX as an asset class in fund management, particularly in hedge funds and pension funds. For the FX daily trading volume distribution by product types and geographic locations, see Figure 1.1.

The price quotation for currencies generally follows the ISO convention, and is the three-letter code used to identify a currency, such as USD for US dollar and GBP for British sterling. For the list of ISO codes of G10 currencies, see Table 1.1. Currencies are always traded against one another. The activity of buying one currency is typically accompanied by the activity of selling another currency. Thus, when a price is quoted, it is always quoted for a currency pair, and is labeled as XXXYYY or XXX/YYY. The first three letters (XXX) represent the *base currency* that is quoted against the second currency (YYY), which is called the *counter currency* or *quoted currency*. In practice, the rate convention is to quote everything in terms of one unit of the US dollar. For instance, the US dollar and Swiss franc rate is quoted as USD/CHF, and is the number of Swiss franc to one US dollar. The exceptions are for euro and sterling, which are quoted respectively as GBP/USD and EUR/USD, the number of US dollar to one pound and one euro.

---

<sup>2</sup>Source: 2010 Triennial Central Bank Survey, Bank for International Settlements

### FX Daily Trading Volume by Product Types



### FX Daily Trading Volume by Locations

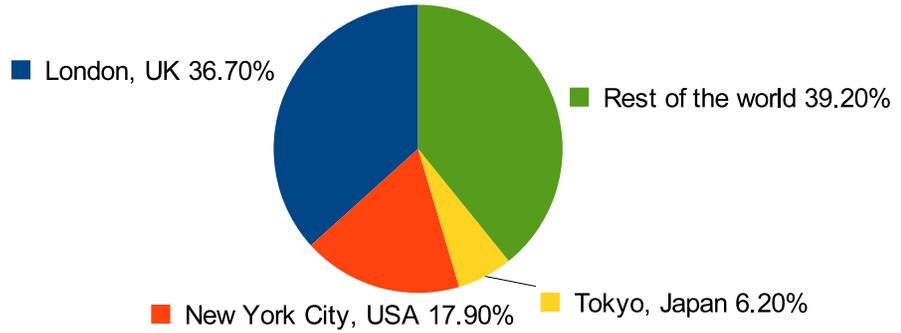


Figure 1.1: FX Daily Trading Volume Distribution

Table 1.1: ISO 4217 Code for G10 Currencies

---

USD	US Dollar
CAD	Canadian Dollar
JPY	Japanese Yen
AUD	Australian Dollar
NZD	New Zealand Dollar
GBP	British Pound
EUR	EURO
CHF	Swiss Franc
SEK	Swedish Krona
NOK	Norwegian Krona

---

### 1.3 FX Spot Exchange Rate

A spot FX trade is a purchase or sale of one currency against another one, with delivery in two business days after the trade date. Non-business days are not included in the count, so a trade on a Friday is settled on the following Tuesday. There are some exceptions to this. For example trading of a US dollar against a Canadian dollar are settled on the next working day. A settlement date that falls on a public holiday in the country of one of the two currencies is delayed for settlement by one day. An FX transaction is possible between any two currencies. However, to reduce the number of quotes that need to be made, the market generally quotes only against the US dollar or occasionally against the sterling or euro, so that the exchange rate between two non-dollar currencies is calculated from the rate for each currency against the dollar. The resulting exchange rate is known as the *cross-rate*. Cross-rates themselves are also traded between banks in addition to dollar-based rates. This is usually because the relationship between any two rates is closer than that of either one against the dollar, for example the Swiss franc moves more closely in line with the euro than it does against the dollar; so in practice one observes that the USD/CHF rate is more or less a function of the EUR/CHF rate.

The spot FX quote is a two-way bid-offer price. The *bid* indicates the rate at which a bank is prepared to buy the base currency against the counter currency; it is the lower rate. The other side of the quote is called the *offer*, which is the rate at which the bank is prepared to sell the base currency against the counter currency. For example, USD/CAD = 1.2324/26 tells that the bid and offer prices for 1 USD are at 1.2324 and 1.2326 CAD. In other words, this expresses the willingness of a bank to buy 1 USD at 1.2324 CAD and to sell 1 USD at 1.2326 CAD. The difference between bid and offer prices, called the *spread*, can be viewed as the risk of making one unit currency transaction. It is the (raw) profit that a market maker can generate upon trading with clients for one unit of the base currency. In this example, the spread for USD/CAD is 0.0002, which is also quoted as “2 pips” (where 0.0001 = 1 pip for USD/CAD). For G10 currencies exclude JPY and GBP, pip resolutions for exchange rates are set at the fourth decimal place. For JPY and GBP, the pip resolutions are set at the third and fifth decimal places respectively. Spread is often used as a measure of the liquidity of the currency pair. In particular, the smaller the spread is, the more liquid the currency pair is. For some of the most liquid currency pairs such as EUR/USD, EUR/GBP, and USD/CAD, the spreads are usually very small, at 0.5 or 1 pip during intra-day busy trading hours. For some of the currencies that are lack of liquidity, the bid-offer spread can be as large as 20 or 30 pips. An example of such currencies is MXN/TRY (Mexican peso and Turkish lira).

## 1.4 FX Market Participants

The participants in the FX market include central banks, commercial and investment banks, funds, corporations, and individuals. Unlike the stock market, access to FX market is divided into several levels. The top tier access level is the interbank market, which is made up of the largest commercial banks and securities dealers. They are responsible for 53% of the global transactions. The next tier is composed of large

hedge funds, pension funds, insurance companies, and multi-national corporations who may need to execute an FX trading for the purpose of hedging currency risk, speculating, foreign asset investment and payment of imports, etc. The third tier is the group of individual investors (both long-term and short-term), who constitute a growing segment of the market and mostly participates indirectly through brokers or market makers.

FX market participants operate with varying perspectives. Each perspective carries a different attitude, goal, investment horizon, risk appetite, and market impact with it. According to [22], these participants can be categorized into five groups according to their perspectives. *Active Hedgers*, who are mostly corporations, are long-term players who seek a profit protection through treasury management. *Market Disruptors*, who are usually governments, are long-term enablers of national, regional, or global economic goals. *Risk Avoiders*, which are usually investment fund specialists, are long-term trend followers with high levels of skills, resources, knowledge, and commitments. *Risk Takers*, who are usually individual traders, are short-term system followers with a wide range of skills, knowledge, resources, and commitment. Lastly, *Market Makers*, who are usually banks and dealers, are the credit suppliers to corporations, governments, funds, other banks, and individual traders.

The key difference among these market participants is their levels of sophistication which include: money management techniques, profit objectives, technologies, quantitative abilities, research abilities, and discipline. In terms of regulation, individual traders have the least amount of external governance; whereas governments, banks, corporations, and investment funds must adhere to a maximum amount of financial regulations and restrictions. Individual traders fall into two groups: sophisticated traders and un-sophisticated traders. In [22], the author states that *“In the zero-sum game of the FX trading, the sophisticated traders impose self-disciplines and use tools and strategies that emulate those of the highly sophisticated institutional participants to extract profits from the novice participants. In the end it is only the sophisticated*

*participants who have the ability to extract positive returns from the FX markets”.*

## 1.5 Interbank Market and Market Making

FX market is a decentralized market. In a centralized market such as New York Stock Exchanges (NYSE) or the Chicago Board of Trade, each transaction is recorded according to price dealt and size traded. There is usually a central physical place back to which all trades can be traced. The FX market, however, is a decentralized market, where there are more than one “exchanges” that record every trade. Instead, each market maker records his or her own transactions and keeps it as his/her proprietary information. According to [26], the primary market makers who make bid and offer prices in the FX market are the largest banks in the world. They deal with each other constantly either on behalf of themselves or their customers. This is why the market on which banks conduct transactions is called the Interbank Market. See Table 1.2 for the list of top 10 FX market participants by the market trading volume<sup>3</sup>. The competition between banks ensures tight spreads and fair pricing. Most individuals are unable to access the pricing available on the interbank market because the interbank participants tend to include the largest mutual funds and hedge funds in the world as well as large multinational corporations who operates in millions (if not billions) of dollars.

Market making simply means being a buyer and a seller at the same time. A FX market maker (usually a bank) is looking for opportunities to buy low and sell high with as many clients as possible. A customer, can be a tourist walking into a local branch or a hedge fund manager calling into the FX sales desk and makes FX trading deals with the market maker. As the counter party of the customer, the market maker is using its own capital to trade with the customer at a certain rate that is agreed between the two parties. From [17] and [2], we know that in order

---

<sup>3</sup>Source: Euromoney FX survey, FX Poll 2010. The Euromoney FX survey is the largest global poll of FX services providers.

Table 1.2: Top 10 FX Market Participants by % of Overall Market Volume

Deutsche Bank	18.06%
UBS AG	11.30%
Barclays Capital	11.08%
Citi	7.69%
Royal Bank of Scotland	6.50%
JPMorgan	6.35%
HSBC	4.55%
Credit Suisse	4.44%
Goldman Sachs	4.28%
Morgan Stanley	2.91%

to facilitate market making business, each bank is structured differently, but most banks will have a separate group known as the Foreign Exchange Sales and Trading Department. This group is responsible for making prices for the bank's clients and for offsetting that risk with other banks. Within the foreign exchange group, there is a sales desk and a trading desk. The sales desk is generally responsible for taking the orders from the clients, getting a quote from the spot trader and relaying the quote to the clients to see if they want to deal on it. This three-step process in the industry is quite common because even though online foreign exchange trading is available, many of the large clients, such as pension funds or big corporations, who deal anywhere from \$10 million to \$100 million at a time (cash on cash), believe that they can get better pricing by dealing over the telephone than over the trading platform. This is because most on-line platforms offered by banks will have a trading size limit due to the desire of the bank to be able to offset the risk.

As the market maker, bank dealers determine their prices based upon a variety of factors including: the current market rate, how much volume is available at the current price level, their views on where the currency pair is headed to and their current inventory positions. If they think that the euros is moving upwards, they may be willing to offer a more competitive rate to clients who want to sell euros

because they believe that once they are given the euros, they can hold onto them for a while for the price to increase. On the other hand, if they think that the euro is headed toward a lower value and the client is giving them euros, they may offer a lower price because they are not sure if they can sell the euro back to the market at the same level at which it would be given to them.

To offset the risk, the bank dealers will turn to the interbank market and try to flatten their positions by putting orders into the market. There are two primary electronic platforms that interbank traders use to put their orders. One is offered by Reuters Dealing and the other is offered by the Electronic Broking Services (EBS), which will be introduced in section 1.6. The interbank market is a credit-approved system in which banks trade based solely on the credit relationships they have established with other banks. All of the banks can see the best market rates currently available; however, each bank must have a specific credit relationship with another bank in order to trade at the rates being offered. The bigger the bank is, the more credit relationships it can have and the better pricing it will be able to access. The same is true for clients such as retail FX brokers. The larger the retail FX broker in terms of the available capital, the more favorable pricing it can get from the interbank market. If a client or a bank is relatively small, it is usually restricted to dealing with only a selected number of larger banks and tends to get less favorable pricing.

## **1.6 Electronic Broking System**

Nowadays, over 90% the FX spot transactions goes through automated electronic order-matching systems. Electronic brokers collect orders from tens of thousands of market players globally by connecting to their networks and match their orders automatically. As such, they are perfectly suited to a decentralized market in need of efficient matching. The foreign exchange market, with its decentralized structure and quickly growing volumes, was one of the earliest adopters of electronic brokers.

Subsequently, many equity markets also adopted electronic brokers.

According to [26], the most popular electronic broker systems are Reuters Dealing 2000 and Electronic Broking Services (EBS) in the FX interbank market. The first, Reuters Dealing 2000-2, was introduced by Reuters in April 1992. A year later, in April 1993, Minex was launched by Japanese banks, with EBS following in September 1993. The EBS Partnership was established by several major market making banks to counter the dominant role of Reuters, and EBS acquired Minex in December 1995 and thereby gained a significant market share in Asia.

The electronic brokers work with the goal of matching the buyer and the seller as efficiently as possible. When a *limit order*<sup>4</sup> is entered, there is first a price priority to ensure that it is always the best prices that are traded on and then a time priority (price-time priority). *Market orders*<sup>5</sup> are given priority according to time of entry, and the system matches the counter-parties automatically. The entry of orders is anonymous, but both parties see the counter-party's identity immediately after the trade. Figure 1.2 is borrowed from [26] and shows the trading screens of both Reuters Dealing and EBS.

Part a of Figure 1.2 shows both Reuters Dealing 2000-1 and Dealing 2000-2. The middle panel contains the D2000-1 system for direct bilateral trading, and the upper panel is the D2000-2 electronic broker. From the D2000-1 panel, we can see that the dealer has been contacted for a quote for USD 4 million against DEM. The dealer replies with the quote "05 08", which is understood to be bid 1.8305 and ask 1.8308. The contacting dealer responds with "I BUY", and the system automatically fills in the line "TO CONFIRM AT 1.8308 I SELL 4 MIO USD". In the lower right corner of the screen, the dealer can see the price and direction of the last trades through the D2000-2 system.

Part b of Figure 1.2 shows the EBS screen. The left half of the EBS screen shows

---

<sup>4</sup>A limit order is an order to buy a specified quantity up to a maximum price or sell subject to a minimum price.

<sup>5</sup>A market order involves buying or selling a specified quantity at the best prevailing price.

a

D2000 TRADE <ETEC> User: <eech>

Setup Trade Modify Display CancelOrders Admin Spo CHF DEM Spo Help

DEM Best Trader

1 usd/dem 1.83 **05 / 06** 1.83 5xR 05 \* 06 50x60

21:24:47 Sell 4M usd/dem 1.8305 ETED\* m1481187 t172192  
 21:22:31 #197 BID usd/dem 1.8305 50M OK...  
 21:22:09 #196 BID usd/dem 1.8304 50M OK...

NO CURRENT PROPOSALS

USD DEM SPOT SELL 1.8308 4MIO **STWB** FORM DEAL BANK B **SEND 1**

DEM 4 PLS FRDS  
 # 05 08  
 1 BUY  
 # TO CONFIRM AT 1.8308 | SELL 4 MIO USD

CALLS 0/24  
 PROPOSALS 0

SPOT -EXTRACTED  
 SPOT SELL  
 4 MIO USD  
 1.8308 USD/DEM  
 4AUG97  
 PAY ??? & ???

TRD **MON READY**

Instrument	Type	Quote	Qty	Status
gbp/USD	Bid	1.6280	50	
gbp/USD	Offer	1.6282	75	
USD/CHF	Bid	1.5094	10	
USD/CHF	Offer	1.5096	10	
USD/DEM	Bid	1.8305	50	

21:27:17	gbp/USD	Bid	1.6280
21:26:15	USD/JPY	Bid	118.26
21:25:44	USD/JPY	Given	118.27
21:25:44	USD/JPY	Offer	118.27
21:25:44	USD/JPY	Bid	118.25
21:24:47	USD/DEM	Paid	1.8305
21:24:47	USD/DEM	Offer	1.8305
21:24:47	USD/DEM	Bid	1.8305
21:22:32	USD/DEM	Bid	1.8305
21:22:09	USD/DEM	Bid	1.8304
21:21:15	USD/DEM	Offer	1.8305
21:20:09	USD/DEM	Given	1.8305

b

BALA MOC Mar/ 17.10 | ABL AMRO Credit Limit | INDUSUEZ Credit Limit | Citibank Credit Limit | BANK OF AMERICA Credit Limit

eur/USD 1.09 14 16 1.09 10-Mar

**14** **14** **16** **16**

bid offer

eur/GBP 0.67 34 35 0.67 10-Mar

**34** **34** **35** **35**

bid offer

USD/CHF 1.45 49 50 1.45 10-Mar

**49** **49** **50** **50**

bid offer

USD/JPY 121. 17 19 121. 10-Mar

**17** **17** **19** **19**

bid offer

GBP/USD 1.61 50 53 1.61 10-Mar

**50** **50** **53** **53**

bid offer

eur/JPY 132. 33 34 132. 10-Mar

**33** **33** **34** **34**

bid offer

Rates

EUR/USD 1.0914-16 GBP/USD 1.6150-53  
 EUR/GBP 0.6734-35 EUR/CHF 1.5928-30  
 USD/CHF 1.4549-50 USD/HKD 7.7462-63  
 USD/JPY 121.17-19 EUR/JPY 132.33-34

Trader Deals

17:03	BUY	10	132.88	LCTC	EUR/JPY
17:03	BUY	10	132.89	SB SG	EUR/JPY
17:03	SELL	10	121.34	TKWD	USD/JPY
17:03	SELL	10	1.4555	JNRB	USD/CHF
17:03	BUY	2	1.4552	JMBP	USD/CHF
17:03	SELL	10	1.6125	SLMC	GBP/USD
17:04	BUY	10	0.6821	WNSB	EUR/GBP
17:04	SELL	10	0.6720	SB SG	EUR/GBP
17:04	SELL	10	132.87	TKWD	EUR/JPY
17:07	BUY	10	1.0918	JMBP	EUR/USD

EBS Deals

16:56	1.0924	Given	EUR/USD
17:03	132.89	Paid	EUR/JPY
17:03	121.17	Given	USD/JPY
17:03	1.4552	Given	USD/CHF
17:03	1.4555	Paid	USD/CHF
17:03	1.6125	Given	GBP/USD
17:04	0.6821	Paid	EUR/GBP
17:04	0.6720	Given	EUR/GBP
17:04	132.89	Given	EUR/JPY
17:07	1.0918	Given	EUR/USD

EUR/USD 1.09 **18** **BID**

03377 Buy 10 @ 1.0918 JMBP

DONE

clear **10** of 10 97-11

Figure 1.2: Reuters Dealing and EBS Trading Screens

the bid and offer (ask) prices. The dealer chooses which exchange rates to display (the base currency is written first). The prices shown are either the best prices in the market or the best available ones (from credit-approved banks only). The upper part of the right half of the screen shows the dealer's own trade. The lower part shows the price and direction of all trades through the system for selected exchange rates. *GIVEN* means that it was traded at the bid price, and *PAID* means it was traded at the ask price. The intuition for this is that the limit order dealer is given the base currency (buys).

More discussions about dealer behaviors (for both electronic and traditional voice brokers), liquidity, transaction costs on electronic FX trading can be found in [26] and [15].

## 1.7 FX High-Frequency Trading

In general, high-frequency trading (HFT) involves the execution of computerized trading strategies characterized by unusually very short position-holding periods, in many cases taking advantage from microstructure inefficiency. In high-frequency trading, programs analyze market data and utilize trading opportunities that may open up for only a fraction of a second to several hours. High-frequency trading, uses quantitative models and computer programs to hold short-term positions in equities, options, futures, ETFs, currencies, and other financial instruments that possess electronic trading capability. High-frequency traders compete on a basis of speed with other high frequency traders, who are not long-term investors (that is, who typically look for opportunities over a period of weeks, months, or years), and compete with each other for very small, consistent profits.

According to the article [6], there are several reasons why the FX market is viewed as the most attractive place for high-frequency trading. First, in the FX market, the spreads are extremely low (at about 1 pip) for most liquid currency pairs such as

EUR/USD, EUR/GBP, etc. If we assume a trader has perfect foresight and can take advantage of every small price spike, he/she can earn (without taking on any leverage) approximately 2 percent of return everyday, or approximately 500 percent during one year. If a trader can not trade at high frequency (but, for example, only once a day), then the annual return potential is only 125 percent. Other things being equal, going to HFT enhances the return potential of an investment strategy because a trader can take advantage of many more price strikes. For those sophisticated investment managers equipped with appropriate computing power and know-how, this is seen as great enticement. Secondly, FX market is evolving at a much faster speed than other markets such as equities or futures. Unlike equity and futures markets, where algorithmic trading is developed as a response to a lack of liquidity, the high levels of liquidity access in the FX market allow the market participants to focus more on generating alphas.

To build a solid FX high-frequency trading environment is an extremely challenging exercise. A HFT engine usually contains the following components: liquidity aggregation, trading strategies manager, execution strategies manager, and risk analytic. Liquidity aggregation involves the utilization of today's advanced network and computer technologies to extend the connections to as many market participants and liquidity venues as possible. Aggregating liquidities from different sources, a HFT engine will have a great view of FX market movements at a very low latency. Information about changes in price, volume, and volatility, etc. are coming in on a millisecond basis. And, as a result, better trading decisions and execution results can be achieved with better liquidity aggregation. Thus, access to the liquidity pool is very crucial in FX trading.

Trading strategies manager is the central brain of a high-frequency trading engine, which contains the strategies that are developed by traders and quantitative modelers. These strategies are usually built based on statistical data analysis, previous trading experiences, and alpha research, etc. It is the trading strategies that conduct all

the real-time market data analysis and make trading decisions on buy or sell certain amounts of currency pairs at certain prices. Execution strategies manager is designed to manage different types of orders (for example, IOC<sup>6</sup> and GTC<sup>7</sup>, etc.) and place orders into the market smartly and efficiently to achieve a high successful rate in its execution. It is very important for a HFT engine to be able to catch the best timing for its execution in this milliseconds competition. Risk analytic calculates the real-time risk exposures and measures of the high-frequency trading activities. It is the tool for traders to monitor auto-trading processes that are initiated by the hedging strategies. Traders rely on risk analytic in terms of performing human-intervening for the high-frequency trading engine.

---

<sup>6</sup>IOC: Immediate Or Cancel type of order

<sup>7</sup>GOC: Good Till Cancel type of order

# Chapter 2

## Exponential Moving Average

### 2.1 Market Data

We begin the discussion in this subsection by considering a real example of FX high-frequency time series. The data series is the USD/CAD spot tick-by-tick bid and ask prices on 2010/05/31. The series starts at 00:00:00.259<sup>1</sup> EST, which is the time stamp of the first quote of the day, and ends at 23:59:58.210 EST, which is the time stamp of the last quote update of the day. Each quote update contains a time stamp with accuracy at 1 millisecond, and size and price for both bid and ask sides at that moment. The data is obtained by aggregating feeds of USD/CAD spot from four different venues, which are Reuters, EBS, HotSpot, and Currenex. A series of data cleaning processes needs to be applied to the original raw data set to obtain a “cleaned” version of the data that can be analyzed. A typical data cleaning processes usually includes steps of removing inverted quotes and removing staled and expired quotes.

Figures 2.1 and 2.2 show two different ways of viewing quote updating frequencies for our sample data set.

---

<sup>1</sup>The last three digits of the time stamp are in the unit of milliseconds. It is at 259 millisecond for this time stamp.

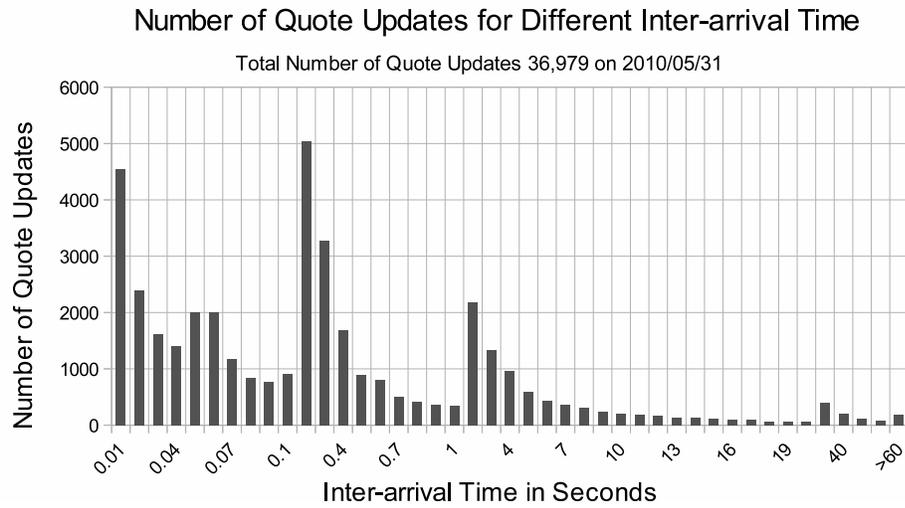


Figure 2.1: Number of Quote Updates for Different Inter-arrival Time on 2010/05/31

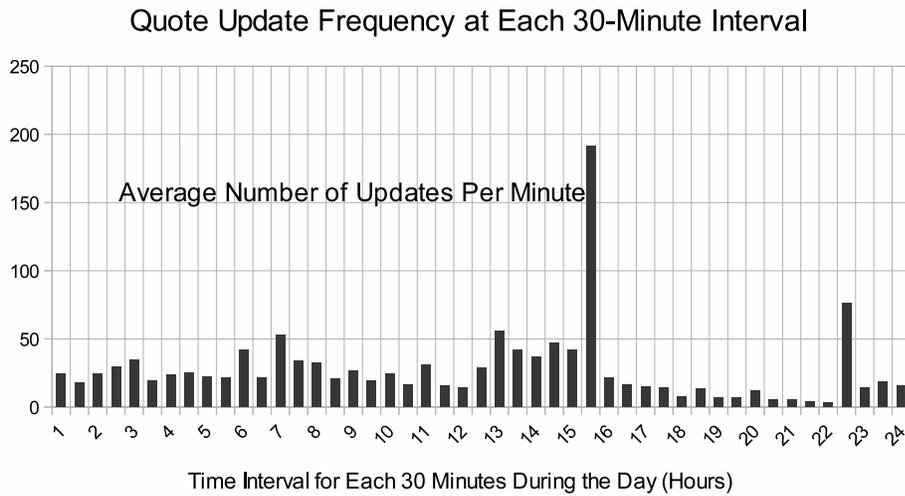


Figure 2.2: Average Number of Quote Updates per Minute for Each 30-Minute Period on 2010/05/31

In particular Figure 2.1 shows the number of quote updates for each different length of *Inter-Arrival Time*, which is defined as the time space between two consecutive quote updates. There are in total 36,979 quote updates during the day of 2010/05/31. Among them, there are 4,549 quote updates with inter-arrival time less than and equal to 10 milliseconds, 17,631 quote updates with inter-arrival time less than and equal to 100 milliseconds, and 5,031 quote updates with inter-arrival time between 100 and 200 milliseconds. Thus, about two-thirds of the total number of quote updates happen at inter-arrival time less than 200 milliseconds. Next Figure 2.2 shows the average number of quote updates per minute for each 30-minute period during the day. We can see that during the North American trading hours (from 7:00 a.m. to 17:00 p.m.), the average number of quote updates ranges from about 20 to 60 per minute, except for a dramatic spike of hitting almost 200 per minute between 15:00 p.m. and 15:30 p.m. In addition, there are also big numbers recorded at 7:00 a.m., 13:00 p.m., and 22:30 p.m., which are the beginning hours of North American morning and afternoon trading sessions and the busy hour in Asia. So, we can see that quote updating can happen at a millisecond level. Frequency of updating can also be different for different periods of time during the day. For example, quote updating speed for USD/CAD between 7:00 a.m. to 11:00 a.m. (busy hours in North America and Europe trading) can be multiple times faster than in the period from 4 p.m. to 5 p.m. Other reasons for a dramatic change in the quote updating frequency include new releases and the announcement of important economic numbers.

Figures 2.3, 2.4, and 2.5 are the time series plots of USD/CAD tick-by-tick prices at different magnitudes of time intervals. Figure 2.3 is the plot of the entire time series (for a 24-hour period) of USD/CAD bid and ask prices. Figure 2.4 is the USD/CAD bid and ask prices for a two-hour period from 10:00 a.m. to 12:00 p.m., during which there are 2,331 quote updates. Figure 2.5 is the plot of bid and ask prices for a ten-minute period from 11:00 a.m. to 11:10 a.m., during which there are 202 quote

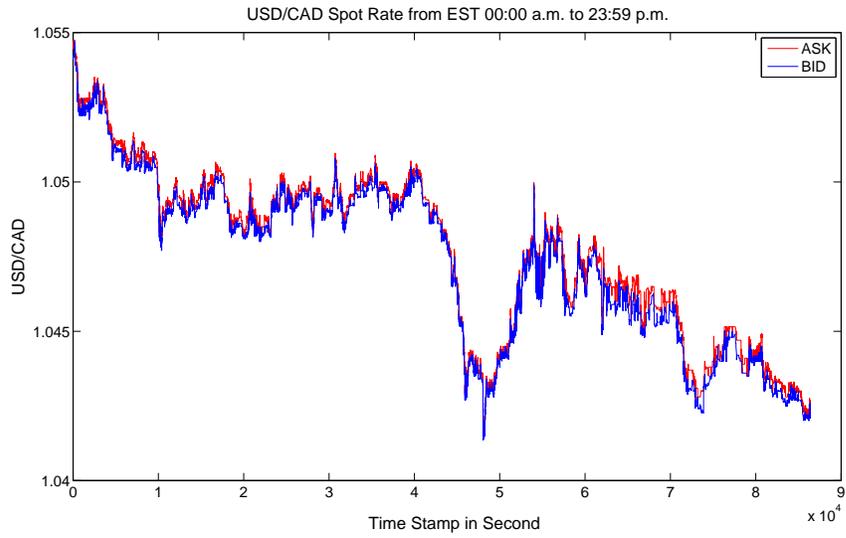


Figure 2.3: USD/CAD Spot Rate from 00:00 a.m. to 23:59 p.m. on 2010/05/31

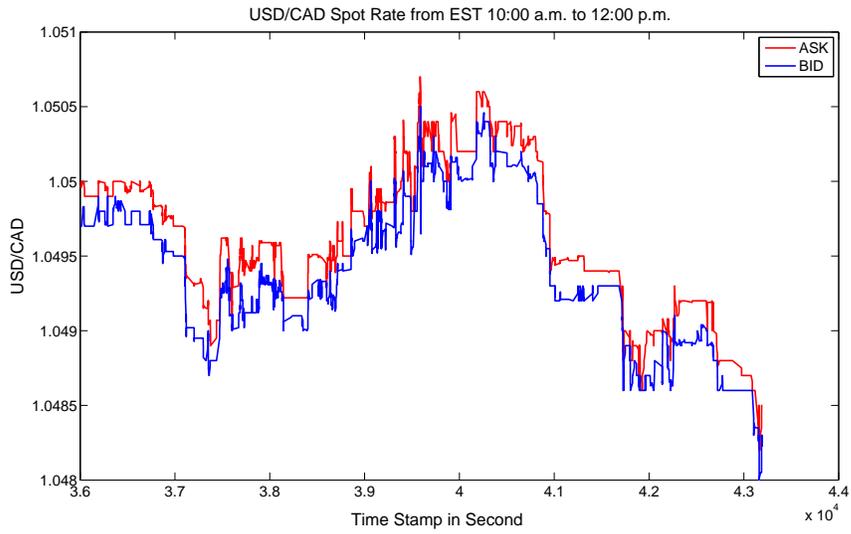


Figure 2.4: USD/CAD Spot Rate from 10:00 a.m. to 12:00 p.m. on 2010/05/31

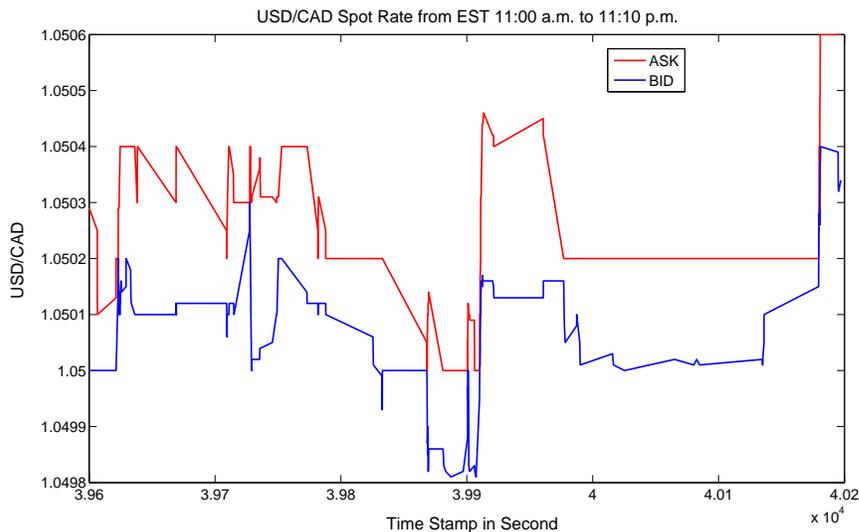


Figure 2.5: USD/CAD Spot Rate from 11:00 a.m. to 11:10 a.m. on 2010/05/31

updates. The first thing that we observe from these plots is that high-frequency data series is a time series with irregular time space, which is referred to as *inhomogeneous* time series. Thus, methods and tools for analyzing *homogeneous* (regular time-spaced) time series need to be modified to handle inhomogeneous high-frequency data analysis. The second thing is that price movements can vary dramatically in a very short period of time, which renders it difficult to make good trading decisions. A market participant has to be very watchful of the fact that both buying and selling decisions need to be made within a few minutes or even seconds. Thus, a market participant must be equipped with both advanced technology and smart trading strategies in order to capture short-term market inefficiency in this high-frequency market.

## 2.2 Exponential Moving Average

A *moving average* process is a widely used in conducting technical analysis on financial data. It is often applied to time series data to smooth out short-term fluctuations

and highlight long-term trends or cycles.

For a continuous function  $x(t)$ , its moving average at time  $t_n$  is defined as an integral

$$MA_{x,\omega} = \frac{\int_{-\infty}^{t_n} \omega(t_n - t)x(t)dt}{\int_{-\infty}^{t_n} \omega(t_n - t)dt}, \quad (2.1)$$

where  $\omega(t)$  is the weighting function defined on non-zero arguments.

The *range* of a moving average is defined as

$$R = \frac{\int_0^{\infty} t\omega(t)dt}{\int_0^{\infty} \omega(t)dt}. \quad (2.2)$$

Specifically, Exponential Moving Average (*EMA*) is a moving average process with the weighting function specified as

$$\omega(t) = \frac{1}{\lambda}e^{-\frac{1}{\lambda}t}, \quad (2.3)$$

where  $\lambda$  is the range of the weight function. This weight function declines exponentially with the time distance of the past observations starting from the present time. The choice of the range value of  $\lambda$  is very important for the exponential moving average operation on time series because it controls the distribution of weights onto past data values. Figure 2.6 shows the EMA weight functions with different range values. We can see that the smaller the range value of  $\lambda$  is, the heavier the weights being assigned to more recent observations, and the faster the weight function decays into past.

If we apply the weight function (2.3) into equation (2.1), we can define the exponential moving average of value  $x$  at time point  $t_n$  as

$$EMA_x(\lambda, t_n) = \frac{\int_{-\infty}^{t_n} \frac{1}{\lambda}e^{-\frac{1}{\lambda}(t_n-t)}x(t)dt}{\int_{-\infty}^{t_n} \frac{1}{\lambda}e^{-\frac{1}{\lambda}(t_n-t)}dt} = \int_{-\infty}^{t_n} \frac{1}{\lambda}e^{-\frac{1}{\lambda}(t_n-t)}x(t)dt \quad (2.4)$$

Based on equation (2.4), EMA can be computed by a recursive method found in [20] and [21]. Unfortunately, the formula, as presented in [20] and [21], are incorrect. Here, we provide a correct version of the formula. In this study, we adopt similar notations as those used in the text book [21]. Let  $Z(t_j)$  represent a raw series with

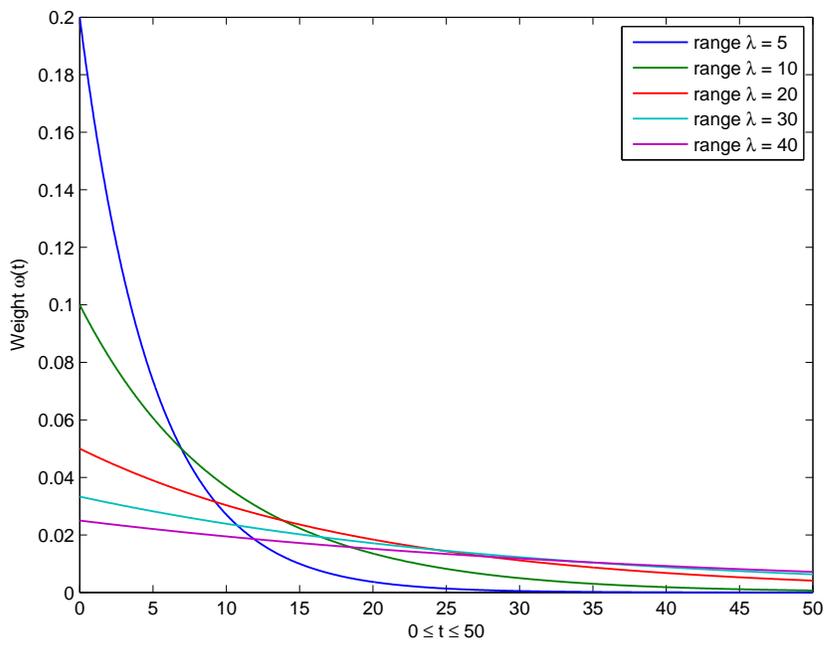


Figure 2.6: EMA Weight Functions with Different Range Values

irregular time spaces at arrival times  $t_j$  where  $j = 0, 1, 2, 3, \dots$ . The sequence of arrival times is required to be monotonically increasing such that  $t_j > t_{j-1}$ . For a time point  $t^*$  such that  $t_{n-1} \leq t^* < t_n$ , the exponential moving average value of time series  $Z(t_j)$  at time  $t^*$  can be obtained by the following recursive formula:

$$EMA_Z(\lambda, t^*) = \mu_1 EMA_Z(\lambda, t_{n-1}) + (\nu_1 - \mu_1)Z(t_{n-1}) + (1 - \nu_1)Z(t_n) \quad (2.5)$$

with  $\mu_1 = e^{-\frac{1}{\lambda}(t^* - t_{n-1})}$  and value of  $\nu_1$  depending on the chosen interpolation scheme for  $Z(t^*)$ , where

$$\nu_1 = \begin{cases} 1 & \text{for the previous tick method} \\ 1 - \frac{t^* - t_{n-1}}{t_n - t_{n-1}} & \text{for the linear interpolation method.} \end{cases}$$

This recursive formula of calculating EMA can be applied to both homogeneous and inhomogeneous time series. It is easy to show that  $0 < \mu_1 < 1$  and  $0 \leq \nu_1 < 1$  given  $t_{n-1} < t^* \leq t_n$ . The values of  $\mu_1$  can be viewed as the weight assigned to the EMA value of the time series at time point  $t_{n-1}$ . Values of  $\nu_1 - \mu_1$  and  $1 - \nu_1$  can be viewed as the weights assigned to  $Z(t_{n-1})$  and  $Z(t_n)$ . Appendix A gives the proof of the above recursive formula. The derivation of this formula assumes the time series to start from time  $-\infty$ , which is not realistic. For practical purpose, here we list an analogous form of the recursive equation that assumes the time series to start from time zero instead.

We let  $Z(t_j)$  with  $j = 0, 1, 2, 3, \dots$  be the raw inhomogeneous time series starting at time point  $t_0 = 0$ . For a time point  $t^*$  such that  $t_{n-1} \leq t^* < t_n$ , the exponential moving average value of time series  $Z(t_j)$  at time  $t^*$  can be obtained by the following recursive formula:

$$EMA_Z(\lambda, t^*) = \mu_2 EMA_Z(\lambda, t_{n-1}) + (\nu_2 - \mu_2)Z(t_{n-1}) + (1 - \nu_2)Z(t_n) \quad (2.6)$$

with  $\mu_2 = \frac{e^{-\frac{1}{\lambda}(t^* - t_{n-1})} - e^{-\frac{1}{\lambda}t^*}}{1 - e^{-\frac{1}{\lambda}t^*}}$  and value of  $\nu_2$  depending on the chosen interpolation

scheme for  $Z(t^*)$ , where

$$\nu_2 = \begin{cases} 1 & \text{for the previous tick method} \\ 1 - \frac{1}{1 - e^{-\frac{1}{\lambda}t^*}} \frac{t^* - t_{n-1}}{t_n - t_{n-1}} & \text{for the linear interpolation method.} \end{cases}$$

The derivation of this formula is provided in Appendix B. EMA can be regarded as an *operator* that transforms one time series into another one:

$$EMA : Z(t_n) \mapsto EMA_Z(\lambda, t_n). \quad (2.7)$$

Due to this recursive formula, the integration need not be computed in practice; instead only few multiplications and additions need to be done for each tick. In this research, we apply the above recursive formula to our FX data.

## 2.3 Application of the EMA Operator

The EMA operator is implemented in matlab according to equation (2.6). For the analysis, we consider a 15-minute (from 10:00 a.m. to 10:15 a.m.) subset of the high-frequency data series on 2010/05/31 as a starting point. Let  $Z(t_k)$  for  $k = 0, 1, 2, \dots, n$  denote  $n$  mid-prices<sup>2</sup> being calculated during this 15-minute time interval. We choose the first observation of the time series as the starting value of the recursive formula. That is, we set  $Z(t_0) = EMA_Z(\lambda, t_0)$ . Then, the EMA values of the mid-prices at each time stamp  $t_k$  for  $1 \leq k < n$  are calculated iteratively.

Figure 2.7 shows the original data series of the mid prices and their EMA values with different values of the range of  $\lambda$  (set at 20, 100, 200, and 600 seconds respectively).

For a value of  $\lambda = 20$  (seconds), we see a small discrepancy between the original data series and the EMA series at the very beginning portion of the plot. Then, the two series merges into almost an identical one, which means that the EMA operator

---

<sup>2</sup>Mid-price = (Ask price + Bid price)/2

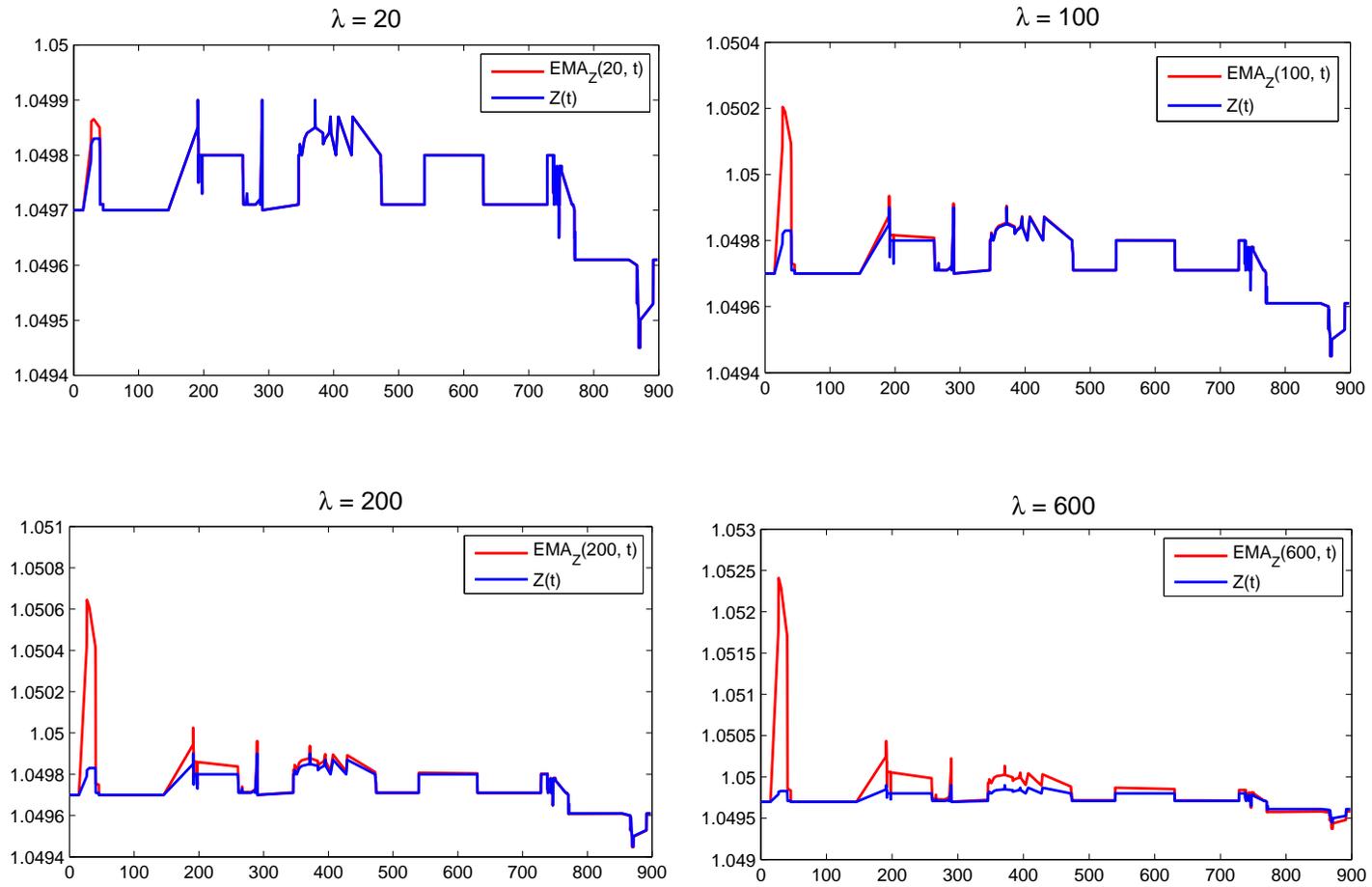


Figure 2.7: Time Series of Mid Price and Its EMA Values with Different Values of Range  $\lambda$

generates estimation very closed to real values. By looking at the other three plots, we can see that as the range of value of  $\lambda$  gets bigger, the discrepancy between the values of the original data and the EMA values are getting bigger, and the longer it takes for the two series to get closed enough to each other. Thus, no matter what value of  $\lambda$  we choose, a built-up time period is necessary for the EMA operator to produce accurate enough values. Empirically, the bigger the range value of  $\lambda$  is, the longer the built-up period is needed for the EMA to produce accurate enough results. This conforms with the rule of thumb given on page 57 in [21]: “*the heavier the tail of the kernel, the longer the required build-up is needed.*”

To get a better picture of how well the EMA operator performs, Figure 2.8 shows the mean square errors (MES) between the true market values and their EMA estimates with different range values of  $\lambda$ . No surprise that we see EMA estimates with larger value of  $\lambda$  has larger MSE values. For each value of  $\lambda$ , MSE starts to decrease and converges after an enough number of observations being made. The larger the  $\lambda$  is, the more observations we need to see MSE starting to decreasing.

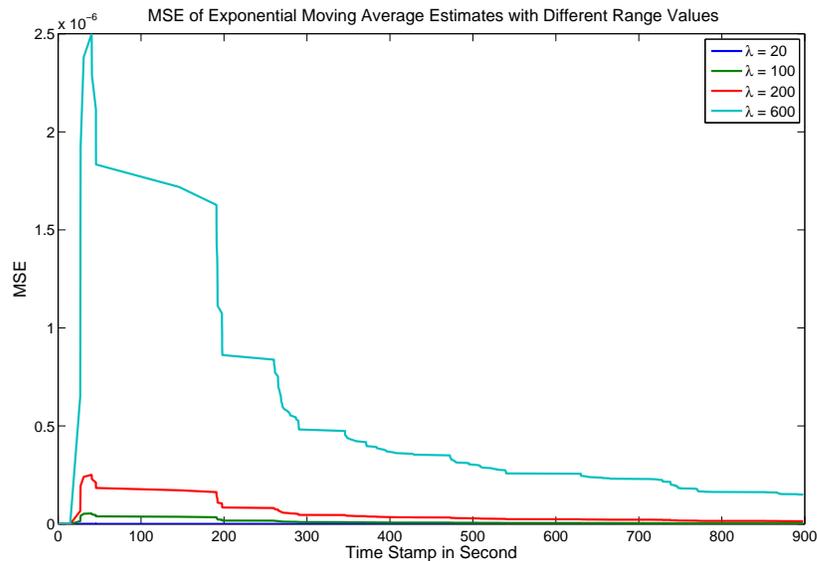


Figure 2.8: Mean Squared Errors of EMA estimates with Different Values of Range  $\lambda$

One possible explanation that Exponential Moving Average is very accurate in estimating high-frequency data is because the time period between two consecutive quote updates is so short that the quote jump is not significant enough to deviate the quote far from its EMA estimate. With EMA operator, a market maker can at least estimate market movements for a very short period (measured in milliseconds) ahead into future. A game of issuing and canceling limit orders within millisecond time intervals can be performed.

# Chapter 3

## Simulation Framework

### 3.1 Motivation

Trading as the counter-party of clients is the core business model of a market maker because client margin spreads (used to) contribute to majority part of profit. As the evolving of advanced technology and market transparency, more and more sophisticated investors start to trade in FX market with access to fast information and liquidity. Market makers can never make money the same way they did 20 years ago. Buying from one client then selling to another one at a higher price can not be done as easily as before. Smart risk hedging strategies must be implemented to help the market maker trade “profitably”. In our opinion, we believe that the hedging strategy should be subjective to client trading flows. This is saying that under different client trading flows, different hedging strategies (or different parameter values for one hedging strategy) should be applied to optimize risk-adjusted returns. According to [1], a market maker should carefully study his/her client trading flows so that non-public information can be extracted from it. For example, client trades can be categorized into different groups such as hedge funds, banks, institutional investors, and retail flows. Transactions done with hedge fund clients provides more useful information than transactions done with retail clients. If a speculative trader from a hedge fund

is buying Euros and selling Dollars, we reach a very different conclusion about the future direction of EUR/USD than if the buying of EUR came from a US importer.

Due to the lack of historical high-frequency data and client trading information, we build a basic simulation framework for market data and client trades in this chapter. Poisson process and Geometric Brownian motion are the natural starting points for event arrival process and asset price simulations.

## 3.2 Poisson Process

A counting process deals with the number of various outcomes of an experiment over a period of time. According to [27], a *counting process* is defined as a stochastic process  $\{N(t), t \geq 0\}$  that has the following properties:

1.  $N(t) \geq 0$ .
2.  $N(t)$  is an integer.
3. If  $s < t$ , then  $N(s) \leq N(t)$ . In other words,  $N(t)$  is non-decreasing.
4. If  $s < t$ , then  $N(t) - N(s)$  is the number of events occurred during time interval  $(s, t)$ .

A Poisson process is a special example of the counting process. A *Poisson Process* with a rate of  $\lambda$  is defined as a continuous-time counting process  $\{N(t), t \geq 0\}$  such that:

1.  $N(t) = 0$  for  $t = 0$ .
2. The process has *Independent Increments*, which means that the numbers of occurrences counted in disjoint intervals are independent from each other.
3. The process has *Stationary Increments*, which means that the probability distribution of the number of occurrences counted in any time interval only depends on the length of the interval.

4. The probability of  $k$  events occurred during a time period of length  $t$  is given by  $P(N(t+s) - P(N(s) = k)) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$ .

The Poisson process is widely used in practice to model events such as the arrival process of incoming calls to a call centre, customer arrival process of a restaurant, the number of cars reaching at a traffic light, etc. A Poisson process with a rate of  $\lambda$  implies that the inter-arrival time between two consecutive events are independently and identically distributed exponential random variables with mean  $1/\lambda$ . An exponentially distributed random variable  $T$  with a mean value  $1/\lambda$  has the cumulative density function given by

$$F(t) = 1 - e^{-\lambda t}, \forall t \geq 0. \quad (3.1)$$

Thus, a simulation of a Poisson process is equivalent to a simulation of series of exponentially distributed inter-arrival time. In this research, we will use it to model both the *client trading arrival process* and the *market data arrival process*.

### 3.3 Geometric Brownian Motion

Geometric Brownian Motion (GMB) has been applied widely in modelling asset price movements in both academic and industry research. In [30], the author assumes the fundamental value of the securities follows a Brownian motion, reflecting the fact that in absence of any trades, the mid-quote price may change due to news about the fundamental value of the security. In section 3.4, we will adopt Geometric Brownian Motion to the high-frequency FX spot exchange rate simulation. Let us lay out the basic framework of modeling an asset price by using GMB.

According to [29], let  $\{W(t), t \geq 0\}$  be a Brownian motion. Let  $\{\mathcal{F}(t), t \geq 0\}$  be an associated filtration, and let  $\{\alpha(t), t \geq 0\}$  and  $\{\sigma(t), t \geq 0\}$  be adapted processes. An *Ito Process*  $X(t)$  can be defined as

$$X(t) = \int_0^t \sigma(s) dW(s) + \int_0^t (\alpha(s) - \frac{1}{2}\sigma^2(s)) ds, \quad (3.2)$$

which has the differential form

$$dX(t) = \sigma(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\sigma^2(t)\right)dt. \quad (3.3)$$

Next let us consider an asset with a price process following an Ito process given by

$$S(t) = S(0)e^{X(t)} = S(0) \exp \left\{ \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right)ds \right\}. \quad (3.4)$$

Then, we can apply Ito's formula to  $S(t)$  in equation (3.4) and obtain

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t), \quad (3.5)$$

or equivalently

$$\frac{dS(t)}{S(t)} = \alpha(t)dt + \sigma(t)dW(t), \quad (3.6)$$

where  $\alpha(t)$  and  $\sigma(t)$  are the instantaneous mean rate of return and volatility respectively. Both  $\alpha(t)$  and  $\sigma(t)$  can be allowed to be time varying or time invariant. By using a Geometric Brownian motion, we assume that the unit incremental amount  $\frac{dS(t)}{S(t)}$  during period  $\Delta t$  is a normal random variable with mean  $\alpha(t)\Delta t$  and variance  $\sigma^2(t)\Delta t$ .

### 3.4 Simulation of Market Data

One of the objectives of this research is to build a simulation framework for market data with different volatility assumptions. High-frequency market data series can be decomposed into two parts: *market data arrival process* and *market price values*. We will illustrate how to simulate them by using a Poisson process and an Ito process separately.

Let  $\{M(t), t = 0, 1, 2, \dots\}$  be a Poisson process with an arriving rate of  $\lambda_M$ . This represents the USD/CAD market data arrival process. Let us assume that for a two-hour simulation period (that is, for  $D = 7200$  seconds), the market data has an average arrival rate of 1 quote update for each 2 seconds (that is,  $\lambda_M = 1/2$ ). Then,

we can simulate a series of inter-arrival time which is exponentially distributed with mean  $1/\lambda_M = 2$  seconds by applying equation (3.1) and the *Inverse Transformation Method*<sup>1</sup>. The process can be stated as follows:

1. Calculate the estimated number of arrival times for period  $D$  by  $n_M = \frac{D}{\lambda_M}$
2. Calculate a series of estimated inter-arrival time  $\widetilde{\Delta t} = \{\Delta t_1, \Delta t_2, \dots, \Delta t_{n_M}\}$  by calculating  $\Delta t_k = \frac{1}{\lambda_M} \log(U(0, 1))$ , where  $U(0, 1)$  is a uniform  $[0, 1]$  random variable for each  $k = 1, 2, \dots, n_M$ .
3. The  $K^{th}$  market data arrival time stamp can be obtained as  $T_K = \sum_{k=1}^{k=K} \Delta t_k$  for  $K \leq n_M$ , and the series of market data arrival time is then given by  $\widetilde{T} = \{T_1, T_2, \dots, T_{n_M}\}$ .

Now we have the simulation results of the market data arrival process. The histogram of simulated inter-arrival time series  $\widetilde{\Delta t}$  is shown in Figure 3.1. There are in total 3,572 inter-arrival time being simulated with about 2,200 of them are between 0 to 2 seconds.

The next step is to simulate the USD/CAD market mid-price value at each market data arrival time (simulated above) by an Ito process. Equation (3.5) is implemented in matlab as a function with five inputs: the initial asset mid-price  $P_{mid}(0)$ , a fixed value of drift  $\alpha$  per unit of time, a fixed value of volatility  $\sigma$  per unit of time, a series of simulated inter-arrival time  $\widetilde{\Delta t}$ , and a series of  $N(0, 1)$  distributed scores. Then, at each market data arrival time stamp  $T_K$  for  $K = 1, 2, \dots, n_M$ , we calculate the market mid-price as

$$P_{mid}(T_K) = P_{mid}(0) \exp \left\{ \sum_{k=1}^K \sigma Z(0, \Delta t_k) + \sum_{k=1}^K \left( \alpha - \frac{1}{2} \sigma^2 \right) \Delta t_k \right\}, \quad (3.7)$$

where  $P_{mid}(0)$  is the starting value of the process and  $Z(0, \Delta t_k)$  is a normal random variable with mean 0 and variance  $\Delta t_k$ . If we assume that the market spread value

---

<sup>1</sup>Inverse Transformation Method: if  $Y$  has a uniform distribution on  $[0, 1]$  and if  $X$  has a cumulative distribution denoted as  $F_X$ , then the cumulative distribution function of the random variable is given by  $F_X^{-1}(Y)$  is  $F_X$

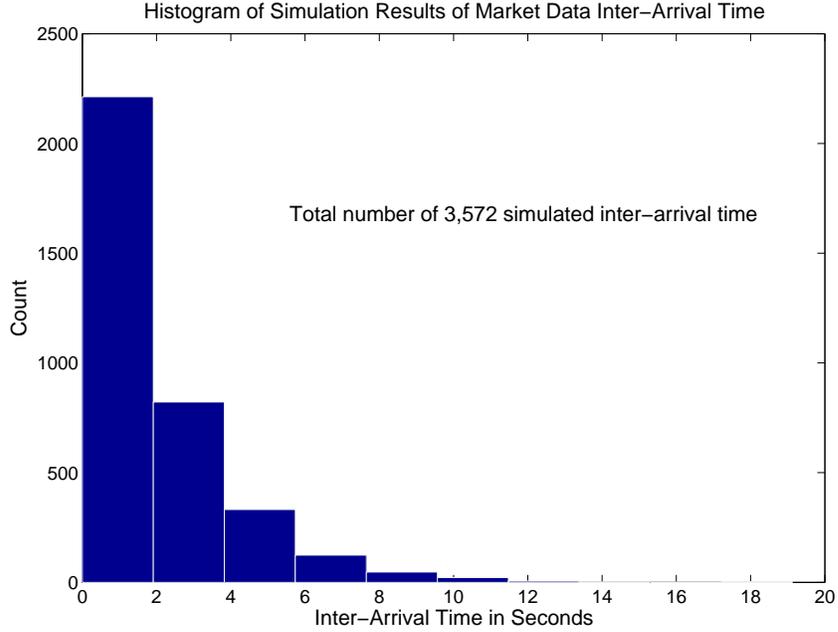


Figure 3.1: Histogram of Simulation for USD/CAD Market Data Inter-Arrival Time

during the simulation period is fixed at  $\delta$ , then the Market bid and ask prices can be obtained as

$$P_{bid}(T_K) = P_{mid}(T_K) - 0.5\delta \quad (3.8)$$

$$P_{ask}(T_K) = P_{mid}(T_K) + 0.5\delta. \quad (3.9)$$

Let us assume that the initial mid-price  $P_{mid}(0)$  of USD/CAD is at 1.1212, the drift and volatility for the two-hour simulation period are at 0.01 and 0.5 pips each unit of time (in minute) respectively, and the spread is fixed at 0.5 pips. Given the series of inter-arrival time  $\widetilde{\Delta t}$ , the sample paths of USD/CAD bid and ask prices for a two-hour period can be obtained and shown in Figure 3.2.

### 3.5 Simulation of Client Trades

Similar to the market data process, a client trading process can also be decomposed into two parts: *client trading arrival process* and *client trading amounts*. We again

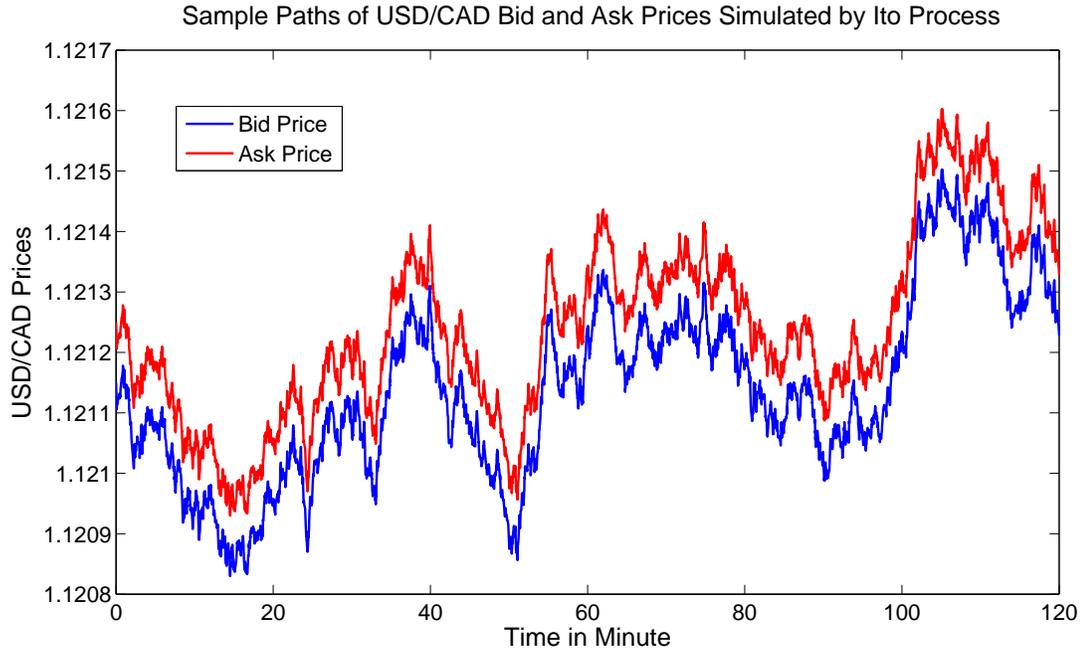


Figure 3.2: Sample Paths of USD/CAD Bid and Ask Prices by an Ito Process

apply a Poisson process in the simulation of client trading arrival process. For the client trading amount, we assume for simplicity that it follows a modified version of the normal distribution with fixed values of mean and standard deviation.

In order to increase the flexibility of the model, we simulate the client buying and selling trading processes separately. Let  $\{N_1(t), t = 0, 1, 2, \dots\}$  and  $\{N_2(t), t = 0, 1, 2, \dots\}$  be two Poisson processes with arriving rates of  $\lambda_{N_1}$  and  $\lambda_{N_2}$  respectively to represent the USD/CAD client buying and selling trades arrival processes. Let us assume that for the two-hour simulation period, the client buying and selling trading arrival processes have an average arrival rate of 1 buying trade and 1 selling trade for each 2-minute interval (that is, we set  $\lambda_{N_1} = \lambda_{N_2} = 1/120$ ). Then, by the same methodology used in the simulation of market data arrival process, we obtain the series of client buying trades arrival time and selling trades arrival time as  $\widetilde{T}_B = (T_{B_1}, T_{B_2}, \dots, T_{B_{n_1}})$  and  $\widetilde{T}_S = (T_{S_1}, T_{S_2}, \dots, T_{S_{n_2}})$  respectively, where  $n_1$  and  $n_2$  are the numbers of client buying and selling trades happening during the simulation

period.

The next step is to simulate the client buying and selling amounts at each client trading arrival time listed in  $\widetilde{T}_B$  and  $\widetilde{T}_S$ . We assume two random variables  $Y_1$  and  $Y_2$  to represent client buying and selling amounts in terms of base currency<sup>2</sup> such that  $Y_1 = |X_1|$  and  $Y_2 = |X_2|$ , where  $X_1 \sim N(\mu_{X_1}, \sigma_{X_1}^2)$  and  $X_2 \sim N(\mu_{X_2}, \sigma_{X_2}^2)$ . By applying the absolute values onto random variables  $X_1$  and  $X_2$ , we simply enlarge the probability density for far-tail values if  $\mu_{X_1}$  and  $\mu_{X_2}$  are enough far from 0. This is a reasonable assumption because we believe that the client trading amount has a distribution with heavier tails than normal distribution. Thus, for the FX spot trading, the market maker's *Base Currency Wealth Process* of trading as the counterparty of its clients at time  $t$  can be defined as

$$W_1(t) = W_1(0) - \sum_{T_{B_i} \leq t} Y_1(T_{B_i}) + \sum_{T_{S_i} \leq t} Y_2(T_{S_i}), \quad (3.10)$$

where  $W_1(0) = 0$  is the initial value of the wealth in the base currency. Since currency trading is always in pair, buying one currency comes with selling another currency and vice versa. Then, given the market bid and ask prices as equations (3.8) and (3.9), the *Counter Currency Wealth Process*  $W_2(t)$  can be defined as

$$W_2(t) = W_2(0) + \sum_{T_{B_i} \leq t} Y_1(T_{B_i})[P_{ask}(T_{B_i}) + \delta_C] - \sum_{T_{S_i} \leq t} Y_2(T_{S_i})[P_{bid}(T_{S_i}) - \delta_C], \quad (3.11)$$

where  $W_2(0) = 0$  is the initial value of the counter currency wealth process and  $\delta_C \geq 0$  represents client margin. As the market maker, we quote price  $P_{ask}(T_{B_i}) + \delta_C$  to the client who wants to sell us the base currency, and quote price  $P_{bid}(T_{S_i}) - \delta_C$  to the clients who want to buy the base currency from us. Client margins are different by client types and requested trading amounts. Usually the larger the requested amount is, the large the client margin is. Processes  $W_1(t)$  and  $W_2(t)$  illustrates how a market maker's position changes according to pure client trading.

---

<sup>2</sup>For a currency pair XXX/YYYY, the client trading amounts are always quoted in the amount of currency XXX.

On the other hand, we can combine the two series of client buying and selling arrival time  $\widetilde{T}_B = (T_{B_1}, T_{B_2}, \dots, T_{B_{n_1}})$  and  $\widetilde{T}_S = (T_{S_1}, T_{S_2}, \dots, T_{S_{n_2}})$  and sort them into one monotonically increasing series  $\widetilde{T}_C = (T_{C_1}, T_{C_2}, \dots, T_{C_{(n_1+n_2)}})$ , which is the time series of client trading arrival time regardless of buying or selling activities. Then, at each client trading arrival time  $T_{C_k}$  for  $k = 1, 2, \dots, (n_1 + n_2)$ , we can rewrite equations (3.10) and (3.11), Base and Counter Currency Wealth Processes, into recursive formulas as follows:

$$W_1(T_{C_k}) = W_1(T_{C_{k-1}}) - Y_1(T_{C_k})I[Y_1(T_{C_k}) > 0] + Y_2(T_{C_k})I[Y_2(T_{C_k}) > 0] \quad (3.12)$$

and

$$\begin{aligned} W_2(T_{C_k}) &= W_2(T_{C_{k-1}}) + Y_1(T_{C_k})[P_{ask}(T_{C_k}) + \delta_C]I[Y_1(T_{C_k}) > 0] \\ &\quad - Y_2(T_{C_k})[P_{bid}(T_{C_k}) - \delta_C]I[Y_2(T_{C_k}) > 0] \end{aligned} \quad (3.13)$$

with starting values of  $W_1(T_{C_0}) = 0$  and  $W_2(T_{C_0}) = 0$ . Indicator functions  $I[Y_1(T_{C_k}) > 0]$  and  $I[Y_2(T_{C_k}) > 0]$  tell us if it is a client buying or selling trade at time  $T_{C_k}$ . In our research, we assume that only one event of either client buying or selling trade can happen at each time stamp. Thus events  $\{Y_1(T_{C_k}) > 0\}$  and  $\{Y_2(T_{C_k}) > 0\}$  are complement of each other.

The market maker's real-time P&L<sup>3</sup> of trading as the counter party with clients can be calculated as

$$PL(t) = W_2(t) + W_2^*(t), \quad (3.14)$$

where

$$W_2^*(t) = \begin{cases} W_1(t)P_{bid}(t) & \text{if } W_1(t) \geq 0 \\ W_1(t)P_{ask}(t) & \text{if } W_1(t) < 0. \end{cases} \quad (3.15)$$

Note that symbol  $t$  in above equations for P&L calculation can be substituted by symbol  $T_{C_k}$  for  $k = 1, 2, \dots, (n_1 + n_2)$ . P&L is measured in the unit of counter currency.

---

<sup>3</sup>P&L is an abbreviation term for Profit and Loss.

If we set the parameter values  $\mu_{X_1} = \mu_{X_2} = 500,000$ ,  $\sigma_{X_1}^2 = \sigma_{X_2}^2 = 1,000,000$ ,  $\delta_C = 0.00005$ , and use the simulation results of market bid and ask prices for USD/CAD in the previous section 3.4, then we obtain sample paths of market maker's Base and Counter Currency Wealth Processes  $W_1(T_{C_k})$  and  $W_2(T_{C_k})$  for  $k = 1, 2, \dots, (n_1 + n_2)$ . Figure 3.3 shows the simulation results of  $W_1(T_{C_k})$  and  $W_2(T_{C_k})$ . There is no surprise that these two paths are nearly mirrors of each other. The P&L values  $PL(T_{C_k})$  for  $k = 1, 2, \dots, (n_1 + n_2)$  are also calculated and shown in Figure 3.4.

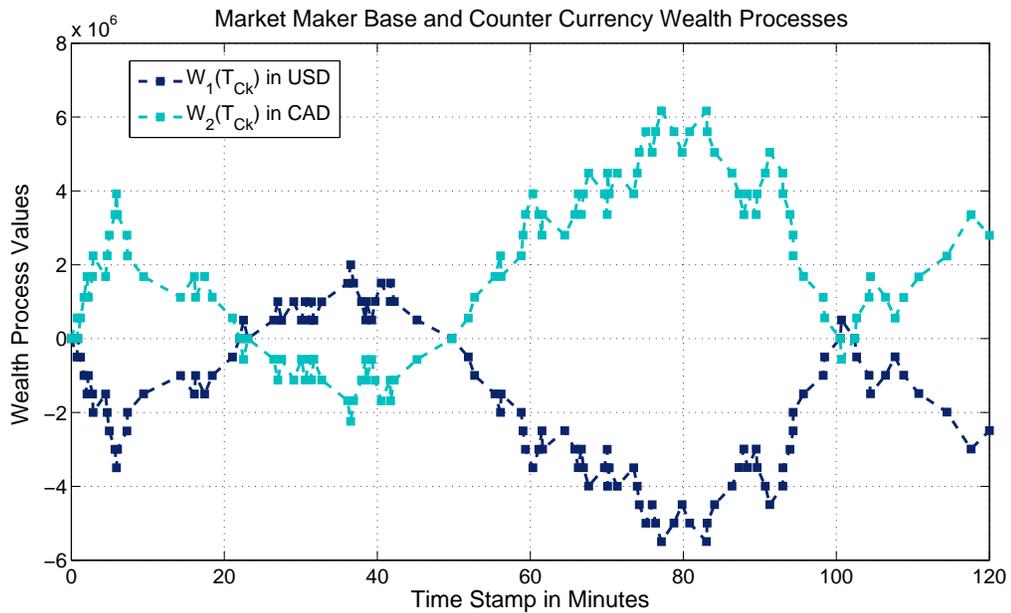


Figure 3.3: Sample Paths of Market Maker's Base and Counter Currency Wealth Processes

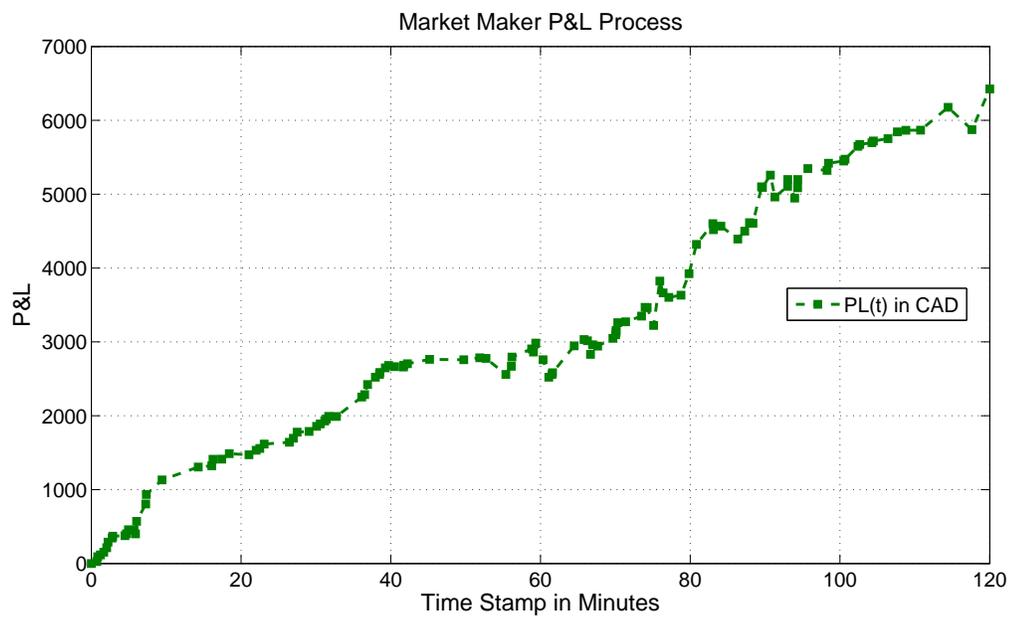


Figure 3.4: Sample Path of Market Maker's P&L

# Chapter 4

## Risk Hedging

### 4.1 Hedging Strategy

In Section 1.5, we explained how a market maker trades with clients by both buying and selling with its own capital. The most ideal situation would be buying from one client at market bid and selling the same amount to another client at market offer at the same time. But this is rarely the case in practice because it is very hard to get two clients to request the same amount of trades on each side at the same time. Thus, a market maker will have to hold positions (positive or negative) for a period of time, and this introduces a substantial amount of market risk to market maker's portfolio. Therefore it is important for a market maker to actively trade during the day for an effective risk management.

In this chapter, we introduce a basic risk hedging strategy that generates trades based on the market maker's Base Currency Wealth Process  $\{W_1(C_k), k = 1, 2, \dots, (n_1 + n_2)\}$ . The intuition underlying this strategy is that a market maker is usually unwilling to hold very big positions at any time due to a market risk exposure. Note that the definition of "big" is subjective to market maker's risk tolerance. This can be determined by many factors such as client trading flows, accessibility to liquidities, efficiency in risk management, etc. For example, a global market player with advanced

technologies to access deep liquidities and smart risk hedging strategies may allow its EUR position staying at 100 million during the day; but a second-tier market player, who does not have the same level of technologies and strategies, may only allow its trader to hold the amount of EUR less than 10 million.

Let two amounts  $U_B$  and  $U_N$  be such that  $U_B > 0$  and  $U_N < 0$  respectively represent the maximum allowable positive and negative amounts in the base currency of the trading pair that the market maker can hold. Then, once the market maker has its Base Currency Wealth Process breaches  $U_P$  or  $U_N$ , a risk hedging trade will be issued to off-load the risk to a lower level. This is pre-defined by the market maker based on his/her risk tolerance. Let us define the two lower levels  $L_P$  and  $L_N$  such that  $U_P > L_P > 0$  and  $U_N < L_N < 0$ . Then, at each client trading arrival time  $T_{C_k}$  for  $k = 1, 2, 3, \dots, (n_1 + n_2)$ , the market maker's *Risk Adjusted Base Currency Wealth Process*  $AW_1(T_{C_k})$  can be defined as a recursive formula as

$$AW_1(T_{C_k}) = AW_1(T_{C_k}^-) + H(T_{C_k})I[H(T_{C_k}) \neq 0], \quad (4.1)$$

where

$$AW_1(T_{C_k}^-) = AW_1(T_{C_{k-1}}) - Y_1(T_{C_k})I[Y_1(T_{C_k}) > 0] + Y_2(T_{C_k})I[Y_2(T_{C_k}) > 0] \quad (4.2)$$

and the Risk Hedging Trading Amount being

$$H(T_{C_k}) = \begin{cases} L_P - AW_1(T_{C_k}^-) & \text{if } AW_1(T_{C_k}^-) > U_P \\ L_N - AW_1(T_{C_k}^-) & \text{if } AW_1(T_{C_k}^-) < U_N \\ 0 & \text{otherwise.} \end{cases} \quad (4.3)$$

The starting value of the process is  $AW_1(T_{C_0}) = 0$ , and the indicator function  $I[H(T_{C_k}) \neq 0]$  equals to 1 when the Risk Hedging Trading Amount is non-zero. Similar to market maker's Counter Currency Wealth Process  $W_2(T_{C_k})$ , the market maker's *Risk Adjusted Counter Currency Wealth Process*  $AW_2(T_{C_k})$  can be defined

as

$$AW_2(T_{C_k}) = AW_2(T_{C_k}^-) + H(T_{C_k}) \left\{ P_{ask}(T_{C_k}) I[H(T_{C_k}) > 0] + P_{bid}(T_{C_k}) I[H(T_{C_k}) < 0] \right\}, \quad (4.4)$$

where

$$\begin{aligned} AW_2(T_{C_k}^-) &= AW_2(T_{C_{k-1}}) + Y_1(T_{C_k}) [P_{ask}(T_{C_k}) + \delta_C] I[Y_1(T_{C_k}) > 0] \\ &\quad - Y_2(T_{C_k}) [P_{bid}(T_{C_k}) - \delta_C] I[Y_2(T_{C_k}) > 0] \end{aligned} \quad (4.5)$$

and function  $H(T_{C_k})$  is given by equation (4.3).

Market maker's Risk Adjusted P&L measured in the counter currency can be calculated as

$$PL_A(t) = AW_2(t) + AW_2^*(t), \quad (4.6)$$

where

$$AW_2^*(t) = \begin{cases} AW_1(t) P_{bid}(t) & \text{if } AW_1(t) \geq 0 \\ AW_1(t) P_{ask}(t) & \text{if } AW_1(t) < 0. \end{cases} \quad (4.7)$$

Again, the symbol  $t$  in above equations can be substituted by the symbol  $T_{C_k}$  for  $k = 1, 2, \dots, (n_1 + n_2)$ . In fact, one crucial assumption we have made about this hedging strategy is that *the FX market is liquid enough so that the market maker can successfully execute the risk hedging trade of amount  $H(T_{C_k})$  defined by equation 4.3 at time  $T_{C_k}$  with no market impact.* For the market makers with relatively small risk tolerance, this assumption is reasonable.

## 4.2 Implementation of the Hedging Strategy

The hedging strategy introduced in the previous section 4.1 is implemented in matlab. To carry out the test, we apply this strategy to the client trading process obtained from the simulation exercise conducted in section 3.5. Figure 4.1 shows the market maker's Risk Adjusted Wealth Processes  $AW_1(T_{C_k})$  and  $AW_2(T_{C_k})$  when we assume  $U_p = 4,000,000$ ,  $L_p = 2,000,000$ ,  $U_N = -4,000,000$ , and  $L_N = -4,000,000$ . This

means that whenever the client trade leads the wealth process to go beyond  $\pm 4$  million USD, the hedging strategy will issue a hedging trade to bring its position back to  $\pm 2$  million USD. We can see that the Risk Adjusted Base Currency Wealth Process  $AW_1(T_{C_k})$  is bounded between  $\pm 4$  million USD.

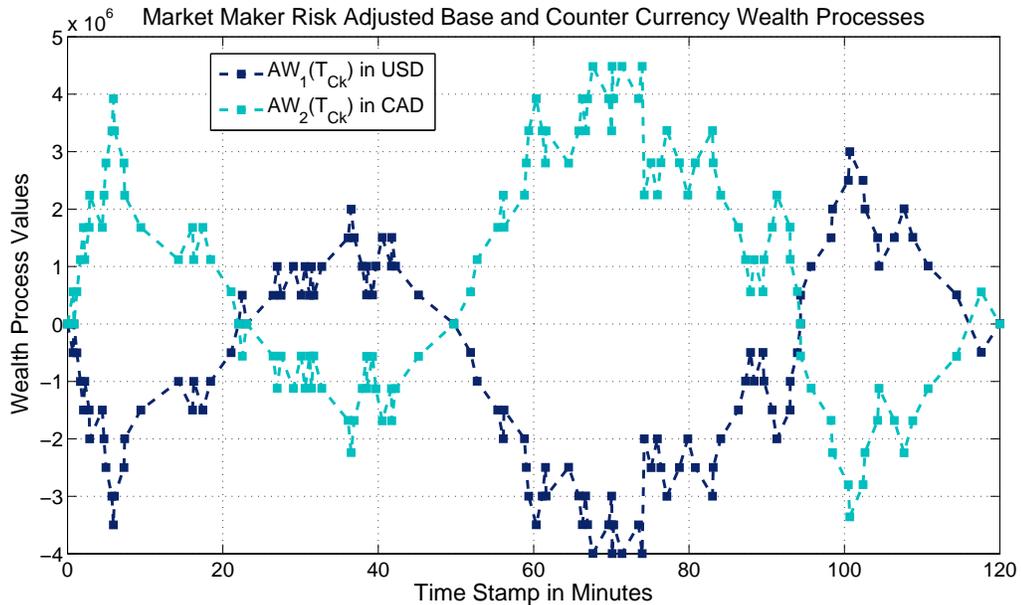


Figure 4.1: Sample Paths of Market Maker’s  $AW_1(T_{C_k})$  and  $AW_2(T_{C_k})$

Figure 4.2 shows the comparison between the P&L with and without risk hedging. At about 80<sup>th</sup> minute, we start to see discrepancies between the P&Ls. This is because that there is no risk hedging trade happening before that time. This plot also shows that our current risk hedging strategy does not necessarily produce better P&L. But at least this gives us something to start with.

### 4.3 Scenario Analysis

After a risk hedging strategy is introduced, the first question would be “*What impact could it have on the market maker’s P&L?*” In order to answer this question, we will

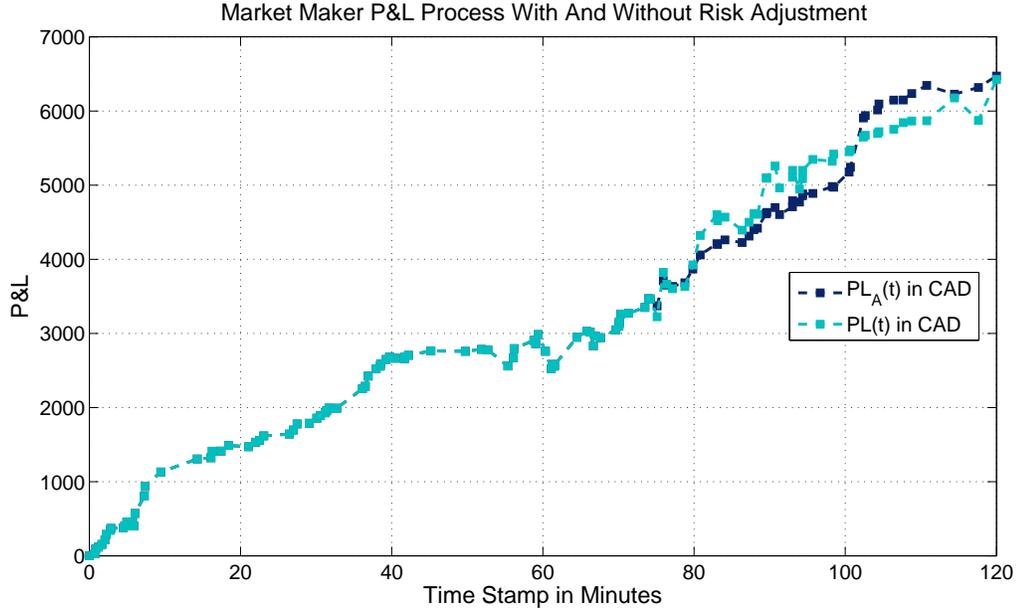


Figure 4.2: Sample Path of Market Maker's  $P\&L(t)$  and  $PL_A(t)$

run the simulation exercise for the market data process and client trading process under three different scenarios as follows:

1. Balanced client buying and selling flows under flat market condition.
2. Intensive client selling flow under downward market condition.
3. Intensive client buying flow under upward market condition.

Scenarios 2 and 3 are more like stress tests for our strategy. For each scenario, we calculate P&L values (equations (3.14) and (4.6)) before and after risk hedging for each sample path, and compare their difference. Sharp ratios are also calculated to compare the returns before and after risk hedging. It is a measure of the excess return per unit of risk in an investment asset or a trading strategy. In [28], it is defined as

$$S = \frac{E[R - R_f]}{\sqrt{\text{VAR}[R - R_f]}}, \quad (4.8)$$

where  $R$  and  $R_f$  are the asset return and risk-free return respectively. In our situation, we let  $R_f$  equal to 0 since we are looking at absolute return.

### 4.3.1 Balanced Client Buying and Selling Flows under Flat Market Condition

In this subsection, we assume that we are under a flat market condition. Then, given a pre-defined client flow, which contains balanced client buying and selling trades, we compute and compare the P&L with and without risk hedging on 5,000 different market data paths. For a five-hour trading horizon, we use the following parameter assumptions for the pre-defined client trading flow:

1. For client buying trades, we assume Poisson arrival process with rate  $\lambda_{N_1} = 1/120$ ,  $\mu_{X_1} = 500K$ , and  $\sigma_{X_1}^2 = 500K$ .
2. For client selling trades, we assume Poisson arrival process with rate  $\lambda_{N_2} = 1/120$ ,  $\mu_{X_2} = 500K$ , and  $\sigma_{X_2}^2 = 500K$ .
3. Client margin  $\delta_C = 0.5$  pip.

For each of the 5,000 sample paths of market (mid-price) data process, we assume Poisson arrival process with rate  $\lambda_M = 1/2$ . Its value follows a geometric Brownian motion given by equation (3.7) with initial value  $P_{mid}(0) = 1.1212$ , drift  $\alpha = 0$  pip per minute, and volatility  $\sigma^2 = 0.5$  pip per minute. Market bid-ask spread remains at  $\delta = 1$  pip. For risk hedging strategy, we set the risk barriers values of  $U_P = 4M$ ,  $L_P = 1M$ ,  $U_N = -4M$ , and  $L_N = -1M$ .

Figures 4.3 and 4.4 show the simulation results of market maker's P&L without and with risk hedging respectively. Figure 4.5 shows the difference between them, and Figure 4.6 shows the comparison of sharp ratio values before and after risk hedging.

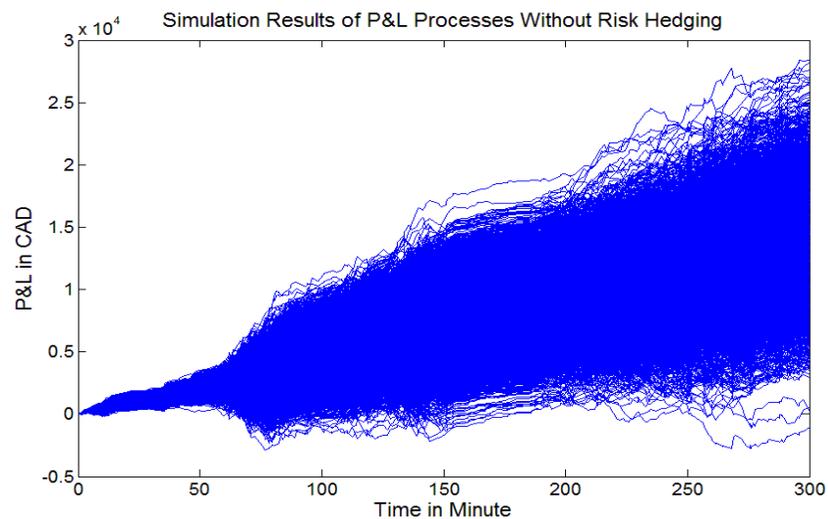


Figure 4.3: P&L Without Risk Hedging in Flat Market with Balanced Client Flow

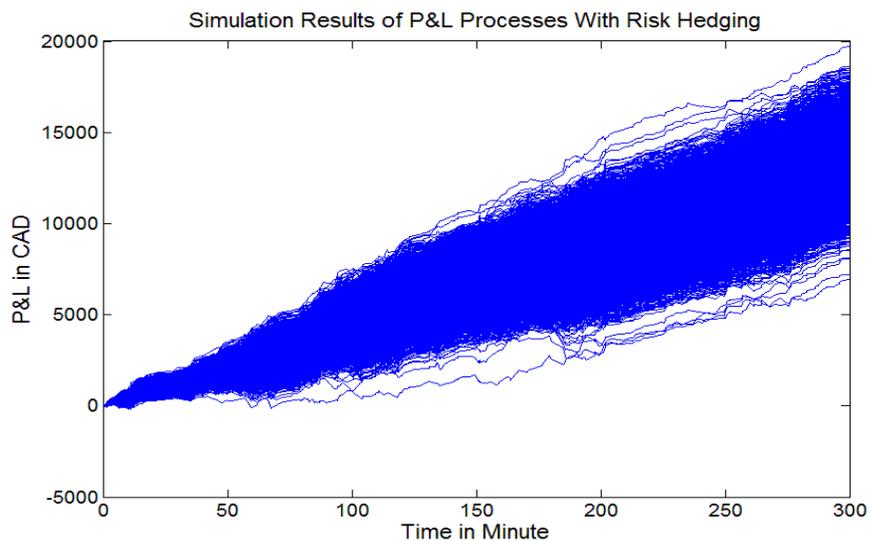


Figure 4.4: P&L With Risk Hedging in Flat Market with Balanced Client Flow

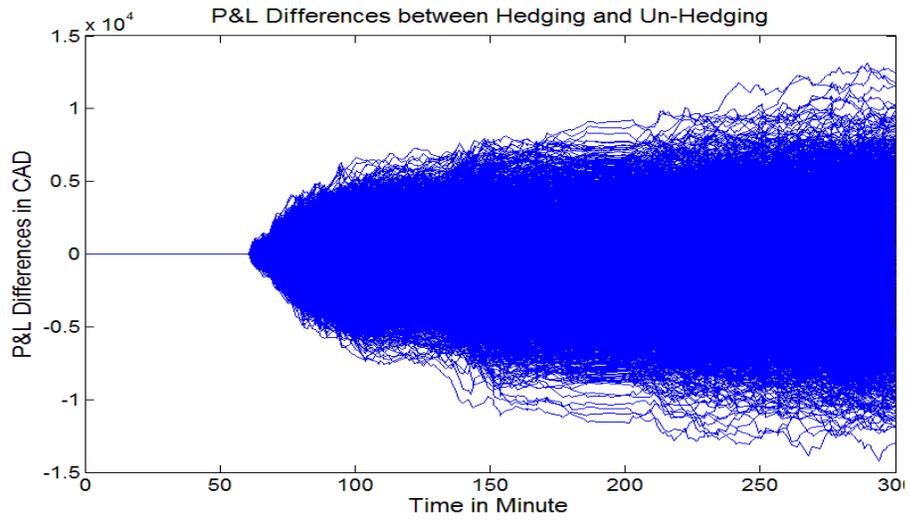


Figure 4.5: P&L Differences in Flat Market with Balanced Client Flow

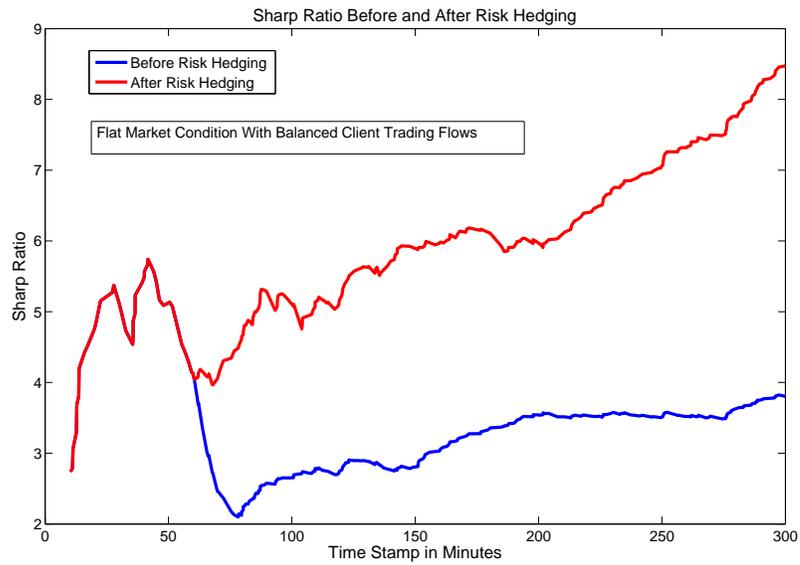


Figure 4.6: Sharp Ratio Comparison in Flat Market with Balanced Client Flow

From inspecting Figure 4.5 alone, it is difficult to tell that using an active risk hedging strategy generates more revenue for the market maker than not hedging. For the 5,000 sample paths, we see that about 50% of the time that the risk hedging strategy generates less revenue than the un-hedging strategy. The amounts of out-performance and under-performance almost cancel each other. But if we compare figure 4.3 and figure 4.4, we can see that the P&L process without the risk hedging strategy has a much wider range than the P&L process with the risk hedging strategy does during the simulation horizon. At the end of the period, the P&L process without the risk hedging strategy has its range from  $-5,000\text{CAD}$  to  $28,000\text{CAD}$ , while the P&L process with the risk hedging strategy has its range from  $5,000\text{CAD}$  to  $19,000\text{CAD}$ . Figure 4.6 shows that the sharp ratio of risk-hedging P&L starts perform better once the first risk-hedging trade is initiated. The better sharp ratio values are mainly resulted from return variance reduction. This tells us that being active in risk hedging is not necessarily more superiors than being in non-risk hedging, but it substantially reduce the probability of getting very low (or even negative) P&L in trading with the clients. This makes sense because more risk taking imposes more chances in both winning in a big way and losing in a big way. The risk hedging strategy is helpful if the market maker is seeking stability in his/her revenue generating.

### **4.3.2 Intensive Client Selling Flow under Downward Market Condition.**

In this subsection, we assume that we are facing a downward market condition and experiencing intensive client selling trades during the trading horizon. P&Ls are calculated on 5,000 simulation paths for risk hedging and un-hedging. For the five-hour trading horizon, we use the following parameter assumptions for the predefined client trading flow:

1. For client buying trades, we assume a Poisson arrival process with the rate of

$\lambda_{N_1} = 1/120$ ,  $\mu_{X_1} = 500K$ , and  $\sigma_{X_1}^2 = 500K$ .

2. For client selling trades, we assume a Poisson arrival process with a rate of  $\lambda_{N_2} = 1/60$ ,  $\mu_{X_2} = 1M$ , and  $\sigma_{X_2}^2 = 1M$ .
3. Client margin  $\delta_C = 0.5$  pip.

For each of the 5,000 sample paths of market (mid-price) data process, we assume a Poisson arrival process with a rate of  $\lambda_M = 1/2$ . Its value is assumed to follow a geometric Brownian motion given by equation (3.7) with initial value  $P_{mid}(0) = 1.1212$ , drift  $\alpha = -0.5$  pip per minute (negative drift means downward market), and volatility  $\sigma^2 = 1$  pip per minute. Market bid-ask spread remains at  $\delta = 1$  pip. For the risk hedging strategy, we set the risk barriers values of  $U_P = 4M$ ,  $L_P = 1M$ ,  $U_N = -4M$ , and  $L_N = -1M$ .

Figure 4.7 shows the market maker's P&L results when the risk hedging strategy is not applied. Since this is in a downward trend market, and client selling trades are arriving with large amounts at twice the speed of buying trades, the market maker is loosing a large amount of money. Figure 4.8 shows the simulation results of market maker's P&L with risk hedging, and Figure 4.9 shows the difference of P&L between hedging and un-hedging. We can see that by imposing a hedging strategy, we substantially reduce the probability of loosing money. Nearly half of the sample paths end up in the positive region of P&L. Even if the P&L is negative, it is much less negative than the P&L without a risk hedging strategy. By figure 4.10, we see that the sharp ratio of risk-hedging P&L is contained at a certain level around  $-1$ , while the sharp ratio of non-risk hedging P&L is with a big downward slope.

### 4.3.3 Intensive Client Buying Flow under Upward Market Condition.

In this subsection, we assume that we face an upward market condition and experience intensive client buying trades during the trading horizon. This represents the opposite

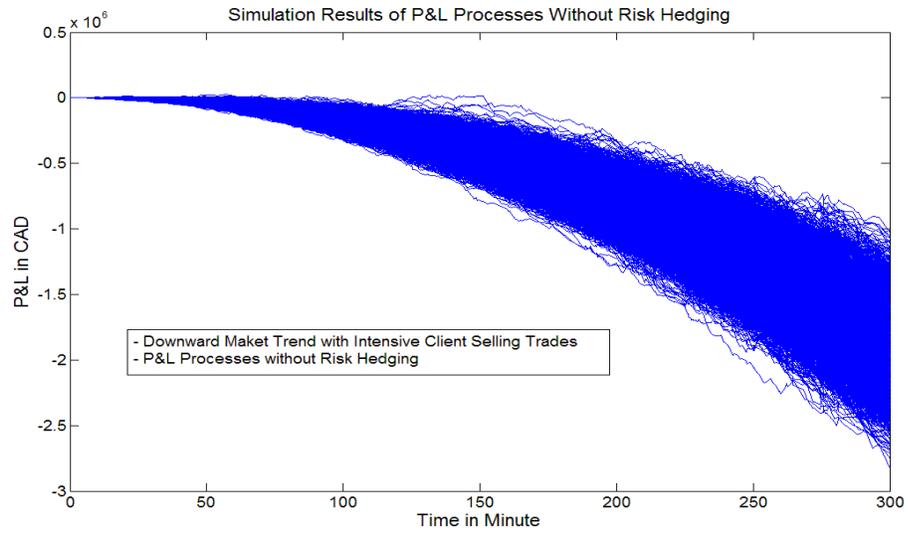


Figure 4.7: P&L Without Risk Hedging in Downward Trend Market with Intensive Client Sell

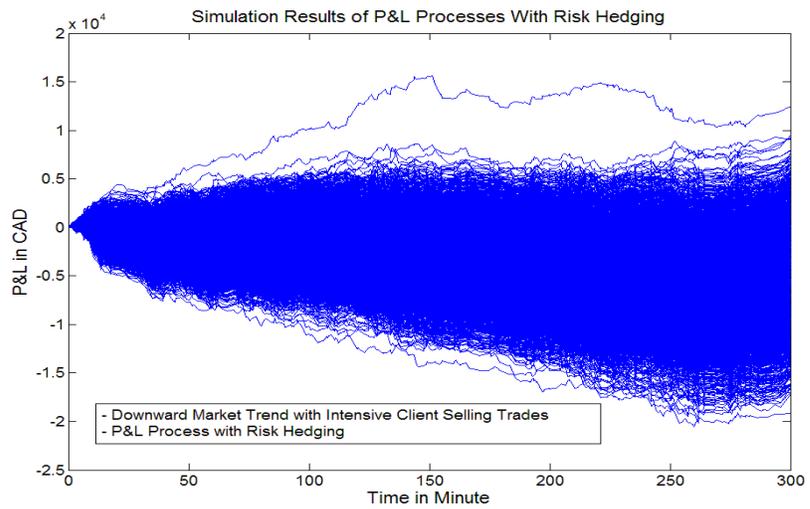


Figure 4.8: P&L With Risk Hedging in Downward Trend Market with Intensive Client Sell

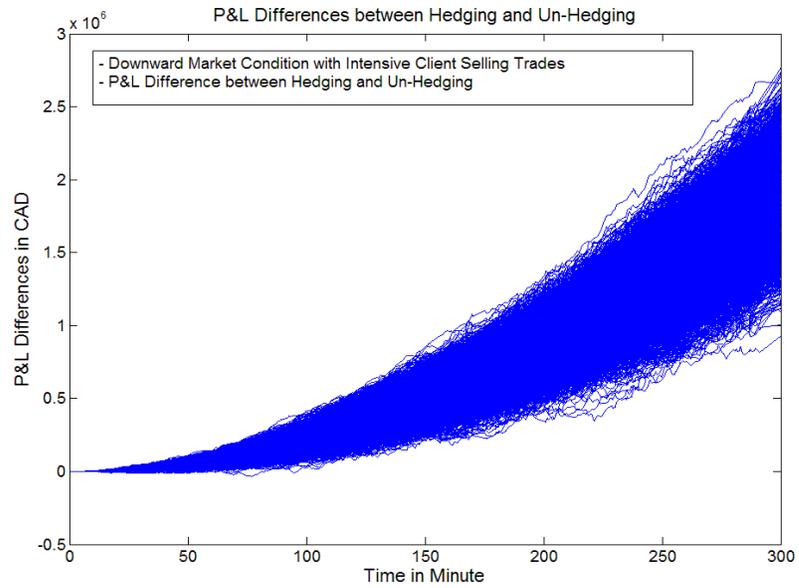


Figure 4.9: P&L Differences in Downward Trend Market with Intensive Client Sell

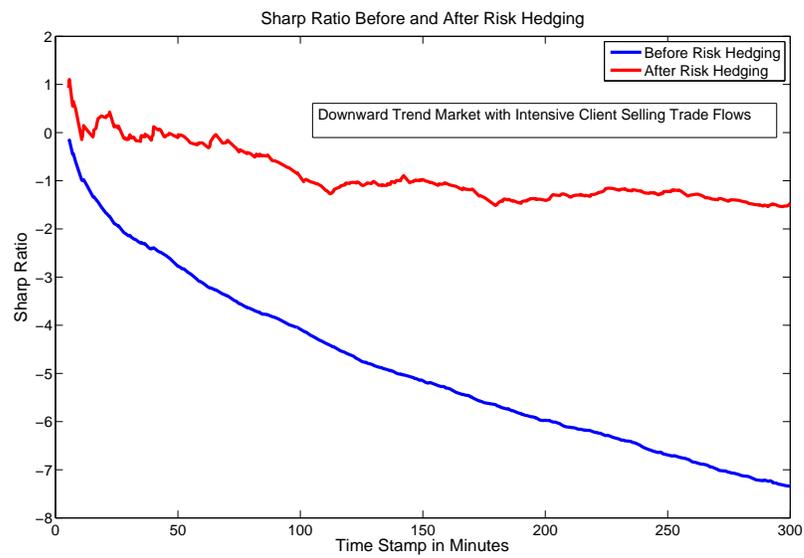


Figure 4.10: Sharp Ratio Comparison in Downward Trend Market with Intensive Client Sell

scenario to the scenario used in the pervious section. P&Ls are calculated on 5,000 simulation paths for risk hedging and un-hedging strategies. For the five-hour trading horizon, we use the following parameter assumptions for the predefined client trading flow:

1. For client buying trades, we assume a Poisson arrival process with a rate of  $\lambda_{N_1} = 1/90$ ,  $\mu_{X_1} = 1.5M$ , and  $\sigma_{X_1}^2 = 1M$ .
2. For client selling trades, we assume a Poisson arrival process with a rate of  $\lambda_{N_2} = 1/120$ ,  $\mu_{X_2} = 500K$ , and  $\sigma_{X_2}^2 = 500K$ .
3. Client margin  $\delta_C = 0.5$  pip.

For each of the 5,000 sample paths of market (mid-price) data process, we assume a Poisson arrival process with a rate of  $\lambda_M = 1/2$ . Its value is assumed to follow a geometric Brownian motion given by equation (3.7) with initial value  $P_{mid}(0) = 1.1212$ , drift  $\alpha = 0.3$  pip per minute (positive drift means upward market), and volatility  $\sigma^2 = 1$  pip per minute. The market bid-ask spread is set to remain at  $\delta = 1$  pip. For the risk hedging strategy, we set the risk barriers values of  $U_P = 4M$ ,  $L_P = 1M$ ,  $U_N = -4M$ , and  $L_N = -1M$ .

There is no surprise that we obtain similar results as the ones obtained in the previous section. Figures 4.11 and 4.12 show the simulation results of market maker's P&L without and with a risk hedging strategy respectively, and Figure 4.13 shows the difference between them. We can see that without any risk hedging strategy, the market maker suffers huge losses when the upward market rally happens with intensive client buying trades. Once the risk hedging strategy is imposed, the probability of getting negative P&L is substantially reduced. P&L differences between the risk hedging and un-hedging strategies are positive for all sample paths. Figure 4.14 shows that the risk-hedging P&L achieved positive sharp ratios while non-risk hedging P&L has a negative and decreasing sequence of sharp ratios.

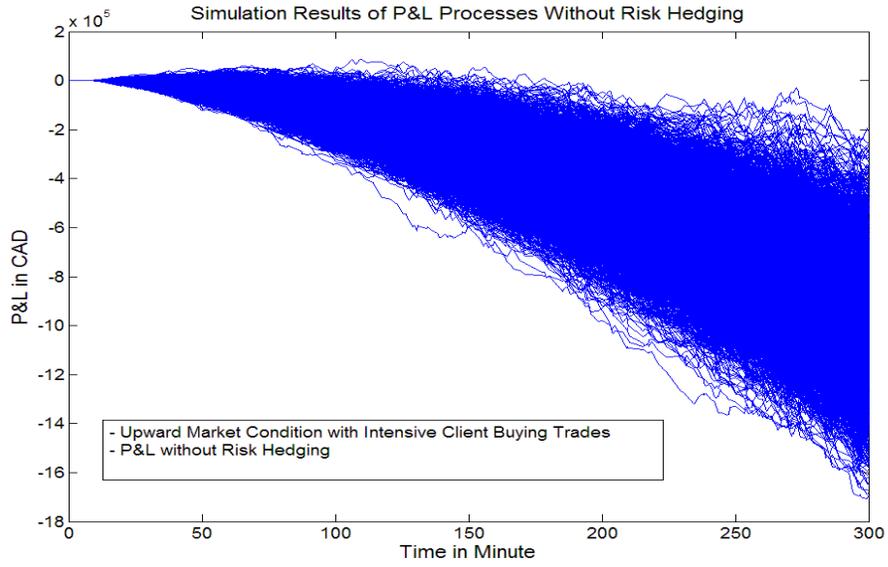


Figure 4.11: P&L Without Risk Hedging in Upward Trend Market with Intensive Client Buy

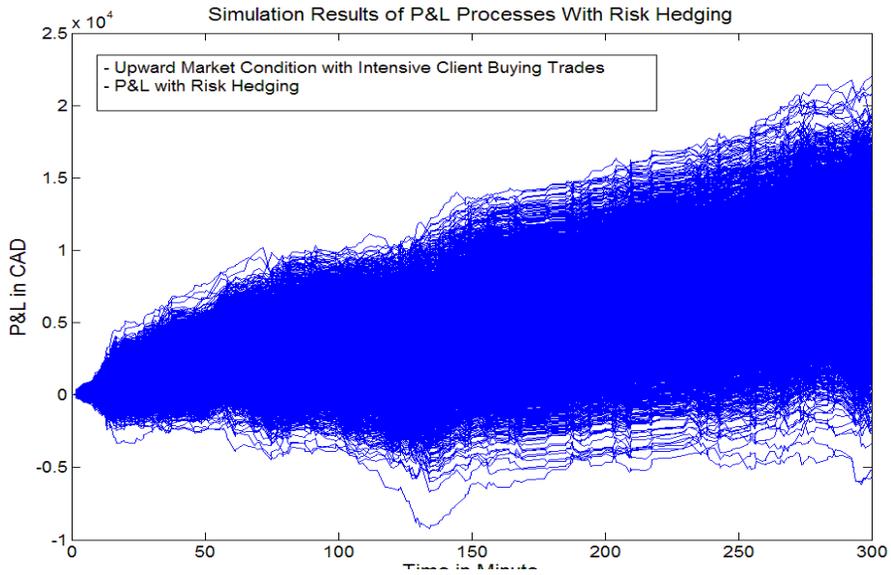


Figure 4.12: P&L With Risk Hedging in Upward Trend Market with Intensive Client Buy

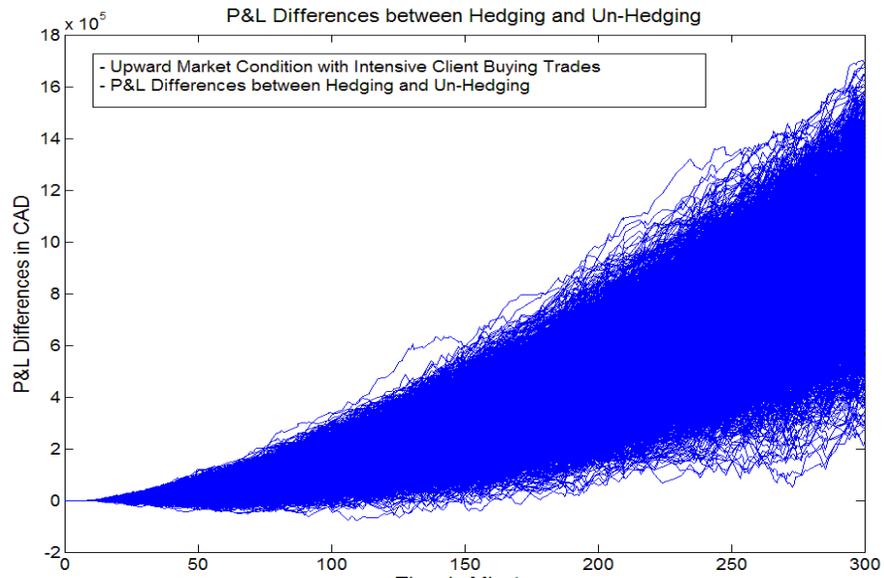


Figure 4.13: P&L Differences in Upward Trend Market with Intensive Client Buy

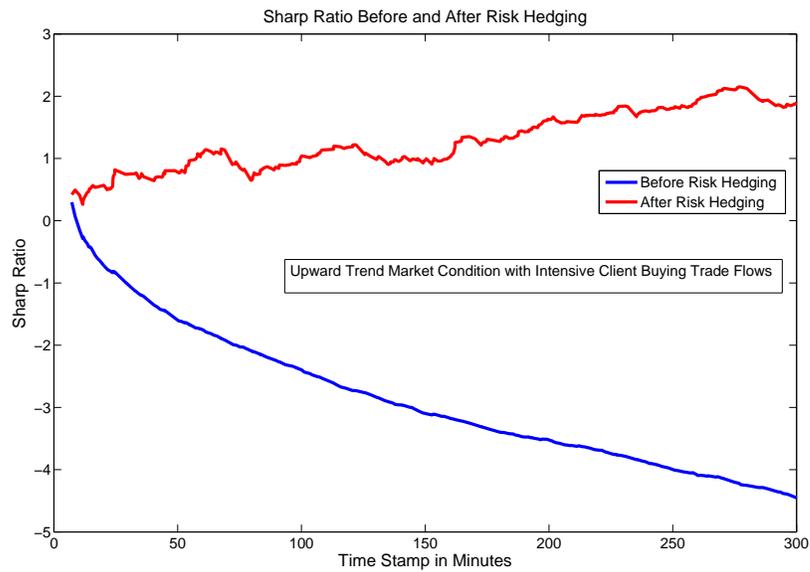


Figure 4.14: Sharp Ratio Comparison in Upward Trend Market with Intensive Client Buy

# Chapter 5

## Tail Risk Analysis

### 5.1 Overview of Extreme Value Theory

The last years of 2007, 2008, and 2009 have seen the greatest financial crisis since the Great Depression 1929. This has led to numerous criticisms about the existing risk management systems and motivated the search for more appropriate methodologies able to cope with rare events that have heavy consequences. The typical question one would like to answer is: “If things go wrong, how wrong can they go?” The problem is then to model the rare phenomena that lie outside the range of available observations. Extreme value theory (EVT) provides a framework to formalize the study of behavior in the tails of a distribution. Critical questions relating to the probability of a market crash or boom require an understanding of the statistical behavior expected in the tails. EVT allows us to use extreme observations to measure the density in the tail. This measure can be extrapolated to parts of the distribution that have yet to be observed in the empirical data. It can also be mapped onto distributions with specific tail behavior. In this way we can simulate a theoretical process that captures the extreme features of the empirical data and improve the accuracy of estimated probabilities of extraordinary market movements. Extreme Value Theory has been well established in many fields such as insurance and engineering. Text books [25]

and [7] gives introduction of Extreme Value Theory from its basic foundation to its applications in insurance and finance industries.

This chapter is composed by two parts. In the first part, we draw down the foundation work of Extreme Value Theory, and introduce the empirical estimation methods for its *shape*, *location*, and *scale* parameters. A dataset of EUR/USD exchange rates with five-minute frequency will be applied for the empirical estimation. In the second part of this chapter, we will introduce the concept of Value-at-Risk (VaR), which is the most popular and important quantity for risk management. We will discuss an approach to VaR calculation using the Extreme Value Theory.

Firstly, let's introduce the basics of Extreme Value Theory. Assume that a series of random variables  $X_k$  for  $k = 1, 2, 3, \dots, n$  are independent and identically distributed with a common cumulative distribution function (c.d.f)  $F(x)$ . The range of random variables  $X_k$  for  $k = 1, 2, 3, \dots, n$  is  $[l, u]$ . For  $X_k$  being log returns, we have  $l = -\infty$  and  $u = \infty$ . Let  $M_n = \max\{X_1, X_2, \dots, X_n\}$  be the maximum of the random sample of size  $n$ . Then, the c.d.f. of  $M_n$  is given by

$$\begin{aligned}
 F_n(x) &= Pr(M_n \leq x) \\
 &= \prod_{j=1}^n Pr(X_j \leq x) \\
 &= [F(x)]^n \\
 &= F^n(x).
 \end{aligned} \tag{5.1}$$

In practice, the c.d.f.  $F(x)$  is unknown and, hence the c.d.f.  $F^n(x)$  of  $M_n$  is unknown. However, as  $n \rightarrow \infty$ ,  $F^n(x) \rightarrow 0$  if  $x < u$  and  $F^n(x) \rightarrow 1$  if  $x \geq u$ , where  $u$  is the upper boundary of the range. Therefore, the limiting distribution of  $F^n(x)$  is degenerate. To deal with this, we need to normalize  $F^n(x)$ . Suppose there exists a sequence of constants  $a_n > 0$  and  $b_n \in \mathbf{R}$  such that:

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) = F^n(a_n z + b_n), \tag{5.2}$$

then the limiting distribution of  $F^n(a_n z + b_n)$  is given by

$$\lim_{n \rightarrow \infty} F^n(a_n z + b_n) = G(z). \quad (5.3)$$

Finding the limiting distribution  $G(z)$  is called the Extremal Limit Problem. Finding the  $F(x)$  that have sequences of constants as described above leading to  $G(z)$  is called the Domain of Attraction Problem. In articles [10] and [24], authors gave the limiting law for the maxima  $M_n$  with  $n$  being the sample size. The theorem is as follows: *let  $M_n^i$  be a sequence of i.i.d. random variables for  $i = 1, 2, \dots$ . If there exist constants  $a_n > 0$ ,  $b_n \in \mathbf{R}$ , and some non-degenerate distribution function  $G$  such that*

$$Z^i = \frac{M_n^i - b_n}{a_n} \xrightarrow{d} G, \quad (5.4)$$

then  $G$  belongs to one of the three standard extreme value distributions:

$$\text{Frechet : } \Phi_\alpha(z) = \begin{cases} 0 & \text{for } z \leq 0 \text{ and } \alpha > 0 \\ e^{-z^{-\alpha}} & \text{for } z > 0 \text{ and } \alpha > 0 \end{cases} \quad (5.5)$$

$$\text{Weibull : } \Psi_\alpha(z) = \begin{cases} e^{-(-z)^\alpha} & \text{for } z \leq 0 \text{ and } \alpha > 0 \\ 1 & \text{for } z > 0 \text{ and } \alpha > 0 \end{cases} \quad (5.6)$$

$$\text{Gumbel : } \Lambda(z) = e^{-e^{-z}} \text{ for } z \in \mathbf{R} \quad (5.7)$$

where parameter  $\alpha > 0$  is the shape parameter, which captures the weights of the tail in the distribution of the parent random variable  $X$ . This theorem is known as the **Fisher-Tippet Theorem**. Constants  $a_n > 0$  and  $b_n \in \mathbf{R}$  are referred to as *scale* parameter and *location* parameter respectively.

Intuitively, these three standard extreme value distributions represent three possibilities for the decay of the density function in the tail. *Frechet* distribution represents tails that decay by a power, as in the cases of the stable Paretian, Cauchy and Student t distributions. They are no longer integrable when weighted by the tail probabilities, hence leading to "fat tails". *Weibull* represents tails that can decay with a finite tail

index; this will be a thin tailed distribution with a finite upper endpoint. *Gumbel* represents tails that can decay exponentially with all finite moments; these are standard cases of Normal, Lognormal, Gamma, etc. Figure 5.1 shows the shapes of the probability density functions for standard *Frechet*, *Weibull*, and *Gumbel* distributions when shape parameter  $\alpha = 1.5$ .

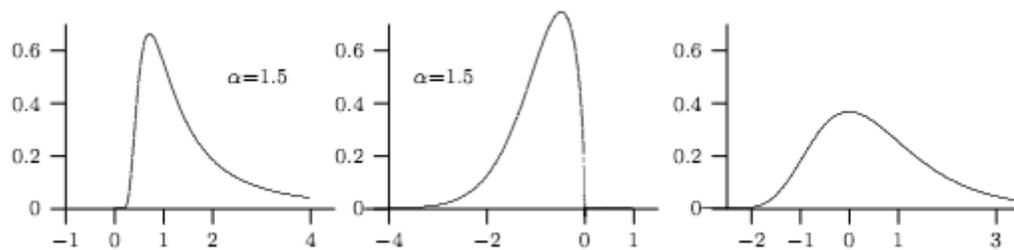


Figure 5.1: Density Functions for Frechet, Weibull, and Gumbel when  $\alpha = 1.5$

In article [13], Jenkinson and Von Mises suggested an one-parameter representation given by

$$G_{\xi}(z) = \begin{cases} e^{-(1+\xi z)^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ e^{-e^{-z}} & \text{if } \xi = 0 \end{cases} \quad (5.8)$$

for these three standard distributions, with  $x$  such that  $1+\xi x > 0$ . This generalization is known as the *Generalized Extreme Value (GEV)* distribution, and is obtained by setting  $\xi = \alpha^{-1}$  for *Frechet* distribution,  $\xi = -\alpha^{-1}$  for *Weibull* distribution, and by interpreting the *Gumbel* distribution as the limiting case for  $\xi = 0$ . We can obtain the probability density function (p.d.f.) of the Generalized Extreme Value distribution

by differentiating the above c.d.f. (5.8), which gives

$$g_\xi(z) = \begin{cases} (1 + \xi z)^{-\left(\frac{1}{\xi}+1\right)} e^{-(1+\xi z)^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ e^{x-e^{-z}} & \text{if } \xi = 0. \end{cases} \quad (5.9)$$

This generalized representation is very useful when maximum likelihood estimates have to be computed when we do not know the type of limiting distribution of the sample maxima in advance.

## 5.2 Maximum Likelihood Methods for EVT

From the previous section, we know that the Generalized Extreme Value distribution contains three parameters  $\xi$ ,  $a_n > 0$ , and  $b_n \in \mathbf{R}$ , which are referred to as *shape*, *scale*, and *location parameter* respectively. In this section, we explain how to use maximum likelihood estimate method in estimating the three parameters.

For one given sample, there is only one minimum value or maximum value observed. We can not estimate these three parameters with only one extreme observation. An alternative approach that has been applied is to divide the sample into non-overlapping sub-samples and apply the extreme value theory to the sub-samples. This approach has been applied in literatures such as [11] and [24]. For a sample of size  $T$ , we divide the sample into  $k$  non-overlapping sub-samples each with  $n$  observations, assuming for simplicity that  $T = nk$ . That is, we divide the sample

$$\{x_1, x_2, \dots, x_T\}$$

into

$$\{x_1, \dots, x_n, |x_{n+1}, \dots, x_{2n}, | \dots |x_{(k-1)n+1}, \dots, x_{nk}\}.$$

We can write the observation value as  $x_{in+j}$ , where  $1 \leq j \leq n$  and  $i = 0, 1, \dots, k-1$ . This tells us that observation  $x_{in+j}$  is the  $j^{\text{th}}$  observation of  $i^{\text{th}}$  sub-sample. When the size of each sub-sample is sufficiently large, we hope that the Extreme Value Theory

will apply to each sub-sample. According to [14] and [24], the choice of  $n$  is subjective to the practical application. For example, for daily stock returns,  $n = 21$  corresponds approximately to the number of trading days in a month, and  $n = 63$  is the number of trading days in a quarter.

Let's define

$$m_n^i = \max\{x_{(i-1)n+1}, x_{(i-1)n+2}, \dots, x_{(i-1)n+n}\}, \quad \text{for } i = 1, \dots, k$$

being the maximum (or  $n^{\text{th}}$  order statistics) of the  $i^{\text{th}}$  sub-sample, where  $n$  stands for the sub-sample size. When  $n$  is sufficiently large,  $z_i = \frac{m_n^i - b_n}{a_n}$  should follow an extreme value distribution, and the collection of sub-sample maximum values  $\{m_n^i | i = 1, 2, \dots, k\}$  can be regarded as a sample of  $k$  observations from that extreme value distribution. This collection of sub-sample maximums is the data set that we will use to estimate the unknown parameter values of the extreme value distribution.

Note that the c.d.f. and p.d.f. functions 5.8 and 5.9 are of normalized maximum  $z_i = \frac{m_n^i - b_n}{a_n}$ . To obtain the p.d.f. of  $m_n^i$ , we simply apply change of variable theory and obtain

$$g_\xi(m_n^i) = \begin{cases} \frac{1}{a_n} (1 + \xi \frac{m_n^i - b_n}{a_n})^{-(\frac{1}{\xi} + 1)} e^{-(1 + \xi \frac{m_n^i - b_n}{a_n})^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ \frac{1}{a_n} e^{\left[\frac{m_n^i - b_n}{a_n} - e^{-\frac{m_n^i - b_n}{a_n}}\right]} & \text{if } \xi = 0. \end{cases} \quad (5.10)$$

And the likelihood function of the sub-sample maximum values can be obtained as

$$L(m_n^1, m_n^2, \dots, m_n^k | \xi, a_n, b_n) = \prod_{i=1}^k g_\xi(m_n^i). \quad (5.11)$$

The log-likelihood function for  $\xi \neq 0$  is

$$\begin{aligned} l(\xi, a_n, b_n) &= \log \left( \frac{1}{a_n} \right)^k \\ &\quad - \left( 1 + \frac{1}{\xi} \right) \sum_{i=1}^k \log \left( 1 + \xi \frac{m_n^i - b_n}{a_n} \right) \\ &\quad - \sum_{i=1}^k \left( 1 + \xi \frac{m_n^i - b_n}{a_n} \right)^{-\frac{1}{\xi}}. \end{aligned} \quad (5.12)$$

The log-likelihood function for  $\xi = 0$  is

$$l(a_n, b_n) = \log \left( \frac{1}{a_n} \right)^k + \sum_{i=1}^k \left( \frac{m_n^i - b_n}{a_n} - e^{-\frac{m_n^i - b_n}{a_n}} \right) \quad (5.13)$$

We see that the MLE estimates will depend on the number of blocks  $k$  and the number of observations  $n$  in each block. According to [25], there is a trade-off between the bias and variance of the estimates. The bias of the MLE is reduced by increasing the block size  $n$ , and the variance of the MLE is reduced by increasing the number of blocks  $k$ . Nonlinear estimation procedure can then be applied to obtain the maximum likelihood estimates of parameters  $\xi$ ,  $a_n$ , and  $b_n$ .

### 5.3 Empirical Analysis on EVT

The data sample we are about to analyze is the daily EUR/USD exchange rate from 5/01/1998 to 06/30/2006. There are in total of 2,058 days. Each data point is marked with a unique date stamp. Before we fit the generalized extreme value distribution by maximum likelihood method, let's conduct some preliminary data analysis. In financial practice, many investors are worried about the investment loss, so let's calculate the daily loss (or negative return) of EUR/USD as follows

$$x_i = - \left( \frac{p_i - p_{i-1}}{p_{i-1}} \right) 100\%, \quad (5.14)$$

where  $p_i$  stands for the day-end price of day  $i$  for  $i = 2, 3, \dots, 2058$ . In our sample, the day-end price is marked with time stamp 23:55 for each day stamp. Figure 5.2 shows the day end-price of EUR/USD for our sample period. Figure 5.3 is the QQ plot of the EUR/USD daily loss observations  $x_i$  for  $i = 2, \dots, 2058$ . The plot suggests that loss observations are having a thicker tail than normal distribution. It suggests a Weibull family of generalized extreme value distribution with  $\xi < 0$  for the sub-sample maximum loss.

Next, let's find the maximum daily loss for each monthly period. In order to make sure that there are  $n = 20$  (number of trading days) observations in each month, we

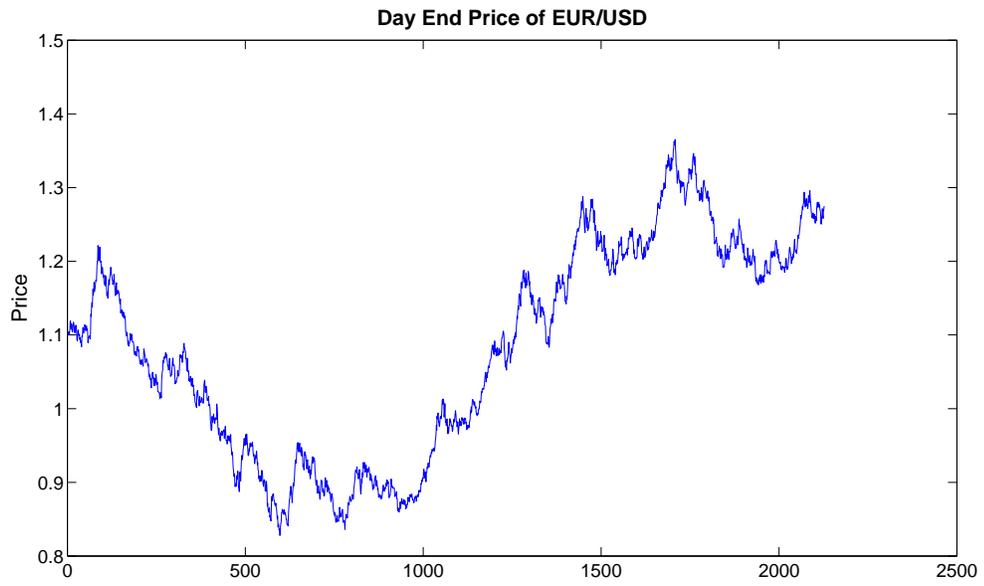


Figure 5.2: EUR/USD Day-End Price

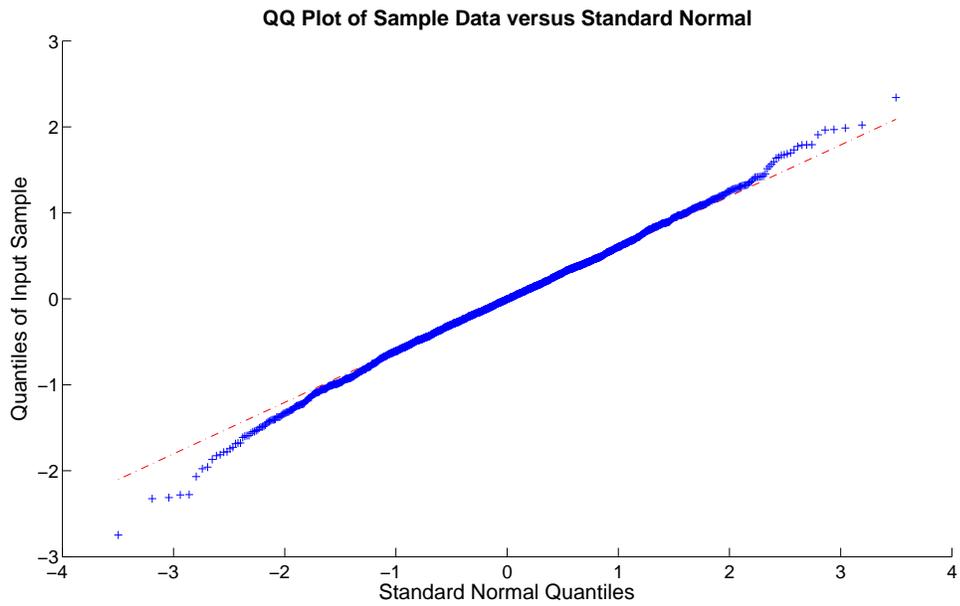


Figure 5.3: Q-Q Plot of EUR/USD Daily Loss

will count the first one or two days of the next month into this month if there is only a number of 18 or 19 days available for this month. We will also use the first 20 days data if the month has more than 20 trading days. This gives us in total of  $k = 98$  observations. Figure 5.4 shows maximum daily loss in each monthly block. We see that the largest daily loss in a monthly block is 2.34% in 2000 September.

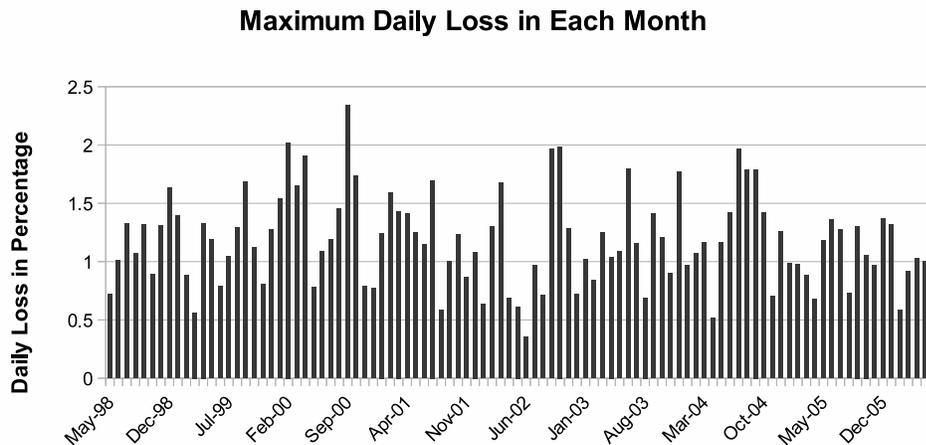


Figure 5.4: EUR/USD Maximum Daily Loss for Each Month

Then, we can calculate the Maximum Likelihood Estimates for the parameters of the generalized extreme value distribution by using these 98 monthly block maxima daily loss on EUR/USD. By inputting the data into MATLAB, we get the estimates of shape parameter  $\hat{\xi} = -0.132$ , scale parameter  $\hat{a}_n = 0.3513$ , and location parameter  $\hat{b}_n = 1.0108$ . Since the shape parameter  $\hat{\xi} < 0$ , the monthly maximum daily loss follows a Weibull distribution. Figure 5.5 compares the empirical cdf for the sample and the theoretical cdf of the generalized extreme distribution with estimated parameters.

Now We can apply the same procedures with different combinations of  $n$  and  $k$ . Thus, we can see the effects of number of observations on MLE estimates. Table 5.1

shows the MLE estimates for the generalized extreme value distribution of the daily maximum loss with different block sizes. We can see that shape and location parameters  $\xi$  and  $b_n$  are quite sensitive to the number of observations in each block, whereas the scale parameter  $a_n$  is less sensitive compare to the other two.

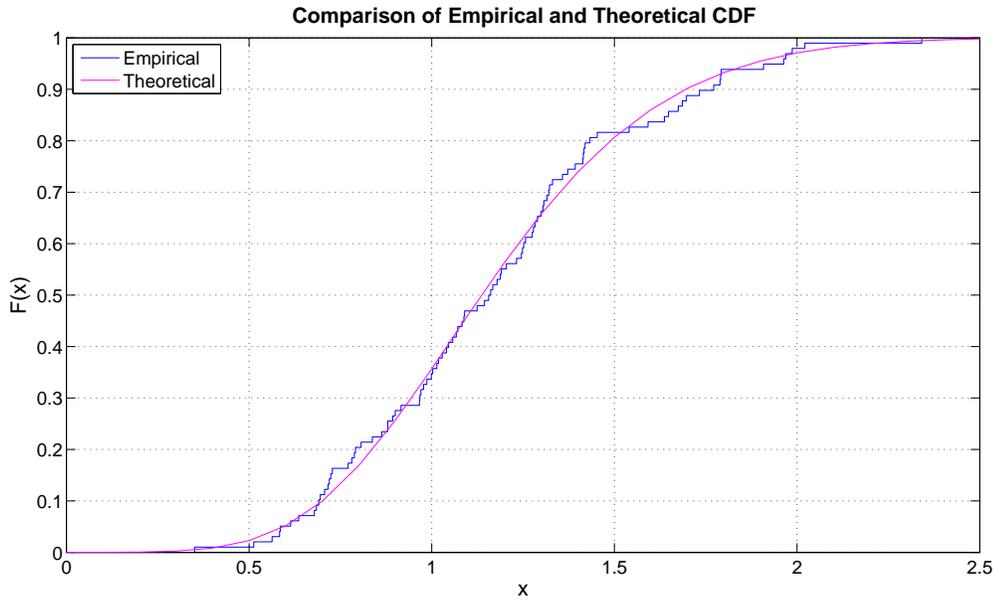


Figure 5.5: Empirical and Theoretical CDF Comparison for Maximum Daily Loss in Monthly Block

Table 5.1: MLE Estimates for Generalized Extreme Value Distribution with Different Block Size

Frequency	# of Obs	# of Blocks	$\hat{\xi}$	$\hat{a}_n$	$\hat{b}_n$
Bi-Weekly	$n = 10$	$k = 196$	-0.1323	0.3689	0.7856
Monthly	$n = 20$	$k = 98$	-0.1320	0.3513	1.0108
Quarterly	$n = 60$	$k = 32$	-0.0039	0.2643	1.3087
Semi-Annually	$n = 120$	$k = 16$	-0.1257	0.2647	1.5467
Annually	$n = 240$	$k = 8$	-0.3158	0.2767	1.7597

## 5.4 Value-at-Risk (VaR)

Value-at-Risk (VaR) is a widely used risk measure in today's financial industry. It is an attempt to provide a single number to summarize the total risk in a portfolio of financial assets. It is an accepted methodology used by corporate treasurers, fund managers, and financial institutions. Central banks regulators also use VaR in determining the capital a bank is required to keep to reflect the market risk it is bearing.

The definitions and concepts of VaR can be found in many books and literatures such as [12] and [4]. In [4], VaR is defined as the maximum loss which can occur with  $X\%$  confidence over a holding period of  $n$  days for a portfolio. Thus, when using VaR as a measure of risk, we are interested in making a statement of the following form: *“We are  $X$  percent certain that we will not lose more than  $V$  dollars in the next  $N$  days for a certain portfolio.”* The variable  $V$  is the VaR of the portfolio. For example, if a daily VaR is stated as \$100,000 with a 95% level of confidence for a portfolio, it means that we are 95% confident that the portfolio will not lose more than \$100,000 during a day. We can see that VaR is a very important risk measure in helping banks to set up capital requirements for preventing extreme market risk events. According to [12], the Basel Committee on Bank Supervision, (the committee of world's bank regulators), requires VaR to be calculated with  $N = 10$  and  $X = 99$  for the bank's trading book on a daily basis. The capital it requires the bank to hold is the multiplier  $k$  times the VaR measure.  $k$  is chosen on a bank-by-bank basis by the regulators and must be at least 3.0.

Now let's define VaR under a probabilistic framework. Since investors usually would like to think of risk from a loss perspective, we let a continuous random variable  $X$  represent the loss (or negative return) of a financial instrument during a certain period of time  $h$  with c.d.f.  $F_h$ . Thus, a VaR value with  $p$  level of confidence for a period of  $h$  can be defined as the ( $p^{th}$  quantile of the distribution function  $F_h$ ). That

is

$$\text{VaR}_{h,p} = F_h^{-1}(p), \quad (5.15)$$

where  $F^{-1}$  is the inverse function of the distribution function  $F$ . Equivalently, we have

$$P(X > \text{VaR}_{h,p}) = 1 - F_h(\text{VaR}_{h,p}) = 1 - p. \quad (5.16)$$

To quote a valid VaR statement, we must include three components: a time period, a confidence level, and a loss amount. According to [4], calculation of VaR involves the following factors in its practical applications:

1. The confidence level  $p$  such as  $p = 95\%$  or  $p = 99\%$ .
2. The time horizon  $h$ . This might be set by a regulatory committee such that  $h = 1$  day or  $h = 10$  days.
3. The frequency of the data, which might not be the same as the time horizon  $h$ . Daily observations are often used.
4. The c.d.f.  $F_h(x)$  for the return random variable.
5. The amount of the financial position or the mark-to-market value of the portfolio.

Among these factors, it is the c.d.f.  $F_h(x)$  that draws most research attentions. The most commonly used VaR models assume that random variable  $X$ , asset return (or log-return) follows a normal distribution. This assumption itself is a huge risk to those practitioners who use it because reality suggests that returns of most financial products are fat-tailed.

## 5.5 An Extreme Value Approach to VaR

In this section, we discuss an approach to VaR calculation using the Extreme Value Theory. In section 5.2, we derived the maximum likelihood function for estimating

parameter values for a generalized extreme value distribution. By fitting the model with the sample data, we perform an MLE calculation in Matlab and obtain estimates of shape parameter  $\hat{\xi}$ , location parameter  $\hat{b}_n$ , and scale parameter  $\hat{a}_n$  for the generalized extreme value distribution of sub-sample maximum values. If we plug these estimates into the c.d.f. equation 5.8 with  $z = \frac{m_n - b_n}{a_n}$ , we obtain the estimate of the c.d.f. of random variable  $M_n$ , the sub-sample maximum value under the limiting generalized extreme value distribution. It is given as follows:

$$\hat{F}_n(m_n) = \begin{cases} e^{-(1+\hat{\xi}\frac{m_n-\hat{b}_n}{\hat{a}_n})^{-\frac{1}{\hat{\xi}}}} & \text{if } \hat{\xi} \neq 0 \\ e^{-e^{-\frac{m_n-\hat{b}_n}{\hat{a}_n}}} & \text{if } \hat{\xi} = 0. \end{cases} \quad (5.17)$$

Suppose  $m_n^*$  is the  $p^{th}$  quantile of the sub-sample maximum under the limiting generalized extreme value distribution, we can then rewrite equation 5.17 as

$$p = Pr(M_n \leq m_n^*) = \begin{cases} e^{-(1+\hat{\xi}\frac{m_n^*-\hat{b}_n}{\hat{a}_n})^{-\frac{1}{\hat{\xi}}}} & \text{if } \hat{\xi} \neq 0 \\ e^{-e^{-\frac{m_n^*-\hat{b}_n}{\hat{a}_n}}} & \text{if } \hat{\xi} = 0, \end{cases} \quad (5.18)$$

and solve for value  $m_n^*$ . We then obtain the  $p^{th}$  quantile of the sub-sample maximum under the limiting generalized extreme value distribution as

$$m_n^* = \begin{cases} \hat{b}_n - \frac{\hat{a}_n}{\hat{\xi}} [1 - [\log(p)]^{\hat{\xi}}] & \text{if } \hat{\xi} \neq 0 \\ \hat{b}_n - \hat{a}_n [\log[-\log(p)]] & \text{if } \hat{\xi} = 0. \end{cases} \quad (5.19)$$

According to [18] and [4], the case of  $\hat{\xi} \neq 0$  is of major interest in financial applications.

The next step is to make explicit the relationship between sub-sample maxima random variable  $M_n$  and the sub-sample loss (or negative return) random variable  $X_j$  for  $j = 1, 2, \dots, n$ . The relationship is established upon a strong assumption of the financial asset returns in the sub-sample. That is, we assume most asset returns are either serially uncorrelated or have weak serial correlations in the sub-sample. Let  $V^*$  denote the  $p^{*th}$  quantile of the loss random variable  $X$ , then we can write down

following equation

$$\begin{aligned}
[p^*]^n &= [Pr(X \leq V^*)]^n \\
&= \prod_{j=1}^n Pr(X_j \leq V^*) \\
&= Pr(M_n \leq V^*) \\
&= \begin{cases} e^{-(1+\xi \frac{V^*-b_n}{a_n})^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ e^{-e^{-\frac{V^*-b_n}{a_n}}} & \text{if } \xi = 0. \end{cases} \tag{5.20}
\end{aligned}$$

The 2<sup>nd</sup> equal sign in the above equation is based on assuming that asset returns are i.i.d. The last equal sign is given by replacing  $m_n^*$  by  $V^*$  in equation (5.18). By taking power of  $\frac{1}{n}$  on both sides of the above equation, we obtain

$$\begin{aligned}
p^* &= Pr(X \leq V^*) \\
&= \begin{cases} e^{-\frac{1}{n}(1+\xi \frac{V^*-b_n}{a_n})^{-\frac{1}{\xi}}} & \text{if } \xi \neq 0 \\ e^{-\frac{1}{n}e^{-\frac{V^*-b_n}{a_n}}} & \text{if } \xi = 0. \end{cases} \tag{5.21}
\end{aligned}$$

Then, the  $p^{*th}$  quantile of the loss random variable  $X$  can be obtained by solving the above equation for  $V^*$  as follows:

$$V^* = \begin{cases} b_n - \frac{a_n}{\xi} [1 - [-n \log(p^*)]^{-\xi}] & \text{if } \xi \neq 0 \\ b_n - a_n [\log[-n \log(p^*)]] & \text{if } \xi = 0. \end{cases} \tag{5.22}$$

Consequently, if  $X$  is the loss amount random variable over a time period  $h$ ,  $V^*$  is the VaR with  $p^*$  confidence level for a period of time  $h$ . By MLE estimates given in Table 5.1, we can calculate the VaR value with 95% confidence level for the next  $h = 10$  days period as follows:

$$\begin{aligned}
\text{VaR}_{10\text{day}, 0.95} &= 0.7856 - \frac{0.3689}{-0.1323} [1 - [-10 \log(0.95)]^{0.1323}] \\
&= 1.021318.
\end{aligned}$$

If one holds a long position of 1,000,000 EURUSD, the estimated VaR with 95% confidence level and 10 days period is equal to  $100,000 \times 0.01021318 = 10,213.18$

EUR. This is saying that in the next 10 days, we have 95% confidence level to say that our daily loss will not exceed 10,213.18 EUR if we hold 1 Million EUR.

Similarly, a VaR value with 99% confidence level with  $h = 20$  days can be calculated as

$$\begin{aligned} \text{VaR}_{20\text{day}, 0.99} &= 1.0108 - \frac{0.3513}{-0.1320} [1 - [-20\log(0.99)]^{0.1320}] \\ &= 1.518747. \end{aligned}$$

As expected, with higher confidence level and longer time period, we obtained a bigger VaR value.

## 5.6 Considering Volatility Clustering

Volatility has been a crucial ingredient in modeling financial time series, designing trading strategies and implementing risk management. In empirical finance, it is often found that asset return volatility is highly persistent in the sense that periods of high volatility tend to be followed by high volatility and periods of low volatility tend to be followed by low volatility. This behavior is well-known as **heteroskedasticity**. There are many articles providing empirical supports for the argument, such as such as [9], [23], and [19]. In this section, we will extend our tail risk estimation to include a stochastic volatility structure for the asset loss (or equivalently, negative return). In 1982, Article [8] proposed ARCH (autoregressive conditional heteroscedasticity) to model volatility dynamics by taking weighted averages of past squared forecast errors. In 1986, article [3] introduced a generalization method (GARCH), which extended the original ARCH model to allow lagged conditional variances enter as well.

In this section, we will calculate the tail risk measurement VaR with stochastic volatility dynamics being modeled by GARCH. According to [19], research shows that econometric models of volatility dynamics that assume conditional normality, such as GARCH models, do yield VaR estimates reflecting the current volatility background.

### 5.6.1 GARCH ( $p, q$ ) Model

Let  $X_t$  for all  $t \in N$  be a strictly stationary<sup>1</sup> time series representing daily observation of loss on a financial asset. Then,  $X_t$  is said to be a GARCH( $p, q$ ) process if it satisfies equation

$$X_t = \mu_t + \sigma_t Z_t \quad (5.23)$$

with

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i (X_{t-i} - \mu_{t-i})^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (5.24)$$

where  $\omega > 0$ ,  $\alpha_i > 0$ ,  $\beta_j > 0$  for all  $i$  and  $j \in N$ . Parameters  $\mu_t$  and  $\sigma_t$  are the **conditional mean** and **conditional variance** of  $X_t$  based on past information  $\mathcal{F}_{t-1}$ , which is the  $\sigma$ -field up to time  $t - 1$ . Process  $\{Z_t \text{ for all } t \in N\}$  is a series of independent and identically distributed random variable with mean 0 and variance 1.

The ( $p, q$ ) in parentheses is a standard notation in which the first number  $p$  refers to how many past squared error terms are included, while the second number  $q$  refers to how many moving average lags of past conditional variance terms are included. If we let  $q = 0$ , we obtained ARCH( $p$ ) model. In this research, we will consider the case where  $p = q = 1$ , which gives the mostly applied GARCH(1, 1) model.

### 5.6.2 Conditional Quantile

Let  $F_X$  and  $F_Z$  denote the marginal distribution functions of random variable  $X_t$  and  $Z_t$  respectively. For a future horizon of  $h$  days, we let  $Y_h = X_{t+1} + X_{t+2} + \dots + X_{t+h}$  represent the total loss random variable for the next  $h$  days. Then,  $F_{Y_h|\mathcal{F}_t}(x)$  is the predictive conditional distribution of loss random variable over next  $h$  day given the information of losses up to and including current day  $t$ . Then, the  $p^{\text{th}}$  **conditional quantile** of the predictive conditional distribution for the loss over the next  $h$  days

---

<sup>1</sup>A process is called strictly stationary if none of its finite moments depend on time.

is obtained by

$$\begin{aligned}
y_h^p &= \inf \{y \in \mathbf{R} : F_{Y_h|\mathcal{F}_t}(y) \leq p\} \\
&= \inf \{y \in \mathbf{R} : F_{X_{t+1}+X_{t+2}+\dots+X_{t+h}|\mathcal{F}_t}(y) \leq p\}.
\end{aligned} \tag{5.25}$$

If we let  $h = 1$  day, we can obtain

$$\begin{aligned}
F_{Y_1}(y) &= F_{X_{t+1}|\mathcal{F}_t}(y) \\
&= Pr(\mu_{t+1} + \sigma_{t+1}Z_{t+1} \leq y | \mathcal{F}_t) \\
&= Pr\left(Z_{t+1} \leq \frac{y - \mu_{t+1}}{\sigma_{t+1}} | \mathcal{F}_t\right) \\
&= Pr\left(Z_{t+1} \leq \frac{y - \mu_{t+1}}{\sigma_{t+1}}\right) \\
&= F_Z\left(\frac{y - \mu_{t+1}}{\sigma_{t+1}}\right)
\end{aligned} \tag{5.26}$$

since  $\{Z_t, \text{ for all } t \in N\}$  is a series of i.i.d. random variables.

Then, the  $p^{th}$  conditional quantile for the 1-step predictive conditional distribution for the loss over 1 day is given by

$$y_1^p = \mu_{t+1} + \sigma_{t+1}z^p, \tag{5.27}$$

where  $z^p$  is the  $p^{th}$  quantile of the marginal distribution of  $Z_{t+1}$ .

### 5.6.3 Empirical Analysis on GARCH(1,1)

In this section, we will conduct empirical analysis on EUR/USD daily closing rate by GARCH(1,1) model. Autocorrelation plot is most commonly used tool to visualize non-independency of the observations in a time series. Figure 5.6 shows the autocorrelation of EUR/USD daily loss at different time lags. The two horizontal lines are the 95% confidence level lower and upper bounds. For the time lags with autocorrelation value locating beyond the bounds, the null hypothesis that there is no autocorrelation at and beyond these time lags is rejected at a confidence level of 95%.

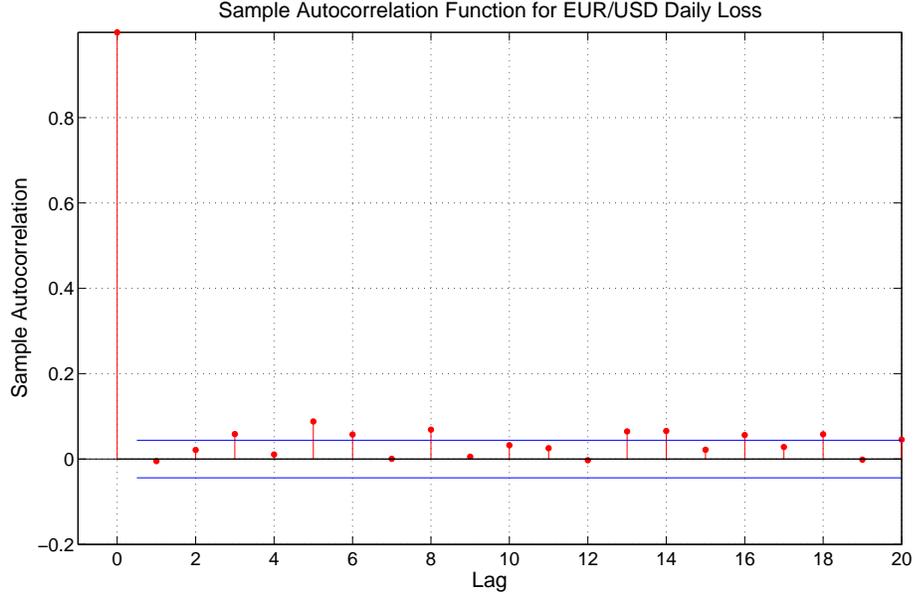


Figure 5.6: Autocorrelation Function of EUR/USD Daily Loss

The next step is to calculate the parameter values  $\mu_{t+1}$  and  $\sigma_{t+1}$ . We consider that EUR/USD daily loss time series is a realization of an AR(1) - GARCH(1,1) process. This means that the conditional mean being modeled by an AR(1) model and the conditional variance being modeled by the GARCH(1,1) model. Hence, we can obtain the following equations for conditional mean and conditional variance:

$$\mu_t = \phi_1 X_{t-1} + \phi_0, \quad (5.28)$$

and

$$\sigma_t^2 = \omega + \alpha_1 (X_{t-1} - \mu_{t-1})^2 + \beta_1 \sigma_{t-1}^2 \quad (5.29)$$

where  $\omega > 0$ ,  $\alpha_1 > 0$ , and  $\beta_1 > 0$ . Then we can fit the AR(1) - GARCH(1,1) model by Maximum Likelihood method with EUR/USD data, and assume that the innovations (or residuals)  $Z_t$  has a standard normal distribution. In Matlab, given the original time series of EUR/USD daily loss observations  $\{x_1, x_2, \dots, x_n\}$  for  $n = 2057$ , we can specify the model and obtain the parameter estimates  $\{\hat{\phi}_0, \hat{\phi}_1, \hat{\omega}, \hat{\alpha}_1, \hat{\beta}_1\}$ . Table 5.2 gives the estimates of these parameter values, and figure 5.7 shows the

plots of Innovations being inferred from the original series, the standard deviations, and the original EUR/USD daily loss series. To see the model effect, figure 5.8 gives the plot of autocorrelation functions of the innovations. We can clearly see that up to 20-day lags, there is no autocorrelation locating outside the lower and upper bounds.

Table 5.2: AR(1)-GARCH(1,1)Parameter Estimates given by MLE with Normal Innovations

$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$
-0.0119	-0.0358	0.0041	0.0177	0.9722

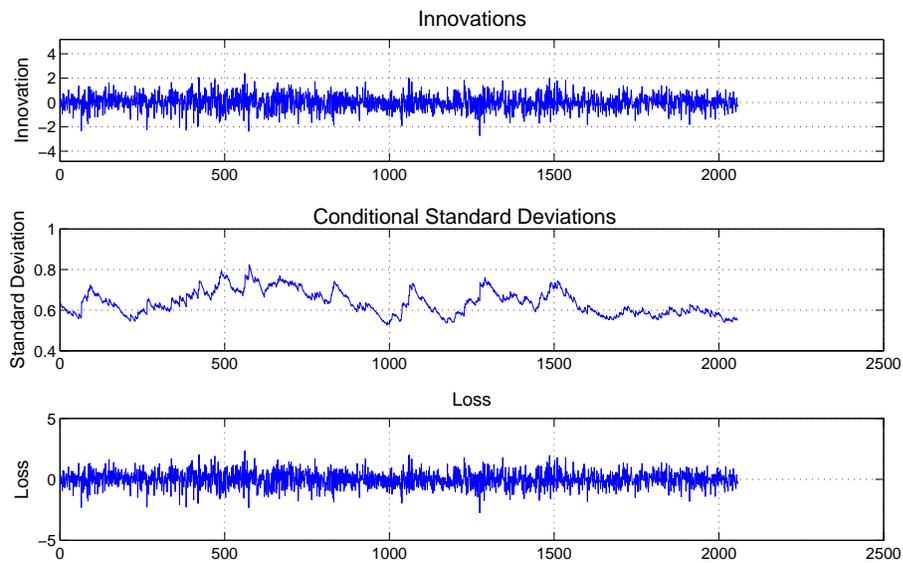


Figure 5.7: Inferred Innovations, Standard Deviations, and Original Time Series

Now we can calculate the conditional mean and conditional variance for day  $t + 1$  by applying equations

$$\hat{\mu}_{t+1} = \hat{\phi}_0 + \hat{\phi}_1 x_t \quad (5.30)$$

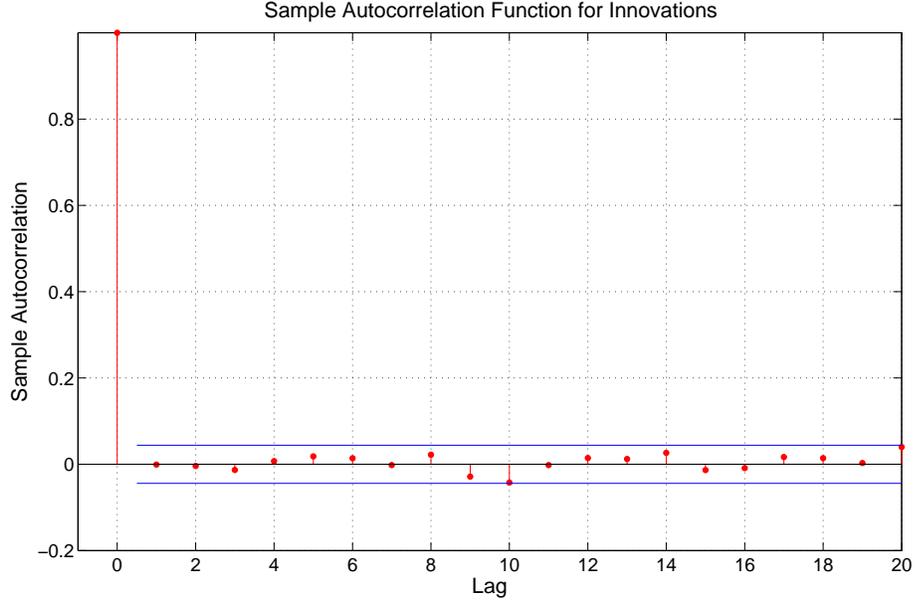


Figure 5.8: Autocorrelation Function of Innovations

and

$$\hat{\sigma}_{t+1}^2 = \hat{\omega} + \hat{\alpha}_1 (x_t - \hat{\mu}_t)^2 + \hat{\beta}_1 \hat{\sigma}_t^2. \quad (5.31)$$

We will obtain  $\hat{\mu}_{t+1} = 0.001407$  and  $\hat{\sigma}_{t+1}^2 = 0.299724$ .

Given the assumption that innovations  $Z_t$  are i.i.d. standard normal variables, the  $p^{th}$  quantile of  $Z$  can be obtained by  $z_p = \Psi^{-1}(p)$ , where  $\Psi^{-1}$  is the inverse of standard normal distribution function. Then, for our EUR/USD daily loss data, we can get the 99% conditional quantile for the 1-step predictive conditional distribution for the loss over 1 day as

$$\begin{aligned} \hat{y}_1^{0.99} &= \hat{\mu}_{t+1} + z^{0.99} \hat{\sigma}_{t+1} \\ &= 0.001407 + 2.326 \sqrt{0.299724} \\ &= 1.275014. \end{aligned}$$

This is also the 1-day VaR of the EUR/USD loss with 99% confidence level.

Another standard approach is to assume that the innovations have a leptokurtic

distribution such as Student's  $t$ -distribution (scaled to have variance 1). An AR(1) - GARCH(1,1) model with  $t$ -innovations can also be fitted with maximum likelihood and an additional parameter  $\nu$  (degree of freedom) can be obtained. By specifying the model in Matlab, we obtain following estimates for parameter values (given in Table 5.3) with assumption of  $t$ -innovations.

Table 5.3: AR(1)-GARCH(1,1)Parameter Estimates given by MLE with Student  $t$  Innovations

$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\nu}$
-0.0076	-0.0420	0.0031	0.0177	0.9747	14

We then obtain estimation values for conditional mean and conditional variance as  $\hat{\mu}_{t+1} = 0.008011$  and  $\hat{\sigma}_{t+1} = 0.299551$ , and consequently get the 99% conditional quantile for the 1-step predictive conditional distribution for the EUR/USD loss over 1 day as

$$\begin{aligned} \hat{y}_1^{0.99} &= \hat{\mu}_{t+1} + z^{0.99} \hat{\sigma}_{t+1} \\ &= 0.008011 + 2.624\sqrt{0.299551} \\ &= 1.444159. \end{aligned}$$

In above equation,  $z^{0.99} = 2.624$  is the 99<sup>th</sup> quantile of a Student's  $t$ -distribution with degrees of freedom  $\hat{\nu} = 14$ . As expected, we obtain a higher value for the conditional quantile of Student's  $t$ -distribution than Standard Normal.

Next, we assume the conditional variance has dynamics given by another forms of GARCH: **GJR-GARCH**. It consider one more quantity than the GARCH model in previous sections. That is, the asymmetric innovations. According to [5], GJR-GARCH is a more generic process of the evolution of the conditional variance given by

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i (\varepsilon_{t-i})^2 + \sum_{k=1}^o \gamma_k (\varepsilon_{t-k})^2 I_{[\varepsilon_{t-k} < 0]} \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (5.32)$$

where the extra parameter  $\gamma_k$  is the coefficient of the asymmetric error square. Integer values of  $p$ ,  $o$ , and  $q$  are the orders for the symmetric error squares, asymmetric error squares, and lagged variance term respectively. If we only focus on modeling the conditional variance  $\sigma_t^2$ , we obtain follow parameter estimates for GJR-GARCH(1, 1, 1) given in table 5.4.

Table 5.4: Parameter Estimates for GJR-GARCH(1,1,1)

Model	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\gamma}_1$	$\hat{\beta}_1$
GJR-GARCH(1,1,1)	0.0050	0.0135	0.0090	0.9696

It is quite obvious that the assumption of normal distribution for innovation is one of the biggest drawback for this empirical method. In previous sections, we already discussed that financial asset returns are with heavy tails. An approach of taking account of heavy tails would be to leverage generalized extreme value distribution for innovations in fitting GARCH Model. This can be an interest area for future research.

# Chapter 6

## VaR for A Trading Strategy

### 6.1 Ideas

This chapter will link the connection between Chapter 4 and Chapter 5 by using VaR as a performance measure of a risk trading strategy. Same as measuring the tail risk of a financial product, VaR can also be used to measure the tail risk of a portfolio's real-time P&L process. If the portfolio is managed by a specific risk trading strategy, measuring VaR of the portfolio's tail risk is equivalent in measuring the performance of a risk trading strategy. In previous chapters from 1 to 4, we explained the FX market making business process by its two major components: client trading and risk hedging. During busy market hours, live FX quotes come into the system in milliseconds, and client trades also come in at a higher frequency than during normal condition market. This leads to the real-time P&L process given by equations 3.14 and 4.6 to be updated at a much higher frequency, which will produce a large sample size of the P&L numbers in a very short time period, and it motives the idea of calculating real-time VaR during the intra-day as a performance measure for the trading strategy. So, the question that we ask will be "What is the extreme loss that the P&L process can have for the next 10 minutes trading period by using current trading strategy given a 99% level of confidence?"

## 6.2 Methodology and Example

The method we are going to apply is very similar to the method given in sections 5.2, 5.5, and 5.3. We will apply generalized extreme value distribution to the block maxima of the P&L process, calculate the MLE estimates of the underlying parameters, and apply equation 5.22 to calculate VaR.

For regulation purpose, VaR is calculated on a daily basis with a time period of  $h = 1$  or 10 days. For our purpose,  $h$  is more likely to be set at 5, 10, or 30 minutes depending on our preference. As the live rate updates are with un-equal time intervals, the first step is to apply linear interpolation method to produce a homogeneous time series of the P&L values. The choice of length for each time interval of the homogeneous time series is subject to the length of the period  $h$  of which VaR is quoted on. The shorter the  $h$ , the finer the time interval we need for linear interpolation in order to have a big enough sample for each block. The choice of  $h$  is really up to the user's preference. But a user should keep in mind that for non-busy trading hours,  $h$  is better to be a long interval than short. The reason is that the rate updates may be very slow during the non-busy hour compare to the busy ones. So does the P&L process. Thus, a very short  $h$  for a non-busy hour may contain only few real observations before applying linear interpolation. Another key point that affects the confidence in VaR calculation is the length of the history of the strategy. The longer the strategy runs, the more observations we have, hence the more confidence we have in the VaR calculation. If we decide to use only current intra-day data for VaR calculation, then we need a built-up time period before reaching a high confidence level in our estimation. By denoting each element of the homogeneous time series of P&L process by  $\{x_1, x_2, \dots, x_T\}$ , where  $T$  is the total sample size, we can apply exactly the same method as mentioned in sections 5.2, 5.5, and 5.3 to calculate VaR.

To give an example, let's run our Limit Position Trading Strategy, which was introduced in chapter 4, to simulate one P&L sample path for a 30-hour trading period for USD/CAD. During this 30-hour period, each of the first, second, and third

10-hour period is with the same assumption values of the three scenarios in section 4.3, which are scenarios of Balanced Client Buying and Selling Flows under Flat Market Condition, Intensive Client Selling Flow under Downward Market Condition, and Intensive Client Buying Flow under Upward Market Condition respectively.

The simulation result provides a sample with a number of 2,131 observations. Each is at the time stamp of an incoming client trade. Since we assume our risk hedging trades can be executed immediately in the market without any delay, the time stamps of risk hedging trades are identical to the time stamps of those client trades which hit the position limit and trigger the auto-hedging trades<sup>1</sup>. The time series of the sample observations is a non-homogeneous one, which has the time interval from 0 to 1,800 minutes. This mean that on average we have more than one observations per minute. Thus, we apply linear interpolation method to transfer the original non-homogeneous time series into a homogeneous one with time interval of 2-minute period between two consecutive points. Figure 6.1 shows the original and the interpolated time series with 2-minute interval period.

The ultimate objective for a market maker is to realize positive increments in the P&L process over a portfolio. Given the homogeneous time series of the P&L observations  $\{y_1, y_2, \dots, y_N\}$ <sup>2</sup>, we will perform the extreme value theory analysis and VaR calculation on the time series  $\{x_1, x_2, \dots, x_{N-1}\}$  for which  $x_k = y_{k+1} - y_k$  for  $k = 1, 2, \dots, N - 1$ . Figure 6.2 shows the P&L increment time series  $\{x_1, x_2, \dots, x_{N-1}\}$ , and figure 6.3 is the Q-Q plot of these observations, which suggests that the P&L increment has a heavy tail distribution.

Then we perform exactly the same procedures as in section 5.3 to model the block maxima by generalized extreme value distribution and to calculate the MLE estimates of parameters. Since we have in total of 30-hour trading history and a homogeneous time series of P&L increments with 2-minute time interval between two consecutive

---

<sup>1</sup>For the client trades simulation, we didn't apply any client spread in price. Thus, the P&L sample path doesn't contain client margin

<sup>2</sup>The interpolated time series with 2-minute interval period between two consecutive observations.

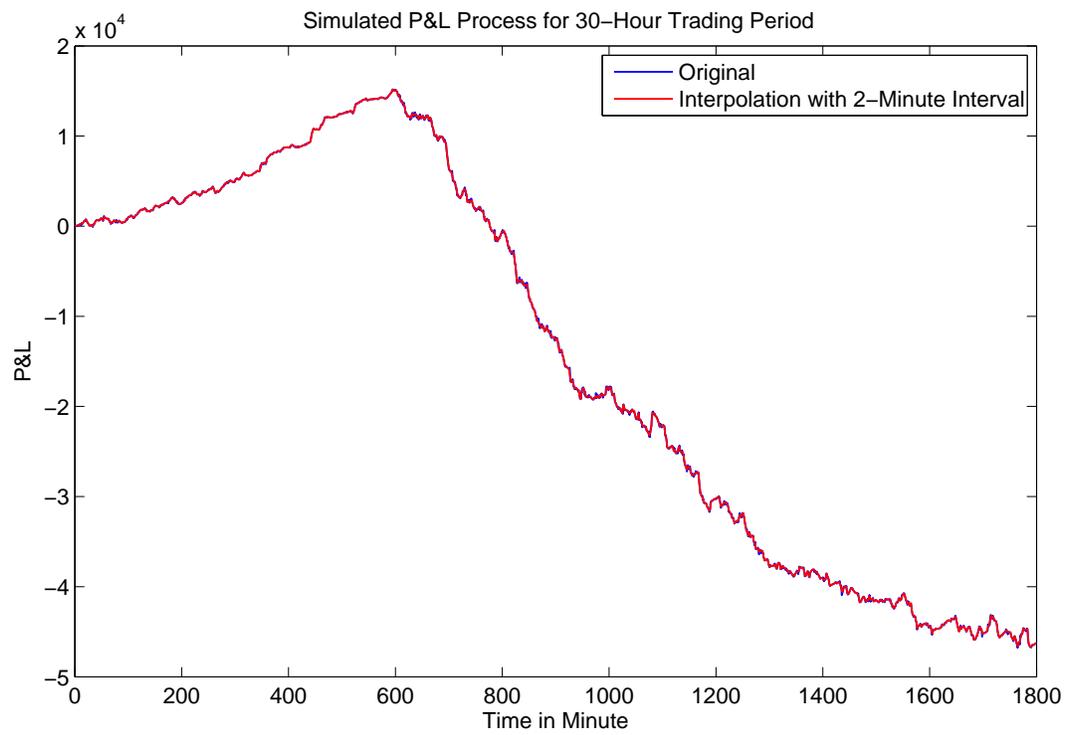


Figure 6.1: Simulated P&L Sample Path for a 30-hour Trading Period

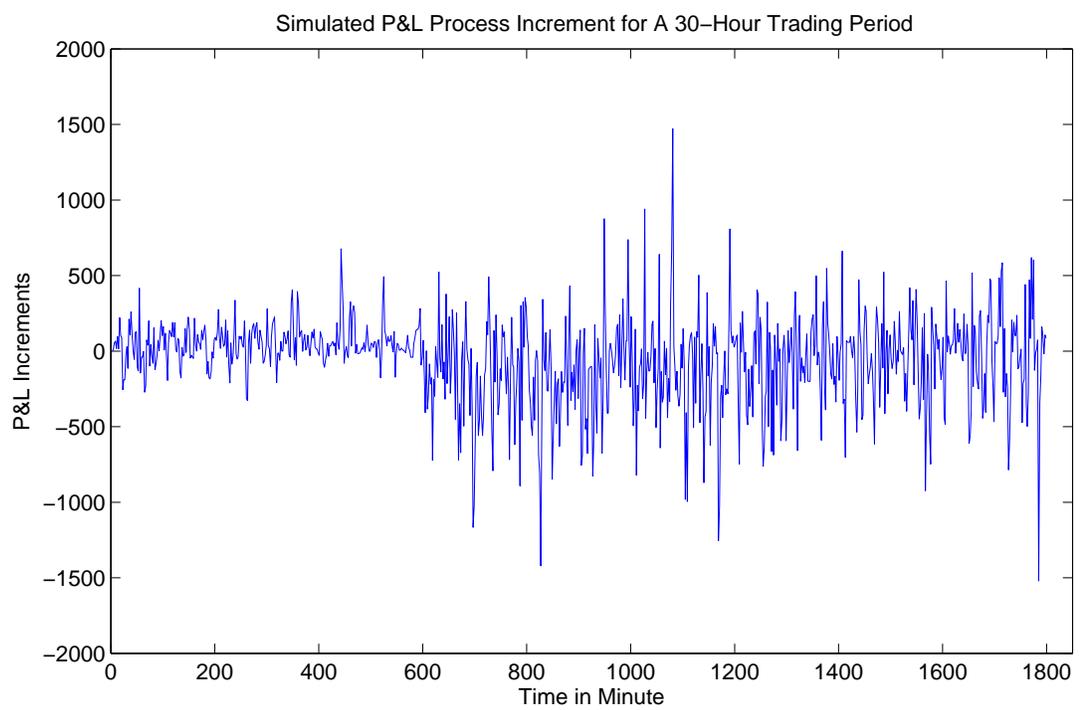


Figure 6.2: P&L Increments for a 30-hour Trading Period

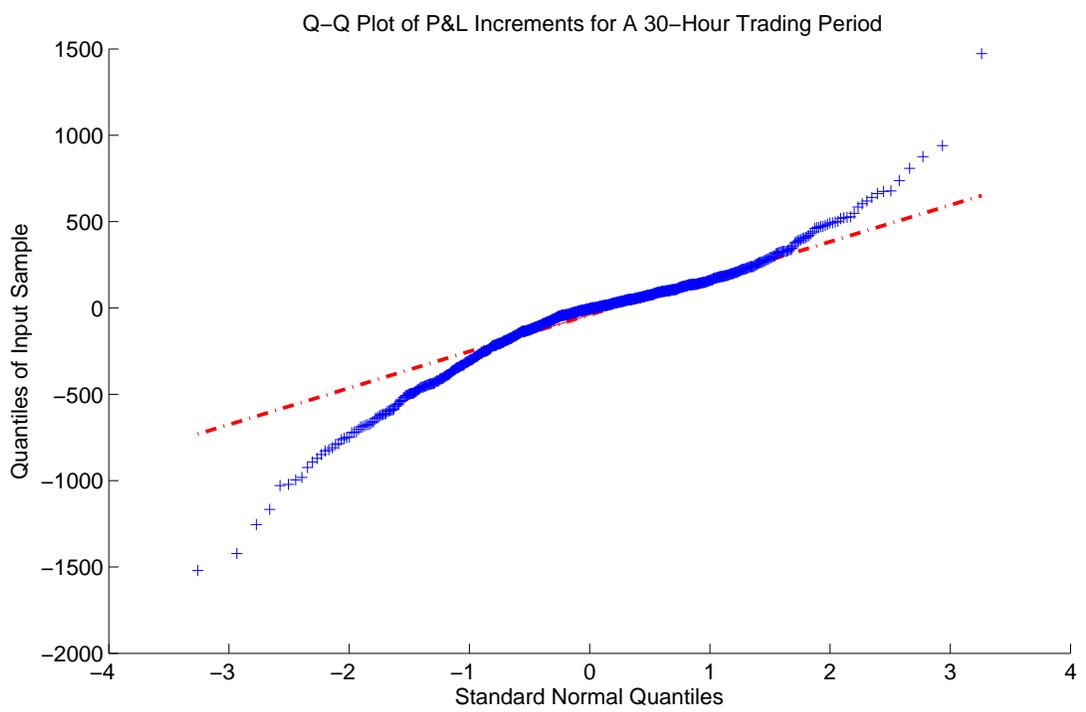


Figure 6.3: Q-Q Plot for P&L Increments for A 30-hour Trading Period

points, we divide the P&L increment time series into  $k = 30$  blocks with  $n = 30$  observations in each block. Figure 6.4 shows the maximum 2-minute loss (or negative P&L increment) in each hourly block.

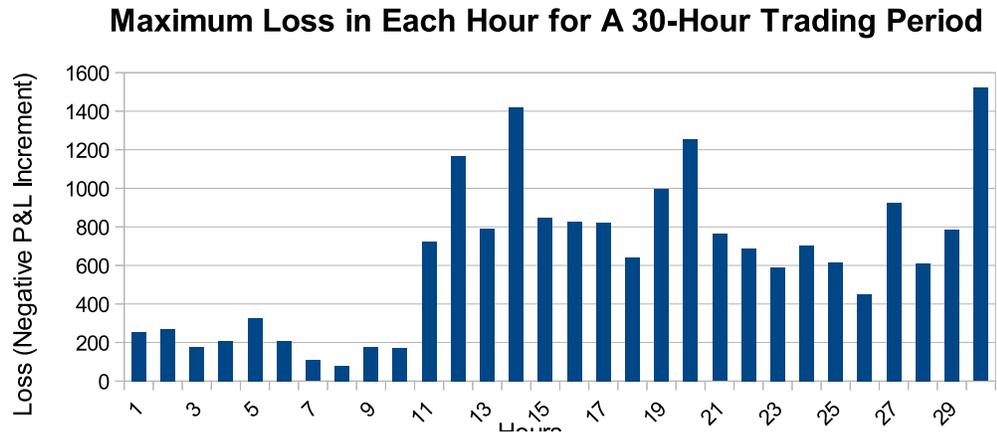


Figure 6.4: Maximum 2-Minute Loss in Hourly Block for A 30-Hours Period

By fitting the maximum block loss observations to the generalized extreme value distribution, we obtain the maximum likelihood estimates of shape parameter  $\hat{\xi} = -0.1070$ , scale parameter  $\hat{a}_n = 337.2511$ , and location parameter  $\hat{b}_n = 471.8377$ . Figure 6.5 compares the empirical cdf for the sample and the theoretical cdf of the generalized extreme distribution with estimated parameters.

Then, according to equation 5.22, we can calculate

$$\begin{aligned} \text{VaR}_{1\text{hour}, 0.95} &= 471.8377 - \frac{337.2511}{-0.1070} [1 - [-30 \log(0.95)]^{0.1070}] \\ &= 323.0779. \end{aligned}$$

This tells us that with a 95% confidence level that 2-minute P&L loss will not exceed 323.0779 in the next one-hour trading period.

With this method, it is possible to use VaR on a high-frequency level as a risk mea-

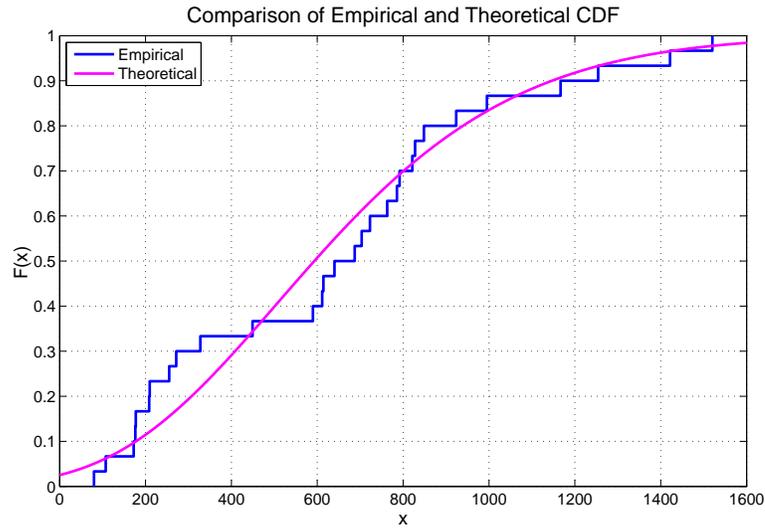


Figure 6.5: Empirical and Theoretical CDF Comparison for Maximum 2-Minute Loss in Hourly Block

surement for a high-frequency trading strategy. As the sample size increases during the trading period, the MLE estimations will be more and more accurate. Moreover, VaR can be calculate on a real-time basis so that it provides a good indication on how the strategy performs. If the VaR value is getting bigger and bigger, this may be a signal for us to shut down the strategy or at least to be cautious in operations. The choice of the block interval size can also be adjusted to serve different purposes of different groups. For traders, they may want the block interval to be as small as possible such as 5-minute, 3-minute, or even 1-minute. For regulators, they may only care about the 1-day or 10-day VaR.

# Chapter 7

## Conclusions and Future Work

In this research paper, we looked into literatures about global foreign exchange market in terms of its structure, product type, participants, and evolvments. We also looked into the FX high-frequency data structure, and implemented the Exponential Moving Average operator in matlab for processing the tick-by-tick data. Empirical analysis on the real data showed that a built-up period was needed for the EMA operator to produce accurate enough estimates. The bigger the range value of the operator had, the long the built-up period was needed.

In the second part of this research, we investigated how a market maker could effectively manage real-time risk as the counter-party of clients trading. In order to conduct the analysis, we introduced a framework for market high-frequency data simulation and client trading flow simulation. By using a Poisson process and a Geometric Brownian motion, we simulated the market data arrival process and the market data value respectively. The client trading flow can be simulated by a Poisson process and a modified Gaussian distribution. Then, market maker's Base Currency Wealth Process and Counter Currency Wealth Process were successfully defined with client trading amounts and market prices. In Chapter 4, we introduced a basic risk hedging strategy, which limits the position that a market maker can take during the trading horizon. Simulation results showed that when we set the risk limit at a static

level, the risk hedging strategy did not necessarily generate more revenue than a non-risk hedging strategy, but it helped reducing the downside risk substantially when the market faced an upward or downward rally.

In the third part of this research, we looked into the Extreme Value Theory and its extension to Value-at-Risk Calculation. By using eight years daily EUR/USD exchange rate data, we performed Maximum Likelihood Estimation methods to calculate estimates for shape, scale, and location parameters. Different estimates were obtained for different combinations of number of blocks and block sizes. Then, we extend the generalized extreme value distribution to VaR calculation based on the assumption that financial product has independent daily returns. Lastly, we applied GARCH(1,1) method for modeling volatility dynamics and calculated conditional quantile for the predictive conditional distribution function of asset loss.

One interesting avenue for future work is to identify links among risk limits, client trading flow, and market movements. Then, the risk limit can be dynamically adjusted according to a real-time market event and client flows so that optimal risk-adjusted returns can be obtained. For example, we can fix a client trading flow, and to search for the relationship between the risk limits and market volatility. Another interesting direction is to apply the Exponential Moving Average operator in risk hedging strategy. For example, we can use EMA operator to estimate where the market will be in the next couple of seconds or milliseconds, so that when we place risk hedging trades, we can place limit orders rather than market orders to avoid paying market spreads. The third interesting area is to leverage generalized extreme value distribution for innovations in fitting GARCH Model. This will be very helpful in removing the assumption that the underlying innovations are normally distributed.

# Bibliography

- [1] M. Austin. Adaptive systems for foreign exchange trading. *Quantitative Finance*, pages 37–45, August 2004.
- [2] E. C. Bank. *Review of the Foreign Exchange Market Structure*. First edition, March.
- [3] T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, pages 307–327, 1986.
- [4] J. Danielsson. Value at risk and extreme returns. *Working Paper*, page London School of Economics, January 2000.
- [5] J.-C. Duan. Approximating the gjr-garch and egarch option pricing models analytically. *Working Paper*, February 2004.
- [6] eFOREX. High frequency fx trading: Technology, techniques and data. *eFOREX*, pages 1–4, July 2007.
- [7] K. C. Embrechts, P. *Modelling Extremal Events for Insurance and Finance. Application of Mathematics*. Springer, second edition, 1997.
- [8] R. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, pages 987–1007, 1982.
- [9] R. Engle. Garch 101, the use of arch/garch models in applied econometrics. *Journal of Economic Perspectives*, pages 157–168, 2001.

- [10] B. Gnedenko. Sur la distribution limite du terme maximum of d'une serie aleatoire. *Annals of Mathematics*, pages 423–453, 1943.
- [11] S. Grimshaw. Computing the maximum likelihood estimates for the generalized pareto distribution to data. *Technometrics*, pages 185–191, 1993.
- [12] J. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, six edition, June 2005.
- [13] A. Jenkinson. The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Quarterly Journal of the Royal Meteorological society*, pages 158–171, 1955.
- [14] K. G. Koedijk. The tail index of exchange rate returns. *Journal of International Economics*, pages 93–108, 1990.
- [15] J. Labuszewski. Fx market growth and trends. *CME Research & Product Development*, 2006.
- [16] M. Levinson. *The Economist: Guide to Financial Markets*. Profile Books Ltd, fourth edition, 2006.
- [17] K. Lien. *The Foreign Exchange Interbank Market*. Investopedia Online, 2008.
- [18] F. M. Login. From value at risk to stress testing: the extreme value approach. *Working Paper*, pages Centre for Economic Policy Research, London, UK, 1999.
- [19] A. J. McNeil. Estimation of tail-related risk measure for heteroscedastic financial time series, an extreme value approach. *Journal of Empirical Finance*, pages 271–300, 2000.
- [20] U. Muller. Specially weighted moving averages with repeated application of the ema operator. Dec 2000.

- [21] R. Olsen. *An Introduction to High-Frequency Finance*. Academic Press, first edition, 2001.
- [22] S. Owens. The six forces of forex. *The Forex Report*, pages 22–37, July 2004.
- [23] A. Pagan. The econometrics of financial markets. *Journal of Empirical Finance*, pages 15–102, 1996.
- [24] F. R. and T. L.H.C. Limiting forms of the frequency distribution of largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, pages 180–190, 1928.
- [25] R. Reiss and M. Thomas. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhauser Verlag, Basel, first edition, 1997.
- [26] D. Rime. New electronic trading systems in foreign exchange markets. *New Economy Handbook*, pages 469–504, 2003.
- [27] S. Ross. *Introduction to Probability Models*. Academic Press, eighth edition, Dec 2002.
- [28] W. Sharp. Mutual fund performance. *Journal of Business*, pages 119–138, 1996.
- [29] S. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer-Verlag, first edition, 2007.
- [30] J. Wang. Optimal trading strategy and supply/demand dynamics. *National Bureau of Economic Research*, April 2006.

# Appendix A

## Proof of EMA Iteration Formula for $t$ Starting from $-\infty$

By equation (2.4), for a time point  $t^*$  such that  $t_{n-1} < t^* \leq t_n$ , we can write

$$\begin{aligned}
 EMA_Z(\lambda, t^*) &= \int_{-\infty}^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt \\
 &= \int_{-\infty}^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt + \int_{t_{n-1}}^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt \\
 &= \int_{-\infty}^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} e^{-\frac{1}{\lambda}(t^*-t_{n-1})} Z(t) dt + \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} \lambda Z(t) \Big|_{t=t_{n-1}}^{t^*} \\
 &= e^{-\frac{1}{\lambda}(t^*-t_{n-1})} \int_{-\infty}^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} Z(t) dt + Z(t^*) - e^{-\frac{1}{\lambda}(t^*-t_{n-1})} Z(t_{n-1}) \\
 &= \mu_1 EMA_Z(\lambda, t_{n-1}) + (\nu_1 - \mu_1) Z(t_{n-1}) + (1 - \nu_1) Z(t_n)
 \end{aligned} \tag{A.1}$$

with  $\mu_1 = e^{-\frac{1}{\lambda}(t^*-t_{n-1})}$  and value of  $\nu_1$  depending on the chosen interpolation scheme for  $Z(t^*)$ , where

$$\nu_1 = \begin{cases} 1 & \text{for the previous tick method} \\ 1 - \frac{t^* - t_{n-1}}{t_n - t_{n-1}} & \text{for the linear interpolation method.} \end{cases}$$

# Appendix B

## Proof of EMA Iteration Formula for $t$ Starting at Zero

By the general definition of moving average, which is given by (2.1), for a time point  $t^*$  such that  $t_{n-1} < t^* \leq t_n$ , we can write

$$\begin{aligned}
EMA_Z(\lambda, t^*) &= \frac{\int_0^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt}{\int_0^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} dt} \\
&= \frac{\int_0^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt}{\int_0^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} dt} + \frac{\int_{t_{n-1}}^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt}{\int_0^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} dt} \\
&= \frac{e^{-\frac{1}{\lambda}(t^*-t_{n-1})} \int_0^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} Z(t) dt}{\int_0^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} dt} + \frac{\int_{t_{n-1}}^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} Z(t) dt}{\int_0^{t^*} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t^*-t)} dt} \\
&= \frac{e^{-\frac{1}{\lambda}(t^*-t_{n-1})} - e^{-\frac{1}{\lambda}t^*}}{1 - e^{-\frac{1}{\lambda}t^*}} \frac{\int_0^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} Z(t) dt}{\int_0^{t_{n-1}} \frac{1}{\lambda} e^{-\frac{1}{\lambda}(t_{n-1}-t)} dt} \\
&\quad + \frac{Z(t^*) - e^{-\frac{1}{\lambda}(t^*-t_{n-1})} Z(t_{n-1})}{1 - e^{-\frac{1}{\lambda}t^*}} \\
&= \frac{e^{-\frac{1}{\lambda}(t^*-t_{n-1})} - e^{-\frac{1}{\lambda}t^*}}{1 - e^{-\frac{1}{\lambda}t^*}} EMA_Z(\lambda, t_{n-1}) + \frac{Z(t^*) - e^{-\frac{1}{\lambda}(t^*-t_{n-1})} Z(t_{n-1})}{1 - e^{-\frac{1}{\lambda}t^*}} \\
&= \mu_2 EMA_Z(\lambda, t_{n-1}) + (\nu_2 - \mu_2) Z(t_{n-1}) + (1 - \nu_2) Z(t_n)
\end{aligned} \tag{B.1}$$

with  $\mu_2 = \frac{e^{-\frac{1}{\lambda}(t^* - t_{n-1})} - e^{-\frac{1}{\lambda}t^*}}{1 - e^{-\frac{1}{\lambda}t^*}}$  and value of  $\nu_2$  depending on the chosen interpolation scheme for  $Z(t^*)$ , where

$$\nu_2 = \begin{cases} 1 & \text{for the previous tick method} \\ 1 - \frac{1}{1 - e^{-\frac{1}{\lambda}t^*}} \frac{t^* - t_{n-1}}{t_n - t_{n-1}} & \text{for the linear interpolation method.} \end{cases}$$