Tracking and Nowcasting Directional Changes in the Forex Market

By

Shuai Ma

A thesis submitted for the degree of Doctor of Philosophy

School of Computer Science and Electronic Engineering

Centre for Computational Finance and Economic Agents

(CCFEA)

University of Essex

March 2022

Abstract

Price changes in financial markets are typically summarized as time series (TS). Directional Change (DC) is an alternative, data-driven way to sample data points. The main objective of this thesis is to find new ways to extract new, useful information from the market. This is broken down into three directions: (1) to summarize price changes with DC, one must first determine the threshold to be used. We ask: could a threshold be too big or too small? If so, how could we determine the range of usable thresholds? (2) Could DC indicators extract volatility information from the market that is not observable under TS? (3) In DC, the start of a new trend is only confirmed in hindsight – to be precise, at the DC Confirmation (DCC) point when the price has reversed by the threshold specified. Could we detect that a new trend has begun before the DCC point? This is known as a nowcasting problem.

This thesis has made three contributions. Firstly, we have created a guideline to determine the range of useable thresholds under DC. This supports the research that follows. Secondly, we have demonstrated how DC indicators could complement TS in tracking the market for volatility information. Thirdly, we have introduced new DC indicators; by using these indicators, we have proposed an algorithm and demonstrated how it could help us nowcast whether a new trend has begun in DC.

Contents

Abstract .... **I**

Contents .... **II**

List of Figures .... **VI**

List of Tables .... **XI**

1. Introduction **1**

1.1. Overview **1**

1.2. Research Objective **4**

1.3. Thesis Structure **6**

1.3.1 Volatility of DC indicators 6

2. Chapter Two: literature review **9**

2.1. Time Series (TS) **9**

2.1.1 ARCH and GARCH model 14

2.1.2 Stochastic Volatility Model 17

2.1.3 Asset Price Models 21

2.1.4 historical volatility and realized volatility 27

2.2. Value at Risk (VaR) **29**

2.3. Directional Change (DC) **29**

2.3.1. Number of Directional Change (NDC) 38

2.3.2. Total Price Movement (TMV) 39

2.3.3. Time for completion of a trend (T) 43

2.3.4. Scaling laws under DC 44

2.4. Profiling **47**

2.5. Nowcasting **48**

3. Chapter Three: Useable range for Threshold **52**

3.1. Background **52**

3.2. The objective of this chapter **52**

3.3. Methodology **53**

3.3.1. Overview of the methodology 53

3.3.2. When is the threshold too small 54

3.3.3. When is the threshold too big? 56

3.4. Example One: EUR/USD Period One **57**

3.4.1. Data selected and range of thresholds to examine 57

3.4.2. Result for Period One 58

3.4.3. When is a threshold too small for profiling Period One 59

3.4.4. When is a threshold too big for profiling Period One 59

3.5. Example Two: EUR/USD Period Two 60

3.5.1. Data selected range of thresholds to examine.. 60

3.5.2. Results for Period Two 61

3.5.3. When is a threshold too small for profiling Period Two? 61

3.5.4. When is a threshold too big for profiling Period Two? 62

3.6. Conclusion 62

4. Chapter Four: Tracking market movement by multiple indicators 64

4.1. Background 64

4.2. Motivation and objective of this chapter 65

4.3. Overview of our approach 65

4.4. Methodology 67

4.4.1. Rolling windows and rolling speed 67

4.4.2. Volatility under TS 67

4.4.3. Volatility under DC 68

4.5. Experimentation 68

4.5.1. Data used 68

4.5.2. Experimental setup 68

4.6. Result 70

4.6.1. Overview 70

4.6.2. The results of the whole period P 70

4.6.3. Overview of the period P under investigation 72

4.6.4. Sub-periods that we focus on.. 73

4.6.5. Observations by SD in the four different periods 75

4.6.6. Observation by median $aTMV_{EXT}$ under DC in the four different periods...... 80

4.6.7. Observation by NDC under DC in the four different DC periods 87

4.6.8. Summary of all the results 92

4.7. Discussion 93

4.7.1. Information tracked by both TS and DC indicators.. 96

4.7.2. Information tracked by SD in TS but not by DC indicators 97

4.7.3. Information tracked by median $aTMV_{EXT}$ but not TS and NDC 97

4.7.4. Information tracked by NDC but not TS and median $aTMV_{EXT}$ 98

4.8. Experimental work on USD/JPY 99

4.8.1 Experimental work set up 99

4.8.2 the result of the whole period for both DC and TS... 101

4.8.3 Observation by DC indicators 102

4.8.4 Observation by TS 104

4.8.5 Observation by both DC and TS 106
4.8.6. Summary of all the USD/JPY results 107
4.9. Experimental work on GBP/USD 108
4.9.1 Experimental work set up 108
4.9.2 The result of the whole period 110
4.9.3 Observation by DC indicators 112
4.9.4 Observation by TS 114
4.9.5. Summary of all the GBP/USD results 115
4.10. Conclusion 116

5. Chapter five: Nowcasting new trends under DC 118
5.1. Background 118
5.2. New indicators, aTMV, Max TMV and Below Max 119
5.2.1. Indicators for DC nowcasting 120
5.2.2. Max 122
5.2.3. Below Max 122
5.3. Data Used and their profiles 125
5.4. Distribution of TMV_{EXT} 128
5.5. Max Below Max (MBM) and its value distribution 135
5.5.1. Max Below Max (MBM) 136
5.5.2. Distribution of Max Below Max (MBM) 137
5.6. Methodology 142
5.6.1. Overview of Nowcasting 143
5.6.2. Criteria for assessing nowcasting performance 145
5.6.3. Recording nowcasting performance statistics 150
5.6.4. Choice of parameters P_{TMV} and P_{BM} 152
5.7. Experimentation 153
5.7.1. Experimental setup 153
5.8. Result 156
5.8.1. Result for the Training period 156
5.8.2. Result for the Nowcasting period 162
5.9. Discussion 177
5.9.1. NCA produces timely nowcasts in terms of Price 178
5.9.2. NCA produces timely nowcasts in terms of Time 179
5.9.3. NCA is robust across market profiles 180
NCA is robust across thresholds 181
5.9.4. Rooms for improvement 182
5.10. Back testing for the necessity about Guideline of threshold 182

5.10.1 Data we used.......... 182
5.10.2 Result.......... 183
5.10.3 Discussion.......... 187
5.11. Experimental work for USD/JPY 188
5.11.1 Experimental work set up.......... 188
5.11.2 Result.......... 195
5.11.3 Discussion.......... 199
5.12. Experimental work for GBP/USD.......... 200
5.12.1 Experimental work set up.......... 200
5.12.2 Result.......... 208
5.12.3 Discussion.......... 212
5.13. Conclusion on Nowcasting 213
6. Chapter Six: Conclusion 216
6.1. Summary 216
6.1.1. The usable range of thresholds for profiling 216
6.1.2. Tracking and monitor markets with multiple indicators.......... 216
6.1.3. Nowcasting with multiple indicators.......... 217
6.2. Contributions 217
6.3. Future work 219
7. Appendix.......... 222
8. Reference 262

List of Figures

Figure 2. 1: Measuring Volatility in DC (Tsang 2017) 42

Figure 2. 2 3% DCs in EUR-USD-ByMinutes-2013 (Tao, 2017).... 43

Figure 2. 3 Directional Changes in foreign exchange rates between US Dollars and Japanese Yen (USD/JPY) with threshold = 2% (E.P.K. Tsang 2017) 49

Figure 4. 1 Rolling windows used in the experiments...................... 67

Figure 4. 2 NDC and median $aTMV_{EXT}$ for the rolling windows EUR/USD data from 25/09/2009 to 18/07/2016 under DC threshold 0.0004 71

Figure 4. 3 SD (log return) for the rolling windows under TS for EUR/USD data 71

Figure 4. 4 SD observed under TS from 03/02/2011 to 28/07/ 2010 (P_{TS}1) 76

Figure 4. 5 SD observed under TS: from 09/11/2011 to 11/01/2012 (P_{TS}2) 77

Figure 4. 6 SD observed under TS: from 03/15/2013 to 18/08/2013 (P_{TS}3) 78

Figure 4. 7 SD observed under TS: from 28/04/2016 to 07/07/2016 (P_{TS}4) 79

Figure 4. 8 Volatility under aTMV from 09/02/2010 to 29/07/2010

($P_{DC}1$) 81

Figure 4. 9 Volatility under median $aTMV_{EXT}$ from 15/11/2011 to 19/01/2012 ($P_{DC}2$) 82

Figure 4. 10 Volatility under median $aTMV_{EXT}$ 84

Figure 4. 11 Volatility under median $aTMV_{EXT}$ 85

Figure 4. 12 Volatility under NDC 09/02/2010 to 29/07/2010 ($P_{DC}1$) 87

Figure 4. 13 Volatility under NDC 15/11/2011 to 19/01/2012 ($P_{DC}2$) 89

Figure 4. 14 Volatility under NDC 03/05/2013 to 29/08/2013 ($P_{DC}3$) 90

Figure 4. 15 Volatility under 24/04/2016 to 20/06/ 2016 ($P_{DC}4$)...... 91

Figure 4. 16 NDC of rolling windows USD/JPY data from 27/09/2009 to 22/07/2015 under DC threshold 0.0004............................... 101

Figure 4. 17 SD (hourly log return) of rolling windows under TS for USD/JPY data 101

Figure 4. 18 Volatility under NDC 23/07/2012-23/01/2013 ($P_{USD/JPY}1$) 103

Figure 4. 19 Volatility under TS 07/04/2010 to 18/06/2010........... 103

Figure 4. 20 Volatility under TS 07/04/2010 to 18/06/2010. ($P_{USD/JPY}2$) 104

Figure 4. 21 NDC from 04/06/2010-06/06/2010............................ 105

Figure 4. 22 Volatility under NDC 03/05/2013-23/07/2013 ($P_{USD/JPY}3$) 106

Figure 4. 23 Volatility under TS 05/05/2013 to 16/07/2013. ($P_{USD/JPY}4$) 106

Figure 4. 24 Median_TMV_{EXT} of rolling windows GBP/USD 25/09/2009 to 22/07/2015 under DC threshold 0.0016 110

Figure 4. 25 SD (hourly log return) of rolling windows under TS for GBP/USD data 110

Figure 4. 26 Volatility under Median TMV_{EXT} 20121205-20130105 ($P_{GBP/USD}1$) 112

Figure 4. 27 Volatility under TS 06/12/2012 to 05/01/2013............113

Figure 4. 28 Volatility under TS 09/04/2010 to 04/06/2010. ($P_{GBP/USD}2$) 114

Figure 4. 29 Median TMV_{EXT} from 06/04/2010-02/06/2010 115

Figure 5. 1 The goal of this chapter in Figure 119

Figure 5. 2 Max.. 121

Figure 5. 3 Below Max and Maximum Below Max....................... 123

Figure 5. 4 The aTMV, Max and Below max values calculated by the EUR/USD data from 18:04:24 3/1/2016 to 18:04:45 3/1/2016 under threshold 0.0016.. 125

Figure 5. 5 Distribution of TMV_{EXT} in the Training period under

threshold 0.0016 130
Figure 5. 6 Distribution of TMV_{EXT} in the Training period under threshold 0.0032 132
Figure 5. 7 Distribution of TMV_{EXT} in Nowcasting period under threshold 0.0016 134
Figure 5. 8 Distribution of TMV_{EXT} in Nowcasting period under threshold 0.0032 135
Figure 5. 9 Distribution of MBM in Training period under threshold 0.0016 138
Figure 5. 10 Distribution of MBM in Training period under threshold 0.0032 140
Figure 5. 11 Distribution of MBM in Nowcasting period under threshold 0.0016 141
Figure 5. 12 Distribution of MBM in Nowcasting period under threshold 0.0032 142
Figure 5. 13 Any nowcast from t2 to t3 is correct and useful (i.e., good). Any nowcast before t2 is incorrect. Any nowcast after t3 is not useful 146
Figure 5. 14 Timeliness in Nowcasting: $Timeliness_{time}$ measures how much time the nowcast gains before the DCC took place; $Timeliness_{price}$ measures how much price the nowcast gains before the DCC price.. 149

Figure 5. 15 Good nowcast point 151

Figure 5. 16 Nogood nowcast point 152

Figure 5. 17 A nowcast on 22/01/2015 which produced the maximum $Timeliness_{Price}$ of 2.154 165

Figure 5. 18 A nowcast between 09:00:12 25 April 2015 to 20:27:19 27 April 2015 which produced the maximum $Timeliness_{Time}$ of 221,194 seconds; The x-axis counts the seconds from the first data point 168

Figure 5. 19 A nowcast on 21/01/2015 which produced the maximum $Timeliness_{Price}$ of 1.961 175

Figure 5. 20 Distribution of TMV_{EXT} in the Training periodT under threshold 0.00002 184

Figure 5. 21 Distribution of MBM in Training $period_{T}$ under threshold 0.00002 185

Figure 5. 22 Distribution of TMV_{EXT} in the Training $period_{USD/JPY}$ under threshold 0.0016 191

Figure 5. 23 Distribution of MBM in Training $period_{USD/JPY}$ under threshold 0.0016 193

Figure 5. 24 Distribution of TMV_{EXT} in the Training $period_{GBP/USD}$ under threshold 0.0032 204

Figure 5. 25 Distribution of MBM in Training $period_{GBP/USD}$ under threshold 0.0032 206

List of Tables

Table 3. 1 Index of (n) and Threshold (Th) used for in this paper 57

Table 3. 2 EUR/USD Period One's mean aTMV and minimum aTMV under different thresholds. 58

Table 3. 3 EUR/USD Period Two's mean aTMV and minimum aTMV under different thresholds. 61

Table 4. 1 The four periods under both DC and TS with significant changes in one of them or both of them 75

Table 4. 2 Summary of the findings for both TS and DC........................ 95

Table 5. 1 Tick-to-tick data used in this chapter, with their profiles compared 127

Table 5. 2 The TMVEXT we pick to use under the distribution of TMVEXT in the Training period under threshold 0.0016.............. 132

Table 5. 3 Distribution of TMVEXT in Training period under threshold 0.0032... 133

Table 5. 4 Distribution of Max Below Max in Training period under Threshold 0.0016 139

Table 5. 5 Distribution of Max Below Max in Training period under Threshold 0.0032 141

Table 5. 6 The result of NCA in the Training period under threshold 0.0016, with PTMV=1.68, PBM=0.68 157

Table 5. 7 The result of NCA nowcast in the Training period under threshold 0.0032, with PTMV=1.60, PBM=0.61 160

Table 5. 8 The result of NCA nowcast in the Nowcasting period under threshold 0.0016, with PTMV=1.68, PBM=0.68 163

Table 5. 9 The result of NCA nowcast in the Nowcasting period under threshold 0.0032, with PTMV=1.61, PBM=0.61 171

Table 5. 10 Summary of key performance measures in all the experiments 178

Table 5. 11 The MBM we pick to use under the distribution of MBM in the Training periodT under threshold 0.00002 186

Table 5. 12 The TMVEXT we pick to use under the distribution of TMVEXT in the Training periodT under threshold 0.00002 186

Table 5. 13 The result of NCA in the Training period under threshold 0.00002, with PTMV=19.57, PBM=0.6457 187

Table 5. 14 Tick-to-tick data used in this section, with their profiles compared.. 189

Table 5. 15 The TMVEXT we pick to use under the distribution of TMVEXT in the Training periodUSD/JPY under threshold 0.0016 192

Table 5. 16 Distribution of Max Below Max in Training period under Threshold 0.0016 194

Table 5. 17 Both Training periodUSD/JPY and Nowcasting

periodUSD/JPY nowcasting result 196

Table 5. 18 Summary of key performance measures of nowcasting iunder USD/JPY 199

Table 5. 19 Tick-to-tick data used in this section, with their profiles compared 202

Table 5. 20 The TMVEXT we pick to use under the distribution of TMVEXT in the Training periodGBP/USD under threshold 0.0032 205

Table 5. 21 Distribution of Max Below Max in Training periodGBP/USD under Threshold 0.0032 207

Table 5. 22 Both Training periodGBP/USD and Nowcasting periodGBP/USD nowcasting result 209

Table 5. 23 Summary of key performance measures of nowcasting under GBP/USD 212

1. Introduction

1.1. Overview

Price changes in a financial market are typically summarized using time series formats. Directional Change (DC) is an alternative, data-driven way to sample date points, which samples a data point whenever the market changes direction (Glattfelder et al 2011).

This research is built on the framework supported using Directional Change. It builds on the research conducted by Tao (2018), which defined a number of indicators for this approach, and the regime change detection research by Tsang and Chen (2018; Chen and Tsang 2018, 2020).

The underlying argument in the first part of this thesis (Chapter 3), supported by the research undertaken, is that to start a Directional Change observation, an observer must determine how big a percentage of a price change in the opposite direction constitutes a directional change. This percentage is called a threshold in DC research. Tao (2018) proposed the collection of statistical measures to profile a given market in each period. We argue that for such statistical measures to be useful, the threshold must be neither too big nor too small.

We argue that when the threshold is too small, every transaction in the opposite direction constitutes a DC; such profiles may potentially capture a lot of noise. On the other hand, when the threshold is too big, there will be few trends in the DC profile. As a consequence, any statistical measures thus collected be based on too few data points to be meaningful. In this research, therefore, data-driven guidelines are proposed for determining when the threshold is too small or too big for profiling a market period.

The second part of this thesis (Chapter 4) builds on the following argument: DC is an additional way to summarize the market, offering another angle from which to observe the market. We argue that market movement should be observed using several different measures of volatility. In Time Series (TS) observations, volatility is commonly defined as the standard deviation of log return over a certain time period. In DC, however, a variety of indicators can be used to define different kinds of volatility. We argue that by using multiple indicators, it is possible to understand the market more thoroughly than by only using TS alone.

Based on this, the current work presents a new research method to track market changes by combining both TS and DC. The value of this approach was then demonstrated by tracking the volatility of the EUR/USD market, (the most frequently used currency trading market), over several years (June 2009 to July 2016), using standard deviation in log returns, under

Time Series (TS), and various indicators under DC.

In this chapter, we will be using the concept of profiling to summarize all information that we collect from the DC data. Profiling is an application of directional changes to turn raw data into information. In order to track market movement in the market, it is necessary to capture the similarities and differences in prices movements within and between periods, based on observing the stylised behaviours of financial series. For this purpose, a computer program that includes some indicators to help analyse market movements is required. These indicators can then be used to measure the stylized behaviours of price movements under directional changes. The outcomes of the program are independent profiles, with each profile illustrating the performance within certain periods of the data under the selected threshold. Comparing and analysing the similarities and differences between the various metrics of these profiles may thus help with tracking and predicting abnormal events in the market.

In the third research chapter (Chapter 5), we investigate the concept of Nowcasting, an idea introduced by Giannone, Reichlin, and Small (2008), which is defined as the development of a very narrow prediction of the financial market featuring only the present, the very near future, and the very recent past. Castle et. al (2013) further noted that “forecasts are made before a period commenced, with ‘Nowcasts’ during the relevant period, and with flash estimates immediately, or shortly after the period ends, when disaggregated information remains incomplete.

Nowcasting thus allows recognition of the current state of the market and allows some awareness of trends in real-time. In DC, trends are only confirmed in hindsight. We ask whether it is possible to nowcast that a new trend has begun by tracking the transactions in real-time. The chapter thus develops a new algorithm with some new DC indictors for nowcasting.

1.2. Research Objective

This study aims to track and nowcast the market movement by using the Directional Change approach. To support that, we first work out what DC thresholds are usable for tracking and nowcasting purposes. Therefore, this research attempt to achieve the following objectives:

Objective 1: Determining appropriate DC thresholds to use

Directional Change (DC) requires the observer to determine how big a percentage change in the opposite direction is significant in terms of determining that a directional change has taken place. This percentage is called a threshold in DC research. This thesis aims to develop guidelines for determining when the threshold is too small or too big to profile a chosen market-period. Following the principle of DC, these guidelines must be data-driven.

Objective 2: With appropriate DC thresholds, track the volatility of the market

In Time Series (TS), volatility is defined as the standard deviation of returns within a time period, while in DC, various indicators are used to define different kinds of volatility. This thesis attempts to establish that the volatility of the market can be better assessed using multiple indicators.

Objective 3: Nowcast new trends

Nowcasting, a relatively new concept introduced by Giannone, Reichlin, and Small (2008), is defined as the narrow prediction of only the present, the very near future, and the very recent past, for a financial market. Directional Change is also a relatively new concept, and this work combines these approaches to develop an algorithm able to predict whether the financial market is moving from one state to another in close to real-time.

The objectives of this thesis are therefore to create guidelines for determining useable thresholds under DC, to produce a new method to tracking market movement by combining DC and TS, and to develop a method for predicting whether the financial market is moving from one state to another in close to real-time using the concept of Nowcasting together with Directional Change.

1.3. Thesis Structure

This thesis is organised as follows:

In Chapter Two, we begin by reviewing the main areas of research on Directional Change and Nowcasting. Besides, this chapter also provides a general overview of the concept of Time Series and its application.

In Chapter Three, we explain how setting the threshold in DC to be too big or too small negatively affect the characteristics of DC profiling. This motivates us to establish guidelines for determining the usable range of thresholds. We recommend that all scientific research on DC profiling should follow these guidelines.

In Chapter Four, DC is offered as an additional way to assess the volatility of a market. We shall demonstrate the benefits of using indicators from both DC and TS. Through empirical studies in the EUR/USD market (the most frequently used currency market), we highlight that volatility may be low under one indicator, while high according to another, but that all such indicators may be useful in the correct circumstances.

1.3.1 Volatility of DC indicators

Under time series, volatility can be measured by the standard deviation of the returns over a period of time (e.g. the 1-days volatility measured by the standard deviation of the last daily returns). Tsang et al (2017) explain that under DC, each trend may take a different amount of time (T) to complete. The smaller T is, the more frequently the market has changed between uptrend and downtrend. We can conclude that 1/T is a measure of volatility in terms of trend-switching frequency. We can also use 1/ Median T to measure the volatility over a period, where Median T is the median T value of all the trends in that period.

Frequency is only one way to measure volatility under DC. TMV measures the magnitude of price change in each trend. The higher the magnitude, the more volatile the market is. Given that theta is fixed within an observation, high TMV indicates high overshoot.

Frequency and magnitude are orthogonal with each other. Figure 2.1 shows hypothetical frequencies and magnitudes in three markets A, B and C. Market A is more volatile than Market B because while they share the same TMV, Market A takes only half of the time that Market B takes to complete each trend. Market B is more volatile than Market C because although they take the same time to complete each trend, the trends in Market B have bigger TMV values than those in Market C.

In Chapter Five, we shall formulate the nowcasting problem in DC. This chapter then identifies some new DC indicators that could be useful for nowcasting new trends in the market. We shall propose a nowcasting algorithm and assess its performance in tracking high-frequency data.

In the final chapter, a conclusion based on the findings within the research chapters is offered, alongside a discussion of their significance. The contribution of the work and possible future directions are also discussed.

2. Chapter Two: literature review

2.1. Time Series (TS)

This research requires a consideration of the concept of time series, which can be described as a series of data points taken over a period registered in a regular sequence and in time order. These are commonly plotted online charts and feature not only in recording and marking trends in financial markets but also in many other fields of study such as statistics, econometrics, weather forecasting, earthquake prediction, applied science, and engineering. The object of time series is to allow the analysis of data to produce useful statistics, particularly for tracking and forecasting, based on past patterns and events, which requires the ability to project the data forward to predict future events.

Time series data can be turned into a mathematical model for the purposes of prediction, or monitoring, to allow past or future trends to be observed. However, time series observations are potentially complex, due to the links between the various observations made over a period, which reflect the stochastic relationships within the data.

According to Pole et al (1994), there are three basic factors within any time series: simple time trends, systematic cyclical variation, and influential or causal variables. Combinations of these can therefore produce several flexible dynamic models for diverse analysis and forecasting purposes (Pole et al, 1994), and these are widely used in financial market analysis.

One of the most important indicators in financial markets is volatility, which measures the risk of a market, therefore, need to be developed to chart this volatility over time, as volatility affects where, and how much a financial or stock market price moves. Volatility can be dramatic, reaching large highs or lows, or maybe relatively stable, reflecting low volatility. Volatility movements have pluses and minuses for traders and investors, and in any case act as a benchmark for financial market activity. The usual measure of volatility is the standard deviation, a standard time series metric that draws on the average price deviation by which a stock differs from the average over a period of time. This work, however, seeks to extend such consideration by researching various methods of measuring volatility in the financial market. Roll (1984) maintained, for example, that volatility is affected by the microstructure of the market, which reflects how exchanges happen in financial markets.

In examining time series, it is common to use the standard deviation of log returns to measure volatility. This requires the use of the formula:

$$\sigma_T = \sigma_{hourly}\sqrt{T}$$

which represents the hourly volatility of a data set for time period T.

observers can set up the volatility as minutely, hourly, daily or monthly as they want, depending on the data set.

However, one must realize that this is just one of many ways to measure volatility. Actual historical volatility, which refers to the volatility of a financial instrument over a specified period, but with the last observation on a date in the past, is realised volatility, calculated as the square root of the realised variance, which is calculated using the sum of squared returns divided by the number of observations.

Actual future volatility refers to the predicted volatility of a financial instrument over a specified period starting at the current time and ending at a future date (normally the expiry date of an option).

Implied volatility similarly has several forms:

Historical implied volatility refers to the implied volatility observed from historical prices of the financial instrument (normally options).

Current implied volatility refers to the implied volatility observed from the current prices of the financial instrument.

Future implied volatility which refers to the implied volatility observed from predictions of future prices of the financial instrument.

For a financial instrument whose price follows a Gaussian random walk or

Wiener process, the width of the distribution increases as time increases, based on an increasing probability that the instrument's price will be further away from the initial price as time goes on. However, rather than increasing linearly, the volatility increases with the square root of time, as some fluctuations are expected to cancel each other out. Thus, the most likely deviation after two periods will not be twice the distance from zero observed after one period.

As observed price changes do not follow Gaussian distributions, other options for analysis such as the Lévy distribution are often used that can capture attributes such as "fat tails". Volatility is, at base, simply a statistical measure of dispersion around the average of any random variable such as market parameters, however.

Financial time series data are now widely available, including very long records of daily closing prices for series such as the Standard & Poors composite equity indices and high-frequency data, consisting of the complete history of transaction times, fees, and quotes for financial securities like individual stocks. However, modelling and statistical analysis of economic time series are relatively recent topics of scientific inquiry. Historically, time series analysis dealt primarily with applications in engineering, physical sciences, and earth sciences. The models developed for these applications were often based entirely on second-order properties of the data as described by the mean and covariance functions. Since a Gaussian process is entirely determined by its second-order properties, it was implicitly assumed that the process was also Gaussian. As a result of the World decomposition, it was then sufficient to consider only linear time series models, and the class of Fractionally Integrated Autoregressive Moving Average (ARIMA) processes provides aa highly flexible and dense type of models from which to model the covariance function. (Torben G. Andersen and etc, 2007)

Financial time series data are now widely available, including very long records of daily closing prices for series such as the Standard & Poors composite equity indices and for high-frequency data, consisting of the complete history of transaction times, prices, and quotes for financial securities like individual stocks. However, modelling and statistical analysis of financial time series are rather recent topics of scientific inquiry. Historically, time series analysis dealt primarily with applications in the fields of engineering, physical sciences, and earth sciences. The models developed for these applications were often based entirely on second-order properties of the data as described by the mean and covariance functions. Since a Gaussian process is completely determined by its second-order properties, it was implicitly assumed that the process was also Gaussian. As a result of the World decomposition, it was then sufficient to consider only linear time series models and the class of Fractionally Integrated

Autoregressive Moving Average (FARIMA) processes provides an extremely flexible and dense class of models from which to model the covariance function. (Torben G. Andersen and etc, 2007)

In this part, I will do a brief introduction to Time Series, and give some introductions to different models of volatility under time series. These will include ARCH model, GARCH model, Stochastic Volatility Model, Asset Price Models, and some other models' development under time series.

2.1.1 ARCH and GARCH model

Financial economists are concerned with modelling volatility in asset returns. This is important as volatility is considered a measure of risk, and investors want a premium for investing in risky assets. Banks and other financial institutions apply so-called value-at-risk models to assess their risks. Modelling and forecasting volatility or, in other words, the covariance structure of asset returns, is therefore important. The fact that volatility in returns fluctuates over time has been known for a long time. Originally, the emphasis was on another aspect of return series: their marginal distributions were found to be leptokurtic. Returns were modelled as independent and identically distributed over time. In a classic work, Mandelbrot (1963) and Mandelbrot and Taylor (1967) applied so-called stable Paretian distributions to characterize the distribution of returns. Rachev and Mittnik (2000) contain an informative discussion of stable

Paretian distributions and their use in finance and econometrics. Observations in return series of financial assets observed at weekly and higher frequencies are in fact not independent. While observations in these series are uncorrelated or nearly uncorrelated, the series contains higher-order dependence.

Models of Autoregressive Conditional Heteroskedasticity (ARCH) and the Generalized ARCH (GARCH) will be introduced in this part.

Timo (2009) said that the ARCH model is the first model of conditional heteroskedasticity.

We let εt be a random variable that has a mean and a variance conditionally on the information set $F_t - 1$ (the σ-field generated by $\varepsilon t - j, j \geq 1$). The ARCH model of εt has the following properties.

First, $E\{\varepsilon_t | F_t - 1\} = 0$ and, second, the conditional variance $h_t = E\{\varepsilon_{2t} | F_t - 1\}$ is a nontrivial positive-valued parametric function of $F_t - 1$. The sequence $\{\varepsilon t\}$ may be observed directly, or it may be an error or innovation sequence of an econometric model. In the latter case,

$$\varepsilon_t = y_t - \mu_t(y_t)$$

where y_t is an observable random variable and $\mu_t(y_t) = E\{y_t | F_{t-1}\}$, the conditional mean of $y\ t$ given F_{t-1}. The sequence $\{\varepsilon\ t\}$ may be observed directly, or it may be an error or innovation sequence of an econometric model.

Engle's (1982) application was of this type. In what follows, the focus will be on parametric forms of h_t, and for simplicity, it is assumed that μ_t (y_t)=0. Engle assumed that ε t can be decomposed as follows: $\varepsilon_t = z_t$ h 1 / 2 t where $\{z_t\}$ is a sequence of independent, identically distributed random variables with zero mean and unit variance. This implies ε t $|F_{t-1} \sim$ D (0, h t) where D stands for the distribution (typically assumed to be a normal or a leptokurtic one). The following conditional variance defines an ARCH model of order q

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 \qquad \text{(X)}$$

where α 0 > 0 , α j ≥ 0 ,j =1 ,...,q − 1 ,and α q > 0 . The parameter restrictions in the formula above form a necessary and sufficient condition for positivity of the conditional variance. Suppose the unconditional variance $E\epsilon_t^2 = \sigma^2 < \infty$ The definition of ε_t through the decomposition involving z_t then guarantees the white noise property of the sequence { ε_t }, since { z_t } is a sequence of variables. Although the application in Engle (1982) was not a financial one, Engle and others soon realized the potential of the ARCH model in financial applications that required forecasting volatility. The ARCH model and its generalizations are thus applied to modelling, among other things, interest rates, exchange rates and stock and stock index returns.

In applications, the ARCH model has been replaced by the so-called generalized ARCH (GARCH) model that Bollerslev (1986) and Taylor (1986) proposed independently of each other. In this model, the conditional variance is also a linear function of its own lags and has the form

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \qquad \text{(Y)}$$

The conditional variance defined by the formula above has the property that the unconditional autocorrelation function of ε_t^2 , if it exists, can decay slowly, albeit still exponentially. For the ARCH family, the decay rate is too rapid compared to what is typically observed in financial time series, unless the maximum lag q in X is long. As Y is a more parsimonious model of the conditional variance than a high-order ARCH model, most users prefer it to the simpler ARCH alternative.

2.1.2 Stochastic Volatility Model

Neil and Torben (2009) point out that stochastic volatility (SV) models are used heavily within the fields of financial economics and mathematical finance to capture the impact of time-varying volatility on financial markets and decision making. The development of the subject has been highly multidisciplinary, with results drawn from financial economics, probability theory and econometrics blending to produce methods that aid our understanding of option pricing, efficient portfolio allocation and accurate risk assessment and management. Time-varying volatility is endemic in financial markets. Black and Scholes (1972, p. 416), suggested that “there is evidence of nonstationary in the variance. More work must be done to predict variances using the information available.” Neil and

Torben (2009) also said that ARCH processes are often described as SV. The feature of ARCH models is that they explicitly model the conditional variance of returns given past returns observed by the econometrician. This very powerful idea that one-step-ahead prediction approach to the volatility model, especially in the field of risk management. It is convenient from an econometric viewpoint as it immediately delivers the likelihood function as the product of one-step-ahead predictive densities.

In the SV theory, the predictive distribution of returns is specified indirectly, via the structure of the model, rather than directly. For a certain number of SV models this predictive distribution can be calculated explicitly but, invariably, empirically study said that realistic representations must be computed numerically. This has some advantages is moves away from direct one-step-ahead predictions. In continuous time it is much convenient and might be more natural to model directly the volatility of asset prices as having its own stochastic process without thinking about the implied one-step-ahead distribution of returns recorded over an arbitrary time interval convenient for the econometrician, for example a month or a year. SV models is not directly available when raise some difficulties as the likelihood function for SV models, most of them are frustration of econometricians in the late 1980s and early 1990s. Neil and Torben (2009). Neil and Torben (2009) also said that only in the 1990s were novel simulation strategies developed to efficiently estimate SV models. These computationally intensive methods enable us, given enough coding and computing time, to efficiently estimate a broad range of fully parametric SV models. And its resulting enriched SV literature has brought us much closer to the empirical realities we face in financial markets.

In the late 1990s, with the development of high-frequency data become wildly used in the econometric analysis of volatility forecasting Neil and Torben (2009) reported that the connection between SV and realized volatility has allowed financial econometricians to combine the enriched information set available through high-frequency data, by an order of magnitude, the accuracy of their volatility forecasts over that traditionally offered by ARCH models based on daily observations.

Eric Renault (2009) suggested that there are moment-based models that are based on the SV model.

Eric (2009) said that based on the Method of Moments (MM) or the Generalized Method of Moments (GMM) often applied since the early days of the Stochastic Volatility (SV) literature. There are at least two justifications for these approaches. First, moments of financial time series have always been of high interest as such moments are associated not only with volatility forecasting but also aspects of the return distribution like heavy tails and return-volatility asymmetries, examples were presented by Rosenberg (1972) and Black (1976). Secondly, besides modelling issues, MM approaches are famous for their simplicity as the exact likelihood function is difficult to evaluate in the context of parametric volatility models within the class of hidden Markov models.

People normally use the Regression Model to Analyze Fluctuations in Variance, these can be the linear regression model for conditional variance such as Rosenberg (1972) is to be the first to realize that fat tails observed in asset prices changes:

$$z_{t+1} = \log \left(\frac{P_{t+1}}{P_t}\right)$$

Where the formula can be explained by a decomposition:

$$z_{t+1} = m_t + \sigma_t \epsilon_{t+1}$$

That the ε_t are serially independent random variables with identical distribution function F (·) having mean equal to zero, variance is equal to one, and kurtosis equal to κ. The variables σ_t, which are the variances of price changes, obey a stochastic process that can be forecasted. The ε_{t+1} are contemporaneously independent of the σ_t".

Then, Drost and Nijman (1993) created the weak GARCH concept to exploit the temporal aggregation properties of linear ARMA models. Where ARMA model or Autoregressive Moving Average model, is used to describe weakly stationary stochastic time series in terms of two polynomials. Petris (2009) give the definition that the first of these polynomials is for autoregression, the second for the moving average, so the model is referred to as the ARMA(p,q) model that:

$$x_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

p is the order of the autoregressive polynomial,

q is the order of the moving average polynomial.

Where:

φ = the autoregressive model's parameters,

θ = the moving average model's parameters.

c = a constant,

Σ = summation notation,

ε = error terms

where the SR–SARV(p) model is introduced.

As aimed by Engle (1995) estimating volatility models by time-series techniques via ARMA models for squared returns is generally very inefficient since these models feature "innovations sequences which not only have time-varying variances but have bounded support which differs over time". This may motivate a preliminary log-transformation of the conditional variance, as also proposed by Rosenberg (1972), before fitting an ARMA model. These we called the Exponential SARV model.

2.1.3 Asset Price Models

In this part, two main asset price model will be introduced. One is the Black-sholes model (BSM), and the other is the capital asset price model

(CAPM).

Peter (2009） proposed discrete-time models which include SV, ARCH, GARCH and their further generalizations have been developed to reflect the called stylized features of financial time series. These properties, which include tail heaviness, volatility clustering and serial dependence without correlation, cannot be captured with traditional linear time series models. Black and Scholes (1973) and Merton (1973) was celebrated of this work based on a geometric Brownian motion model for the asset price S(t) at time t. In 1965, Samuelson introcuced Brownian motion model according to which S (t) satisfies the Itô equation,

$$dS(t)= \mu S(t)dt + \sigma S(t)dW(t) \text{ with } S (0) > 0$$

This equation we define that $\{W(t) ,t \geq 0 \}$ is standard Brownian motion defined on a complete probability space (Ω, F ,P) with filtration {Ft} where F t is the sub-σ-algebra of F generated by $\{ W (s) , 0 \leq s \leq t \}$ and the null sets of F . The solution above becomes:

$$S(t)= S(0) \exp[(\mu - \sigma^2/2) t + \sigma W(t)]$$

so that the log asset price in this model is Brownian motion and the log return over the time-interval (t, t + Δ) become:

$$log\frac{S(t+\Delta)}{S(t)}=\left(\mu-\frac{1}{2}\sigma^2\right)\Delta+\sigma(W(t+\Delta)-W(t))$$

Here we will give the definition and equations for the BSM,

The thesis assumes that the capital structure of the firm is comprised of equity and by a zero-coupon bond with maturity T and face value of D, whose values at time t can be denoted by $E_{\_t}$ and $z_{(t,T)}$ respectively, for $0 \leq t \leq T$. The asset value of the firm is the sum of the equity and debt values. The firm will default if its assets cannot cover the promised payment of debt. Therefore the credit risk can be measured as the spread between the value of the firm's assets and its debt. The value of the firm's assets, V_t, is assumed to follow a geometric Brownian motion under the following measure

$$\mathrm{d}V_t = \gamma V_t dt + \sigma V_t dW_t$$

where γ is the drift rate, σ is the asset volatility, and W is a Brownian motion.

Under the settings, the payoff of the bondholders will be given by min (D, Vt) and the payoff of the shareholders will be given by max, applying the Black-Scholes pricing formula, the value of the equity at time t is given by

$$E_t(V_t, \sigma V, T-t) = e^{-r(T-t)}[e^{r(T-t)} V_t \emptyset(d_1) - D\emptyset(d_2)]$$

where ∅ is the distribution function of a standard normal random variable. d_1 and d_2 can be expressed as follows:

$$d_1 = \frac{\ln\left(\frac{e^{r(T-t)} V_t}{D}\right) + \frac{\sigma_V^2}{2}(T-t)}{\sigma_V \sqrt{T-t}}$$

$$d_2 = d_1 - \sigma_V \sqrt{T-t}$$

This formula represents that the value of equity at time t is given by the value of a risk-free bond minus the value of a put written to equity. The probability of default at time T can be expressed as

$$\mathrm{P}[V_T < D] = \emptyset(-d_2)$$

Over the last several decades, a significant volume of the literature aims to explain the “credit spread puzzle” since the credit spreads that Merton’s model produced are far smaller than estimates of credit spreads derived from actual, traded corporate bonds (Culp, Nozawa et al. 2014).

The advantage of the structural models is that the models provide a way of relating the credit risk of a firm to its capital structure: assets and liabilities. The key idea is that if the value of the firm’s assets goes below a given safety level, the firm is not able to repay its debts. Then the firm is subject to default. The default is induced by observable market information. But the models are based on several assumptions which are far from realistic. The most important restriction of the model is that the default time is fixed, the default can only happen on the maturity of the debt. Secondly, Merton’s model assumes the firms only issue a single zero-coupon bond and the firm value is tradeable. The usual capital structure of a firm is much more complicated than a simple zero-coupon bond. Geske (1979) and Geske and Johnson (1984) extended Merton’s model on compound options which have different maturities. They also considered sinking funds, safety covenants, payout restrictions and subordinated debt. Thirdly, another handicap of the model is that it assumes a constant and flat term structure of interest rates, like Black and Cox (1976), Leland (1994) and Leland and Toft (1996). Jones, Mason et al. (1984) proposed the stochastic interest rates, as well as taxes, would improve the performance of the model. In addition, many works in the literature also considered interest rates as stochastic processes, for instance, Ronn and Verma (1986), Kim, Ramaswamy et al. (1988), Briys and De Varenne (1997) and Saa-Requejo and Santa-Clara (1997). Nielsen, Saà-Requejo et al. (2001) and Longstaff and Schwartz (1995) also considered that the interest rate follows a Vasicek process.

The capital asset pricing model (CAPM) was introduced by The CAPM builds on the portfolio choice model developed by Harry Markowitz (1959). And William Sharpe (1964) and John Lintner (1965) give a further definition of CAPM. The CAPM is often used to measure the performance of mutual funds and other managed portfolios. Jensen (1968) estimated the CAPM time-series regression for a portfolio and used the intercept (Jensen's alpha) to measure abnormal performance. Eugene and Kenneth (2004) suggest that CAPM offers robust and intuitively pleasing predictions about measuring risk and explaining the relation between expected return and risk.

For instance, the CAPM suggests that the risk of a stock should be measured relative to a comprehensive "market portfolio" that in principle can include not just traded financial assets but also consumer durables, real estate, and human capital. Harry Markowitz (1959) creates the portfolio choice that lets the CAPM model build. In Markowitz's model, an investor selects a portfolio at time t-1 and produces a stochastic return at t. The model assumes investors are risk-averse and, when choosing among portfolios, they care only about the mean and variance of their one-period investment return. Markowitz's approach is often called a “mean-variance model.” This is because, as a result, investors choose “mean-variance-efficient” portfolios, in the sense that the portfolios have two parts; one is to minimise the variance of portfolio return, given expected return. The other is to maximise expected return, given conflict. The portfolio model provides an algebraic condition on asset weights in mean-variance efficient portfolios. The CAPM turns this algebraic statement into a testable prediction about the relation between risk and expected return by identifying a portfolio that must be efficient if asset prices are to clear the market of all assets.

There are two assumptions added by Sharpe (1964) and Lintner (1965) into the Markowitz model to identify a portfolio that must be mean-variance-efficient.

At first, when given market-clearing asset prices at t-1, investors agree on the joint distribution of asset returns from t-1 to t. And this distribution is the true one—that is, it is the distribution from which the returns we use to test the model are drawn.

Secondly, borrowing and lending at a risk-free rate is the same for all investors and does not depend on the amount borrowed or lent.

In summary, CAPM assumptions show that the market portfolio market (M) must be on the minimum variance frontier if the asset market is too transparent.

In conclusion, if there are N risky assets, $E(R_i)$ is the expected return on asset i, and the β_{iM} is the market beta of asset i. where we define that

$$\beta_{iM} = \frac{COV(R_i, R_M)}{\sigma^2(R_M)}$$

Under this condition, we summarise that the equation of the CAMP model would be:

$$E(R_i) = E(R_{ZM}) + [E(R_M) - E(R_{ZM})]\beta_{iM}, i = 1, \dots \dots N$$

2.1.4 Historical volatility and realized volatility

In time series, we normally define the volatility as a percentage and interpreted as standard deviation of returns, measures how much a security moves over a certain period. There are different type of volatility in time series such as historical volatility and realized volatility.

Realized volatility is the assessment of variation in returns for an investment product by analysing its historical returns within a defined time period. Assessment of degree of uncertainty and/or potential financial loss/gain from investing in a firm may be measured using variability/ volatility in stock prices of the entity. In statistics, the most common measure to determine variability is by measuring the standard deviation, for example the variability of returns from the mean. It is an indicator of the actual price risk.

The realized volatility or actual volatility in the market is caused by two components a continuous volatility component and a jump component, which influence the stock prices. Continuous volatility in a stock market is affected by intra-day trading volumes. For example, a single high volume trade transaction can introduce a significant variation in the price of an instrument.

Historic Volatility is the standard deviation of the "price returns" over a certain period, this can be hourly, daily, monthly, or other time periods. A "price return" is the natural logarithm of the percentage price changes or we can use the formula: $\ln [P_t / P_{(t-1)}]$.

A volatile market, therefore, has a larger standard deviation and thus a higher historical volatility value. Conversely, a market with small fluctuations has a small standard deviation and a low historical volatility value.

Historic volatility can also be used as a tool by traders who are trading only the underlying instrument. Quantifying the volatility in a market can affect a trader's perception of how far the market can move and thus provides some help in making price projections and placing orders. High volatility may indicate a trend reversal as heavy buying/selling comes into the market and may sharp price reversals.

2.2. Value at Risk (VaR)

Value at Risk (VaR) is a measure of the risk of loss for investments. It estimates how much a given set of investments might lose under a certain probability, given normal market conditions within a set time period such as an hour, a day, or a month. VaR is typically used by firms and regulators in the financial industry to measure the quantity of assets required to cover possible losses.

Philippe (2006) noted that, for a given portfolio, time horizon, and probability(P), the P VaR can be defined informally as the maximum possible loss during the time frame after excluding all worse outcomes whose combined probability is, at most, P. This assumes mark-to-market pricing and no trading within the portfolio. Setting up a portfolio of stocks that has a one-day 0.5% of $10 million means there is a probability of 0.005 that the value of the portfolio will fall by more than $10 million in a day, for example.

In calculating VaR, 0.5%, 1% and 5% probabilities for a fixed time interval (normally one day) are generally used as reference points.

2.3. Directional Change (DC)

The concept of directional change has been fully explored in the available literature. DC is a new way to summarize price changes (Guillaume et al

1997) based on the market movement alternating Uptrends and Downtrends over time. An uptrend terminates when a Downturn DC Event takes place; similarly, a downtrend terminates when an Upturn DC Event takes place. A Downturn (Upturn) DC Event is any event in which the price drops (rises) by a certain threshold (θ) from the highest (lowest) price (P_{EXT}) seen in the previous trend period.

As a Downturn DC Event defines the beginning of a new downtrend, at the end of such a Downturn DC Event, the price will have dropped by the specified threshold from the highest price in the last (and current) trend. That highest point (which would be the lowest point in the case of upturn directional change) is called an Extreme Point. An extreme point is only confirmed in hindsight, however, when DC is confirmed as the price changes by at least the threshold value from the previous extreme point.

A downtrend continues until the next upturn DC event is observed; this defines the lowest price in the current downtrend and starts the next uptrend. The price change from the end of the Downturn DC Event to the lowest price in the current trend is referred to as an Overshoot Event; each trend is thus comprised of both a DC Event and an Overshoot Event.

Tsang et al (2017) indicated that the Theoretical Directional Change Confirmation Point (DCC*) is the minimal or maximum directional change confirmation price for any upturn or downturn directional change event; however, this does not exist in practice in the real market. The reason for using DCC* rather than DCC is because, in the real world, the EXT point and DCC point can be the same point, depending on the threshold. The price of DCC*, however, is defined in the following ways:

In an uptrend: $P_{DCC\uparrow *} = P_{EXT} \times (1+ \theta) \leq P_{DCC\uparrow}$;

In a downtrend: $P_{DCC\downarrow *} = P_{EXT} \times (1- \theta) \geq P_{DCC\downarrow}$,

Where P_{EXT} is the price of directional change extreme point (EXT); P_{DCC} is the price of the directional change confirmation point (DCC); θ is the fixed threshold; and ↑and ↓ here represents the Upturn and Downturn events respectively. $P_{DCC\uparrow *}$ is thus the DCC* price of an upturn directional change event and $P_{DCC\downarrow *}$ is the DCC* price of a downturn directional change event.

Definitions of these factors and an outline of the core topic of directional change are required for the current work, and contrast with time series must be made. In time series, financial market prices are sampled using fixed-time intervals; the daily closing prices at a fixed interval are generally used to record the movement of prices. However, this snapshot of the market may not pick up all the necessary financial information. For example, according to Tsang et al (2016), if only the end of day financial prices had been recorded, the flash crash on 6th May 2010 (also known as the crash of 2:45 or simply the flash crash, was a United States trillion-dollar stock market crash, which started at 2:32 p.m. EDT and lasted for approximately 36 minutes) would not even have been noticed.

In contrast, instead of sampling prices at fixed time intervals as in time series, DC is entirely data-driven, using market price movements exceeding an observer-chosen threshold to dictate when a price must be recorded. (Tsang, et al 2017).

The concept of DC was introduced by Guillaume et al. (1997) as another way to sample data, with chosen sample points driven by data movement. The observer chooses the threshold for when to sample the market data based on their own determination of an appropriate threshold to display significant movement; this causes the market to be defined as a series of alternating uptrends and downtrends in which a change from downtrend to an uptrend occurs when the predefined threshold is reached in terms of market movement (Tsang et al, 2017).

The concept of DC was further developed by Glattfelder et al (2011), who introduced twelve new empirical scaling laws related to foreign exchange data series across thirteen currency exchange rates based on the theory of directional change. Kablan and Ng (2011) then examined capturing volatility using the event-driven approach of directional change within pre-specified thresholds, while Aloud et al (2012) demonstrated that the length of the price-curve coastline, as defined by directional change, showed the long coastline of price changes.

In contrast to time series analysis, in which researchers have developed indicators such as return and volatility to summarize market price changes, there remains a need to develop new indicators for the profiling of markets under DC, however, which forms part of the current work.

DC is a developing concept, and a range of DC indicators to profile various market dynamics allow new ways to examine the data; however, all of these are data-driven. The returns that a time series examines are investigated over a fixed time period, while the returns that DC examines are returns over a directional change event. Given the same period of data points, the DC coastlines are often much longer than time series coastlines for the same period, as DC deliberately captures all of the extreme points (Aloud et al, 2012).

Several studies have explored how to use the concept of DC to understand financial markets better. Depuis and Olsen (2011) examined how to use the DC concept in High-Frequency Trading models (HFT). Masry (2013) examined FX market activity based on the concept of DC; her approach laid the foundations for understanding how FX market activity changes as price movement progress, exploring the idea that small differences in market activity can change price trend conditions significantly. Bisig, Dupuis, Impagliazzo, and Olsen (2012) defined the concept of the Scale of Market Quakes (SMQ), based on the ideas of DC. This suggests that the FX market can be quantified by economic and political event declarations, and that, by analysing the average OS event, the magnitude of a quake can be calculated similarly to assessing the magnitude of an earthquake. The authors claim that because the SMQ arises in response to compelling market events, an analyst will observe larger SMQs when the market is in an unstable period and smaller SMQs during stable periods. This idea of SMQ shows that the measurement of an OS event can be used to quantify the price behaviour, that occurs in the financial market at periods of major economic and political events. This work led to Tsang et al (2016) developing a set of DC-based indicators for profiling the financial market. Glattfelder et al. (2011) also discovered 12 new scaling laws in foreign exchange markets using the DC approach, which they used to study stylised facts in FX markets.

Gillaume et al. (1997) proposed a new scaling law for DC, which they considered offered a new way to measure volatility and to give a description of the evolution of financial prices. Using DC to sample data helps ameliorate the problems described by Tsang (2017) that arise from using time series alone to summarize prices in the financial market, where prices are sampled only at fixed time intervals, with the final transaction price recorded as the daily closing price. However. Tsang (2017) also argued that time series and directional change are not competing methods of studying price dynamics. In his view, these approaches can complement each other and offer different perspectives. According to Tsang (2017), using both ensures that volatility observed under time series can be used alongside the observed frequency and price movement values observed under directional change: 'they all tell part of the story.". DC offers the flexibility that "By sampling different data points, DC sees price movement from an angle different from time series. Under time series, one fixes time (in the x-axis) and measures changes in price (in the y-axis). Under DC, one fixes the threshold in price change (in the y-axis) and lets the data determine when to sample the next extreme point, i.e., letting the data determine the next value on the x-axis. This also determines the time at which the next data point is sampled" (Tsang, 2017).

The continuing rise of global 24-hour financial markets necessitates the use of more flexible systems that can recognise market volatility, fluctuations, and dynamics rapidly and effectively. Directional change offers a different perspective on market dynamics, as it allows for transaction prices to be recorded only when there is a significant change in the price, offering a more continuous view of market fluctuations.

Tsang et al (2017) introduced further refinements to the definition of Directional Change. They defined the Directional Change Extreme Point (EXT) as the starting point for a DC trend, which may be either an Upturn Point or Downturn Point. It can be also seen as the end of a TM event where TM events consist of a total price movement, either a downturn event followed by a downward overshoot event, or an upturn event followed by an upward overshoot event (Glattfelder et al, 2011). The Directional Change Confirmation Point (DCC) is the point at which it is possible to confirm a DC event, while the Theoretical Directional Change Confirmation Point (DCC*) is the minimal or maximum directional change confirmation price for any upturn or downturn directional change event.

Overshoot refers to the price change from the last directional change confirmation price (DCC) to the current price. Tsang et al (2017) used Overshoot Value (OSV) to measure the extent of an overshoot, and thus, instead of using the absolute value of the price change as in time series, the value of OSV is relative to the threshold chosen by the observer. OSV is thus defined as follows:

$OSV= ((Pc - P_{DCC}) \div P_{DCC}) \div \theta$

where Pc is the current price, P_{DCC} is the last directional change confirmation price, and θ is the threshold.

Overshoot values at extreme points (OSV_{EXT}) act as an indicator for measuring the value of an overshoot based on the price distance between fixed points. They measure how far the overshoot exceeds the theoretical directional change confirmation point (DCC*) before reaching the next extreme point (EXT). We define OSV_{EXT} is defined as:

$$OSV_{EXT} = ((P_{EXT} - P_{DCC*}) \div P_{DCC*}) \div \theta$$

Where P_{EXT} is the price at the extreme point that ends the current trend, P_{DCC*} is the price of the theoretical directional change confirmation point of the current trend; and θ is the threshold.

Glattfelder et al. (2011), discovered 12 new scaling laws in foreign exchange markets, which were established using the DC approach, which was used to study stylised facts in FX markets. Gillaume et al. (1997) had proposed a new scaling law for DC, to be considered as a new way to measure volatility and the description of the evolution of financial prices.

Directional Change Indicators

Tsang et al (2017) introduced the idea that, in DC, different indicators may be used to measure different kinds of volatility. To track market changes, both Directional Change and Time Series have specific indicators to measure the volatility of the market. In time series, the log return of standard deviation for a fixed time interval is used as a measure of volatility for that market period, while in DC, several indicators may be used. These include Number of Directional Changes (NDC), Total Price Movement (TMV), Time for completion of a trend (T), Time independent Coastline (C_{DC}), and Time-adjusted return of DC (R_{DC}) under Directional Change. All of these indicators offer means of explaining different features of volatility tracking in the financial market, and defining them can build up a “vocabulary” for new concepts of tracking using DC

2.3.1. Number of Directional Change (NDC)

The Number of Directional Changes (NDC) is an indicator that represents how many Directional Changes occur in a given period; thus, NDC is the total number of DC events that happen over a profiled period. This measures the frequency of DC events; based on the same threshold, a time period with a higher NDC value is more volatile than a time period of the same length with a lower NDC. By recording the NDC within the profiled period, DC offers an additional way to measure the volatility of market price movements.

For example, if two profiles cover time periods of the same length, then the comparative number of directional changes within these gives a measure of comparative volatility. The Number of Directional Changes (NDC) is inversely proportional to T, and a higher NDC value suggests a higher frequency of directional changes in the market, indicating higher volatility more generally.

Figure 2. 1 shows the DC summary of two hypothetical markets, A and B. The x-axis represents the data period (T), and the y-axis represent the price. In this figure, the price moves between the same range in the two markets. However, we can see that twelve trends happened in market A while six trends happened in market B in the same period. Therefore, market A is more volatile than market B.

2.3.2. Total Price Movement (TMV)

Total Price Movement (TMV) can also be used to measure volatility under DC by indicating any big changes in prices.

Definition 2.1: Total Price Movement (TMV)

> In a trend that starts at the price P_{EXT}, the TMV of a price P is the price change rate from P_{EXT} to P, normalized by θ, where θ is the threshold used to generate the DC events[1].

[1] See Section 2.3 for introduction of DC thresholds.

$$TMV_P = (P - P_{EXT}) / P_{EXT} / \theta \tag{2.1}$$

In the TMV definition, price change rates are normalized by the DC threshold θ. This allows us to compare TMV obtained under different thresholds.

The extreme points are special transactions in DC. As we refer to them frequently, we denote the TMV at extreme points TMV_{EXT} (Tsang et al 2017). It is defined below for the sake of completeness.

Definition 2.2: Total Price Movement of an extreme point (TMV_{EXT})

In a trend that starts at the price P_{EXT_i}, and ends with price P_{EXT_i+1}, the TMV of P_{EXT_i+1} is the price change rate from P_{EXT_i} to P_{EXT_i+1}, normalized by the DC threshold θ.

$$TMV_{P\_EXT\_i+1} = (P_{EXT\_i+1} - P_{EXT\_i}) / P_{EXT\_i+1} / \theta \tag{2.2}$$

TMV defined above are signed. They are positive in uptrends (when P_{EXT_i+1} is greater than P_{EXT_i}) and negative in downtrends (when P_{EXT_i+1} is greater than P_{EXT_i}). On many occasions, we are not interested in the sign of TMV. We are only interested in the magnitude of price changes. Therefore, we introduce the term aTMV to refer to unsigned TMV.

Definition 2.3: Absolute Total Price Movement (aTMV)

In a trend that starts at the price P_{EXT}, the TMV of a price P is the *absolute* price change rate from P_{EXT} to P, normalized by θ, where θ is the threshold used to generate the DC events.

$$TMV_P = |P - P_{EXT}| / P_{EXT} / \theta \tag{2.3}$$

Figure 2. 1 shows the DC summary of two hypothetical markets, B and C as well. The x-axis represents the data period (T), and the y-axis represent the price. In this figure, the time moves between the same range in the two markets. However, we can see that in markets B and C we spend the same time to finish a trend, while the TMV in B is twice than TMV in C in the same time period. Therefore, market B is more volatile than market C.

NDC and TMV provide two different perspectives to the volatility of a market: while the NDC measures the frequency of directional change, TMV measures the magnitude of price changes in each trend. All else being equal, the higher the NDC, the more volatile the market is. All else being equal, the greater the magnitude of price changes per trend, the more volatile the market is. To summarize, TMV is one indicator that measures the volatility in a trend. It can also be used to measure the volatility of a market over a certain period.

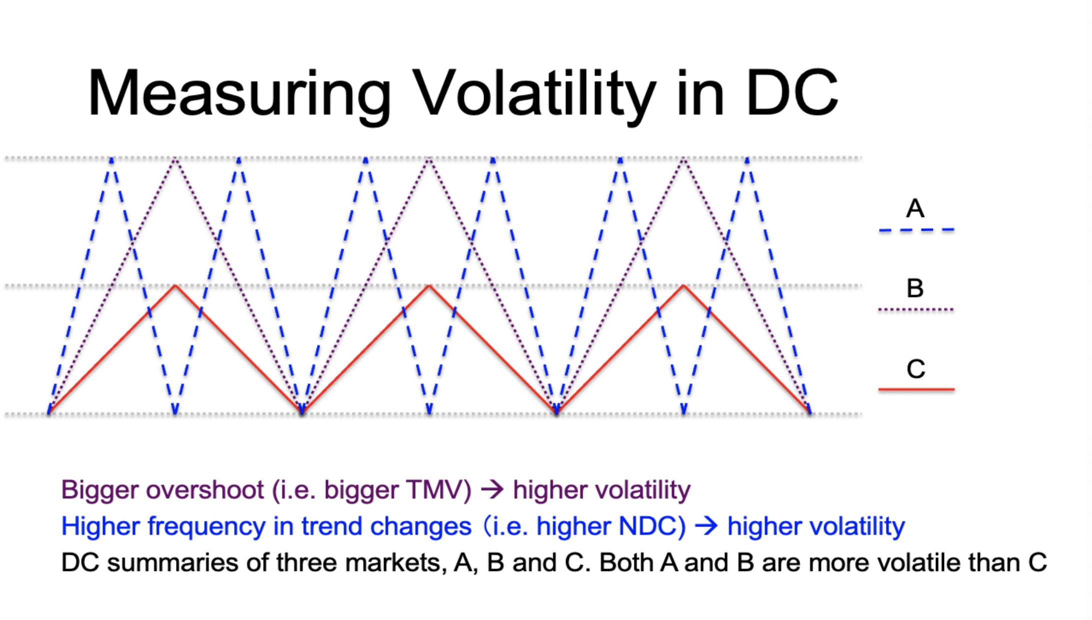

Figure 2. 1: Measuring Volatility in DC (Tsang 2017)

2.3.3. Time for completion of a trend (T)

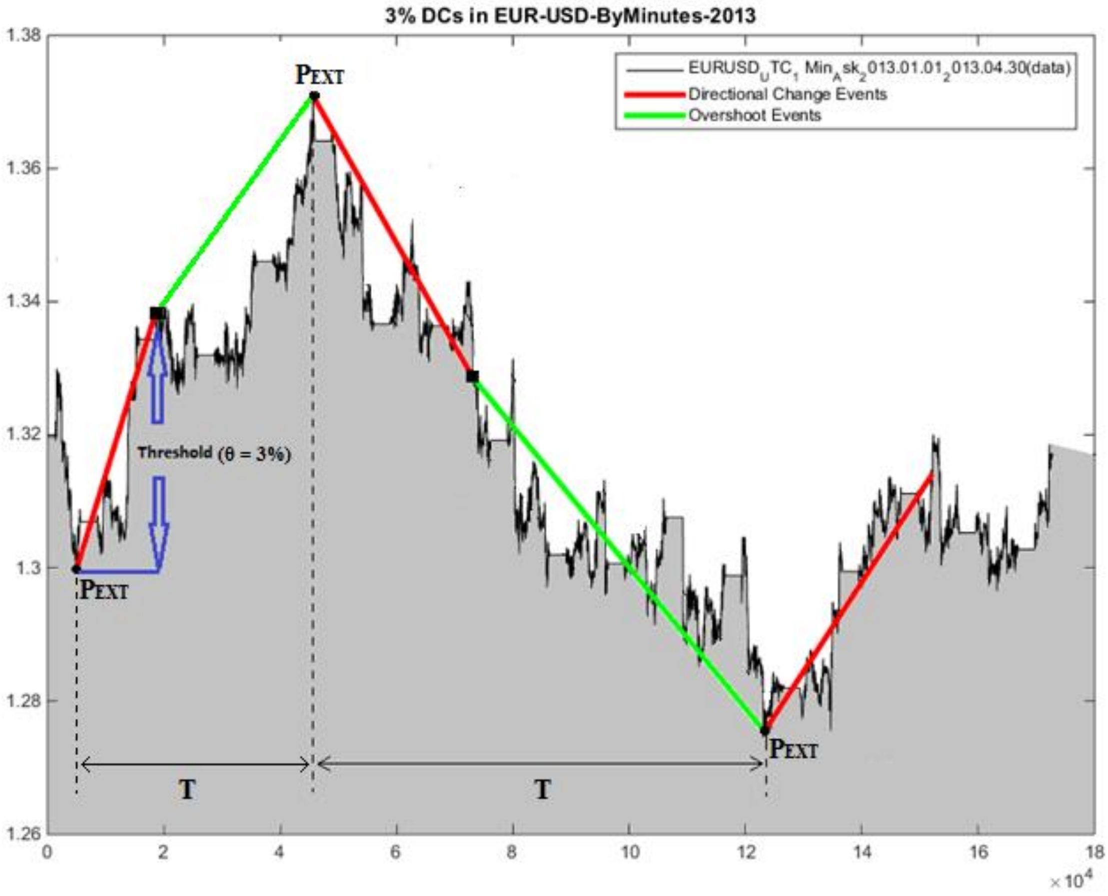

Figure 2. 2 3% DCs in EUR-USD-ByMinutes-2013 (Tao, 2017)

Within Figure 2. 2 when θ is equal to 3%, P_{EXT} is the price at directional change extreme point (solid black squares), and T is the time that it takes between two consecutive directional change extreme points.

Tsang et al (2017) pointed out that DC is defined based on events, thus utilising intrinsic as opposed to physical time (Glattfelder et al, 2011). However, this does not mean that it ignores physical time. The amount of physical time that a trend takes to complete is a significant piece of information, and the indicator T is used to represent the physical time between the extreme points that begin and end a trend.

2.3.4. Scaling laws under DC

Glattfelder and Olsen (2011) discovered 12 independent new empirical scaling laws in foreign exchange data-series under DC that hold for close to three orders of magnitude and across 13 currency exchange rates. Their statistical analysis crucially depends on an event-based approach that measures the relationship between different types of events.

At first, all 12 scaling laws are under the test of the price data of the foreign exchange (FX) market. Glattfelder and Olsen (2011) suggest that in financial markets, the flow of time is discontinuous: over weekends trading comes to a standstill or, inversely, at news announcements, there are spurts of market activity. In law (0a), the confinement of analysing returns as observed in physical time is overly restrictive. Law (0b) is a first attempt at establishing a new paradigm by looking beyond such constraints within financial data, constituting an event-driven approach, where patterns emerge for successions of events at different magnitudes. This alternative approach defines an activity-based time-scale called intrinsic time.

Extending this event-driven paradigm further enables us to observe new, stable patterns of scaling and reduces the level of complexity of real-world time series. In detail, the fixed event thresholds of different sizes define focal points, blurring out irrelevant details of the price evolution. The Figure below depicts how the price curve is dissected into so-called directional-change and overshoot sections.

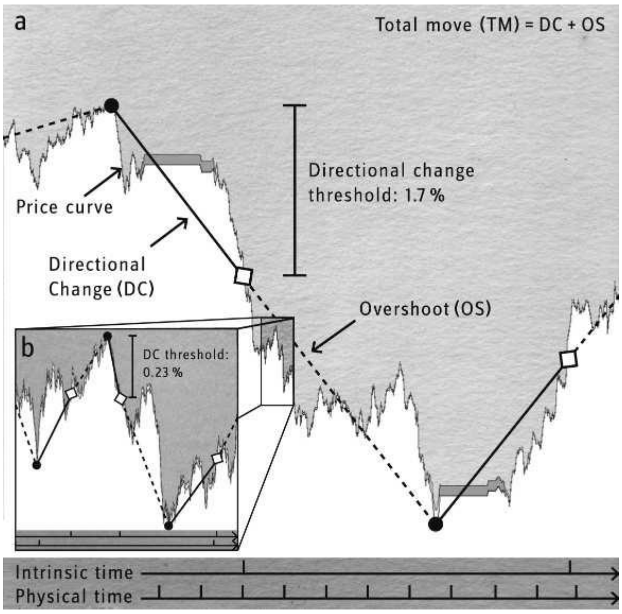

Figure 2. 3 Projection of a (a) two-week, (b) zoomed-in 36 hour price sample onto a re- duced set of so-called directional-change events defined by a threshold (a) Δxdc = 1.7%, (b) Δxdc = 0.23% (Glattfelder and Olsen (2011))

In Figure 2. 3 These directional-change events (diamonds) act as natural dissection points, decomposing a total-price move between two extremal price levels (bullets) into so-called directional-change (solid lines) and overshoot (dashed lines) sections. Glattfelder and Olsen (2011) confirm that time scales depict physical time ticking evenly across different price-curve activity regimes, whereas intrinsic time triggers only at directional-change events, independent of the notion of physical time.

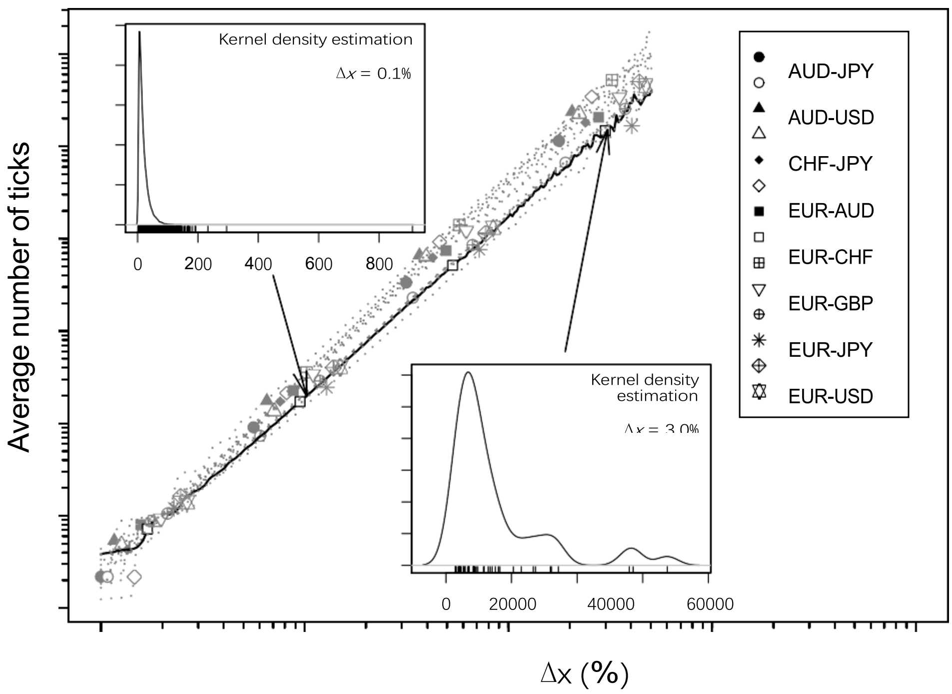

Figure 2. 4 Scaling law (1) is plotted where the x-axis shows the price move thresholds of the observations and the y-axis the average tick numbers. (Glattfelder and Olsen (2011))

Figure 2. 4 shows a tick is defined as a price move of 0.02%. Glattfelder and Olsen (2011) point out what the solid line shows the raw data for EUR-USD. For the remaining 12 currency pairs and the Gaussian random walk benchmark model the raw data is displayed with dots. Insets show the distribution of the EUR-USD observations (drawn above their x-axis) for selected threshold values of 0.1% and 3.0%.

Glattfelder and Olsen (2011) summarized that laws (4) and (5) relate the average numbers of seconds that elapse between consecutive price moves or directional changes, respectively.

Glattfelder and Olsen (2011) also unveil a set of scaling laws emerging from the identification of directional-change events (see figure 2.3) that make up the so-called total-move (TM) segments, which themselves decompose into directional-change (DC) and overshoot (OS) parts. The total price move, waiting time, and the number of ticks can then be written as:

$$\langle|\Delta x^{tm}|\rangle = \langle\left|\Delta x^{dc}\right|\rangle + \langle|\Delta x^{os}|\rangle \quad (6)$$

$$\langle\Delta t^{tm}\rangle = \langle\Delta t^{dc}\rangle + \langle\Delta t^{os}\rangle \quad (7)$$

$$\langle N(\Delta x_{tck}^{tm}\rangle = \langle N(\Delta x_{tck}^{dc}\rangle + \langle N(\Delta x_{tck}^{os}\rangle \quad (8)$$

Glattfelder and Olsen (2011) also propose the decomposition leads to nine additional scaling laws, where the average values are functions of the directional-change thresholds Δx_{dc}

$$\langle|\Delta x^{*}|\rangle = (\frac{\Delta x_{dc}}{c_{x,*}})^{E_{x,*}} \quad (9)$$

$$\langle\Delta x^{*}\rangle = (\frac{\Delta x_{dc}}{c_{t,*}})^{E_{t,*}} \quad (10)$$

$$\langle N(\Delta x_{tck}^{*}\rangle = (\frac{\Delta x_{dc}}{c_{N,*}})^{E_{N,*}} (11)$$

Where * stands for {tm, dc, os} Note that $\left|\Delta x^{dc}\right| = \Delta x_{dc}$ holds by construction.

Glattfelder and Olsen (2011) also considering cumulative price moves instead of the averages in laws (9) leads to another triplet of laws:

$$\Delta x_{cum}^{*} = \sum_{i=1}^{n}|\Delta x_{i}^{*}| = (\frac{\Delta x_{dc}}{C_{cum,*}})^{E_{cum,*}} \quad (12)$$

This concludes the presentation of 12 new scaling laws.

2.4. Profiling

Profiling is an application of directional changes to turn raw data into information. In order to track market movement in the market, it is necessary to capture the similarities and differences in prices movements within and between periods, based on observing the stylised behaviours of financial series. By using directional change theory instead of sampling at fixed intervals, price changes dictate when a price is recorded. DC thus provides a complementary way to extract more information from data and to observe features that may not be recognised in time series. For this purpose, a computer program that includes some indicators to help analyse market movements is required. These indicators can then be used to measure the stylized behaviours of price movements under directional changes. The outcomes of the program are independent profiles, with each profile illustrating the performance within certain periods of the data under the selected threshold. Comparing and analysing the similarities and differences between the various metrics of these profiles may thus help with tracking and predicting abnormal events in the market.

2.5. Nowcasting

Nowcasting is a concept that was introduced by Giannone, Reichlin, and Small (2008). It is defined as any narrow prediction of only the present, the very near future, and the very recent past in a financial market. Castle et. al (2013) further noted that “forecasts are made before a period commenced, with ‘Nowcasts’ during the relevant period, and with flash estimates immediately, or shortly after the period ends, when disaggregated information remains incomplete. For instance, we do not necessarily know what is happening now, very often we only know what is happening in hindsight”.

Even where an uptrend may have ended, an observer cannot know that a downtrend has started, until the next Directional Change confirmation point, which occurs only when the price has dropped from the last peak by the specified threshold.

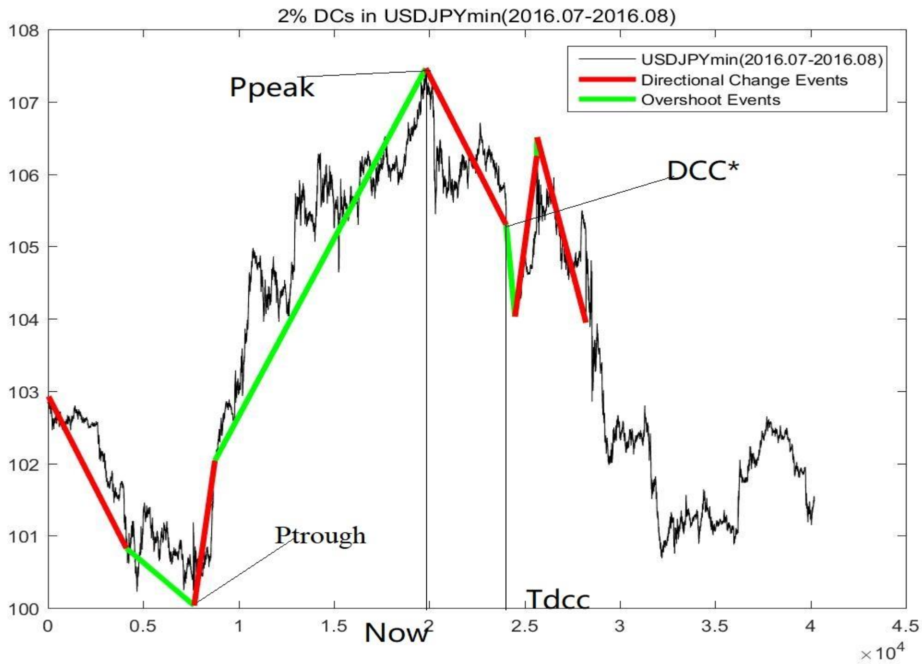

Figure 2. 5 Directional Changes in foreign exchange rates between US Dollars and Japanese Yen (USD/JPY) with threshold = 2% (E.P.K. Tsang 2017)

Figure 2. 5 shows an example of Nowcasting using DC. In this figure, the X-axis represent the time, while the Y-axis represents the price. At the point “Now”, a new trend is started. However, it is only possible to confirm that a new trend has started at time T_{dcc} when the price reaches DCC*. “Nowcasting” refers to the idea of recognising the start of a new trend after Now, before T_{dcc}. The closer this point of recognition can be brought to “Now”, the more useful this nowcast will be. This nowcasting problem will be tackled in this thesis (see Chapter 5).

Bakhach (2018) noted that many studies have concluded that the directional change (DC) framework is useful in analysing FX markets. In his research, he considered the problem of forecasting the change of a trend’s direction from a DC perspective, based on the task of forecasting whether the current trend, whether that is an uptrend or a downtrend, will continue in the same direction for a specific percentage before the occurrence of the next extreme point. He thus introduced an original DC-based independent variable and proved its usefulness in the proposed forecasting problem, as well as addressing the problem of forecasting the change of a trend’s direction within the DC framework.

Bakhach (2018) then formulated a means of predicting the change of direction of a market’s trend under the DC framework. He proposed tracking price movements using two concurrent DC thresholds, STheta and

BTheta. He then attempted to forecast whether a DC trend, as observed under threshold STheta would continue so that its total magnitude would be at least equal to BTheta. To do this, he introduced a new concept named Big-Theta, originating from the DC framework. The idea of Big-Theta is that a DC event of threshold BTheta will embrace at least one DC event of the smaller threshold STheta (with BTheta > STheta). The concept of Big-Theta was thus used to introduce the Boolean variable BBTheta, the value of which expresses whether the total price change of a DC trend, as observed under the threshold STheta will reach BTheta.

Bakhach (2018) also examined the performance of this forecasting approach using eight currency pairs sampled minute-by-minute. The results demonstrated an accuracy of between 62% and 80%.

To the best of the authors' knowledge, this was the first attempt to forecast the change of a trend's direction under the DC framework.

3. Chapter Three: Useable range for Threshold

3.1. Background

Firstly, thresholds constitute a vital part of DC's make-up. As DC is entirely data-driven, it records market movements as they happen, meaning that the observer has to choose what threshold will trigger the significant market movement. Therefore, in this chapter we argue for the utility of DC, and that this utility depends on the choice of threshold. Moreover, DC's data-driven nature makes it particularly suitable for observing the data-heavy, 24-hour global financial markets. However, the question then arises, what range of threshold should one choose for observing the trend in DC? This chapter attempts to answer that question and to provide a framework for making useable range thresholds choices.

3.2. The objective of this chapter

An observer may choose to use any threshold to observe the market. However, if the aim is to collect statistical measures to profile a market over a period, as proposed by Tsang et al. (2017), then choosing the thresholds requires due consideration. A threshold that is too big will not produce enough observable directional changes for statistical purposes. Equally, if a threshold is too small, any small shift in the opposite direction will be seen as a directional change for the observer. This introduces the problem of ‘noise’ into the statistical measures, as will be elaborated below.

The objective of this research is to develop measures to determine when the threshold is too big or too small. However, the objective is not identifying what constitutes an ‘optimal’ threshold for profiling a market period, as it is unlikely that a one-size-fits-all approach would be appropriate for this research. Instead, the objective is to determine the range of usable thresholds for profiling. This chapter proposes a method which will allow the data to show what thresholds are too big or too small when profiling a market period.

3.3. Methodology

This section proposes guidelines to determine the range of useable thresholds for a given data set.

3.3.1. Overview of the methodology

To decide what thresholds are usable for profiling a data set, this section summaries the data as a sequence of trends in DC, with a range of thresholds. The mean, maximum, and minimum aTMVs of the trends are all computed.

Based on Guillaume et al. (1997), the minimum aTMV observed should be close to 1 (to be explained below). Therefore, this study proposes to use the minimum aTMV, as an indicator of whether the threshold used is too small.

As explained in Section 2.3.2, aTMV are normalized by the threshold. Thus, aTMV is threshold independent. Therefore, if a sharp change is observed in the mean aTMV when the threshold is incremented, it signals the danger of using too few trends for profiling. This is the sign of using a threshold that is too big.

3.3.2. When is the threshold too small

As we explained in Section 2, an uptrend is confirmed when there exists a price $P_{DCC\uparrow}$ such that:

$$P_{DCC\uparrow} \geq P_{trough} \times (1 + \theta).$$

The minimum price for an uptrend to be confirmed is $P_{DCC\uparrow *}$, where:

$$P_{DCC\uparrow *} = P_{trough} \times (1 + \theta)$$

Similarly, the maximum price for a downtrend confirmation is $P_{DCC\downarrow *}$:

$$P_{DCC\downarrow *} = P_{peak} \times (1 - \theta)$$

According to Tsang et al. (2017), the Overshoot Value (OSV) at an upward DC confirmation point $DCC_{\uparrow}$ is as:

$$OSV_{DCC\uparrow} = ((P_{DCC\uparrow} - P_{DCC\uparrow *}) \div P_{DCC\uparrow *}) \div \theta$$

$P_{DCC\uparrow}$ is normally close to $P_{DCC\uparrow*}$ in practice. If that is the case, then $OSV_{DCC\uparrow}$ is close to 0. However, if we use a very small threshold, $P_{DCC\uparrow}$ could be significantly larger than $P_{DCC\uparrow*}$. To take an extreme example, suppose price normally moves by steps of 0.001% in a particular market. If a threshold of 0.0001% is used, then $OSV_{DCC\uparrow}$ could become extremely big:

$$OSV_{DCC\uparrow} = 0.001\% \div 0.0001\% = 10$$

If the trend reverses immediately at $DCC_{\uparrow}$, the aTMV of this trend is equal to 10. Empirical studies (Guillaume et al., 1997) shows that this is a very big aTMV. However, this big aTMV is only observed because of the choice of an unreasonably small threshold for this market.

Guillaume et al, (1997) showed that, regardless of the threshold used, aTMVs follow a power-law decade: many trends reverse immediately after DCC. As explained above, aTMVs of these trends are close to 1. Therefore, we expect the minimum aTMV is close to 1. Should the minimum aTMV observed in a market-period be significantly greater than 1, the reason is likely to be caused by the fact that we have chosen a threshold that is too small. This forms the basis of our first guideline for establishing a threshold.

Guideline 1： A threshold should be rejected for being too small for profiling if the minimum aTMV in the market-period is significantly greater than 1.

3.3.3. When is the threshold too big?

According to Olsen's observation (Guillaume et al., 1997), markets exhibit a fractal phenomenon under DC. This means we observe similar profiles under different thresholds. The TMV definition above is normalised by the threshold used. Therefore, according to Olsen's observation, we should observe similar TMV values under different thresholds. If the mean aTMV changes dramatically as we incrementally increase the threshold, then it is likely that the new profile offers too few (dissimilar) trends. In other words, the new threshold used is too big. This is the basis of our second guideline.

Guideline 2: A threshold should be rejected for being too big for profiling if the mean aTMV increases dramatically when the threshold is increased slightly.

3.4. Example One: EUR/USD Period One

3.4.1. Data selected and range of thresholds to examine

This section uses tick data in the EUR/USD exchange market from 21/04/2016 07:31:08 to 06/05/2016 14:31:21.

This data set was named EUR/USD Period One.

To decide what thresholds are useable, we summarized the data in section 3.4.1, by using DC with a range of thresholds. We use 28 thresholds in this chapter. The thresholds and their indices are shown in Table 3. 1.

Index(n)/ Threshold	1	2	3	4	5	6	7
	0.00001	0.00002	0.00003	0.00004	0.00005	0.00006	0.00007
	8	9	10	11	12	13	14
	0.00008	0.00009	0.0001	0.0002	0.0003	0.0004	0.0005
	15	16	17	18	19	20	21
	0.0006	0.0007	0.0008	0.0009	0.001	0.002	0.003
	22	23	24	25	26	27	28
	0.004	0005	0.006	0.007	0.008	0.009	0.01

Table 3. 1 Index of (n) and Threshold (Th) used for in this paper

3.4.2. Result for Period One

Index	Threshold	Mean aTMV	Min aTMV
1	0.00001	4.6271777	1.258416
2	0.00002	4.0133908	1.258447
3	0.00003	3.63958239	1.258495
4	0.00004	3.4211865	1.258384
5	0.00005	2.7369492	1.006707
6	0.00006	2.73067158	1.049274
7	0.00007	2.72268926	1.078696
8	0.00008	2.71337584	1.10412881
9	0.00009	2.51458287	1.000023
10	0.0001	2.43097167	1.00913265
11	0.0002	2.22847085	1.00743
12	0.0003	2.1003618	1.008547
13	0.0004	2.02578999	1.000594
14	0.0005	1.96533547	1.00009
15	0.0006	1.9601325	1.000324
16	0.0007	1.91670702	1.000005
17	0.0008	1.93278493	1.000065
18	0.0009	1.955967	1.001104
19	0.001	1.99373879	1.002411
20	0.002	2.03064767	1.032848
21	0.003	1.90877544	1.021156
22	0.004	2.0913975	1.031619
23	0.005	1.94717667	1.111224
24	0.006	1.96052713	1.055028
25	0.007	1.93929633	1.021187
26	0.008	2.44278567	1.336983
27	0.009	2.17136467	1.188429
28	0.01	1.95422833	1.069586

Table 3. 2 EUR/USD Period One’s mean aTMV and minimum aTMV under different thresholds.

3.4.3. When is a threshold too small for profiling Period One

As we mentioned in Section 3.2, from Table 3. 2, we can see that the value of minimum aTMV (the rightmost column) under thresholds from 0.007 to 0.00005 are all approximately 1. At Threshold=0.00005, the minimum TMV is 1.006707. However, when the threshold value drops to 0.00004, the value of the minimum aTMV rises sharply to 1.258384. In this case, we see any threshold below 0.00005 is unsuitable for profiling EUR/USD Period One.

3.4.4. When is a threshold too big for profiling Period One

Table 3. 2 shows that in EUR/USE Period One, when the threshold is between 0.0002 and 0.005, the mean aTMV values are approximately 2.0 (column 3). Once the threshold increased to 0.008, the value of mean aTMV sharply increase to 2.44278567. The mean aTMV increases to 2.44278567 under threshold 0.008. Under the circumstances, we define that any threshold larger than 0.008 is defined as being unsuitable for profiling EUR/USD Period One.

In conclusion, we defined that the usable range of thresholds for this data set is between 0.00005 and 0.008, both numbers included.

3.5. Example Two: EUR/USD Period Two

3.5.1. Data selected range of thresholds to examine

This section uses tick data in the EUR/USD exchange market from 29/03/2016 13:29:42 to 13/04/2016 10:19:18.

This data set was used EUR/USD Period Two.

The same threshold was used as the one introduced in Section 3.4

3.5.2. Results for Period Two

Index	Threshold	Mean aTMV	Min aTMV
1	0.00001	4.73054856	1.232119
2	0.00002	4.0040592	1.232013
3	0.00003	3.59096243	1.232544
4	0.00004	3.3717238	1.232438
5	0.00005	2.93279468	1.000000
6	0.00006	2.68517048	1.027297
7	0.00007	2.68297783	1.056219
8	0.00008	2.67478211	1.07841
9	0.00009	2.67152819	1.096014
10	0.0001	2.53199258	1.000000
11	0.0002	2.24322587	1.000075
12	0.0003	2.13125317	1.000125
13	0.0004	2.05193953	1.000138
14	0.0005	2.00232298	1.000713
15	0.0006	1.99328215	1.00144
16	0.0007	1.98420177	1.002104
17	0.0008	1.96438704	1.001462
18	0.0009	1.98484554	1.000595
19	0.001	1.97924251	1.000717
20	0.002	1.84791789	1.018877
21	0.003	2.01279488	1.01183
22	0.004	1.88247868	1.038123
23	0.005	1.75531236	1.036448
24	0.006	2.1202938	1.066875
25	0.007	2.419529	1.063466
26	0.008	2.953677	2.953677
27	0.009	2.625491	2.625491
28	0.01	2.362942	2.362942

Table 3. 3 EUR/USD Period Two's mean aTMV and minimum aTMV under different thresholds.

3.5.3. When is a threshold too small for profiling Period Two?

As we mentioned in Section 3.2, from Table 3. 3 we can see that under thresholds 0.007 to 0.00005, the value of minimum aTMV is around 1.

Under threshold 0.00005, the minimum aTMV is 1.0. When the threshold is dropped to 0.00004, the minimum aTMV increases sharply to 1.232438. In this case, we see any threshold below 0.00005 unsuitable for profiling EUR/USD Period Two.

3.5.4. When is a threshold too big for profiling Period Two?

Table 3. 3 shows, we can see that under thresholds 0.0004 to 0.003, the mean aTMV are all approximately 2.0. Under threshold 0.003, the mean TMV value is 2.01279488. Under threshold 0.004, the mean aTMV value decreases sharply to 1.88247868. When the threshold is increased to 0.005, the mean aTMV value decreases to 1.75531236. When the threshold value rises to 0.006, the mean aTMV jumps back to 2.1202938. In other words, the mean aTMV values fluctuate above threshold 0.03. Thus, we see any threshold larger than 0.003 as unsuitable for profiling EUR/USE Period Two.

In conclusion, we defined that the usable range of thresholds for this data set is between 0.00005 and 0.003, both numbers included.

3.6. Conclusion

It is up to the observer to choose an appropriate threshold to observe Directional Changes (DCs) in a given market-period. However, this study argues that if the aim is to use statistical information to profile a market-period, as in the use of DC, thresholds should not be chosen which are either too small or too big. However, such a choice requires careful consideration. This chapter proposed two guidelines to determine the range of usable thresholds for DC profiling.

When the threshold is too small, every transaction in the opposite direction constitutes a DC, and such profiles may capture a significant amount of noise. We shall show with an example in Chapter 5 the problem of using thresholds that are too small. On the other hand, when the threshold is too big, there will be too few trends, not enough noise, in the profile. Statistical measures thus collected are based on too few data points to be significant and cannot be effectively analysed.

It is important to stress that DC-based analysis is entirely data-driven. This means finding the range of useable thresholds is required for each individual data set. By proposing an effective guideline to determine what are the ranges of usable thresholds, in our view, this chapter lays an important foundation for new scientific, computer-based research, in the new area of DC profiling of financial markets.

4. Chapter Four: Tracking market movement by multiple indicators

4.1. Background

Financial markets change constantly, and these changes are normally tracked by different kinds of volatility. In Time Series (TS), volatility is typically defined as the standard deviation of log return over a certain time interval, which can be every minute, every hour, every day, or every month. The standard deviation of log returns can be referred to as SD, yet SD is not the only feasible measure of volatility. Directional Change (DC) introduces as an additional way to measure market volatility by different indicators such as NDC and aTMV, which, as explained in Section 2.3, measure two independent properties of volatility. By using multiple indicators, it is possible to understand markets more fully than when only using DC or TS. This chapter, therefore, presents new research on tracking market changes by combining SD, NDC, and aTMV, offering a new way to track the volatility of the market.

In this chapter, a new method proposed has been developed by combining DC and TS, and this chapter demonstrates that the proposed method allows researchers to effectively follow market tracking, facilitating a summarization of market changes.

4.2. Motivation and objective of this chapter

Every transaction is recorded in financial markets each second while it has happened, some seconds do not have transactions which we do not record. The key to market success is to extract information from that data. Prices change rapidly and knowing what indicators to look for and tracking change in values is necessary to make good decisions in these markets. As explained in Section 2.3, the DC indicators NDC and aTMV capture a great deal of the volatility information in the market; this chapter, therefore, demonstrates how this is achieved.

The approach taken is empirically based, and data-driven, concentrating on what market volatility is revealing through the use of DC and TS indicators to analyse market data. By using multiple indicators, the aim is to track volatility more reliably.

4.3. Overview of our approach

This chapter demonstrates tracking market movements by monitoring both TS and DC volatility indicators under specific rolling windows. The aim is to establish the usefulness of NDC and aTMV, and an explanation of the approach used to do this is therefore required.

This approach begins with the data, which is grouped into rolling windows. Within each window, the SD, NDC, and median $aTMV_{EXT}$ values are calculated. Following we use $aTMV_{EXT}$ to explain how we assess the volatility of a rolling window. The same analysis applies to the other two indicators, NDC and SD.

1. We calculated the median $aTMV_{EXT}$ for each rolling window.
2. We then calculated the standard deviation of the median $aTMV_{EXT}$ over all the rolling windows observed.
3. For example, suppose we observed a median $aTMV_{EXT}$ value of 2.31 (see Section 4.6.6.2 below). The value 2.31 is 5.46 times bigger than the mean of median $aTMV_{EXT}$ of the historical period observed.
4. We say that this median $aTMV_{EXT}$ value 2.31 is significant because:
   - 4.1. It is above 99.999% of the median $aTMV_{EXT}$ values observed in all windows.
   - 4.2. A value that is 5.46 times bigger than the mean has less than 0.00001% probability of happening.
5. Therefore, we use the word 'significant' in general terms and all the word 'significant' in this section means ‘historically significant’.

Within the ensuing analysis, the following questions are asked:

- Do all the indicators agree with each other all the time?

- If not, do they disagree with each other all the time?

- If neither of these cases applies, does any indicator extract particular volatility information in the market that the other indicators miss?

4.4. Methodology

4.4.1. Rolling windows and rolling speed

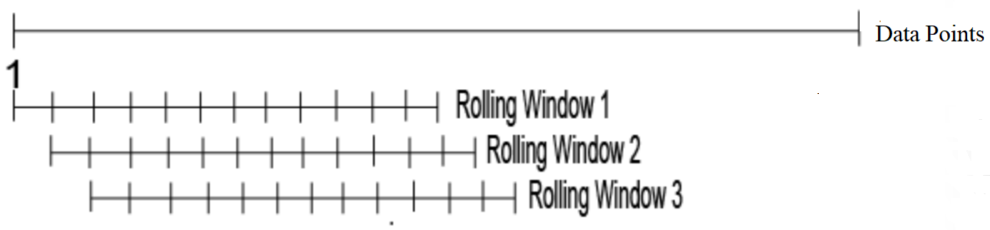

Figure 4. 1 Rolling windows used in the experiments

Figure 4. 1 shows an example of the rolling windows applied. A data set was selected from the market in a chosen period, and the various rolling windows (rolling window 1, rolling window 2, rolling window 3, and so on) were then superimposed to the end. Each window thus contains a certain number of data points under DC and a certain time period under TS. The gap between each rolling window represents the rolling speed.

In this research, rolling speed is therefore defined by a certain number of data points under DC and a certain time period under TS.

4.4.2. Volatility under TS

The standard deviation of log returns, referred to as SD, is used to measure that volatility which draws on the average price deviation by which a time period differs from the average over a period of time.

4.4.3. Volatility under DC

In this chapter, NDC and median $aTMV_{EXT}$ are the indicators used to measure the volatility of DC.

4.5. Experimentation

4.5.1. Data used

EUR/USD data, collected by the second (some seconds showed no trading), from 00:00:10 on 25/09/2009 to 07:13:14 on 18/07/2016 was used; this set contains 112,442,529 data points (transactions) across the period and both under TS and DC were applied to the same raw data, so these data points were the same. For clarity of reference, this period is labelled P.

4.5.2. Experimental setup

In TS, SD is used to measure the volatility of the market. For DC the indicators median $aTMV_{EXT}$ and NDC are used to measure the volatility of the market.

Threshold Setup:

We use the threshold= 0.0004 for all DC calculation

DC rolling windows setup:

In DC, the rolling windows are defined by the number of data points[2]. Each DC rolling window was therefore comprised of 1,000,000 data points, with NDC and median $aTMV_{EXT}$ calculated for this set. The rolling speed was 250,000 data points, as each rolling window started 250,000 data points after the beginning of the previous rolling window[3].

TS rolling windows setup:

For the time series (TS) theory, 16 days was set as the rolling window size, and four days was set as the rolling speed. For each rolling window, the hourly log return was calculated as the SD.

The reason for using 1,000,000 data points per window under DC and 16 days under TS was to make the TS window size and DC window size approximately the same physical duration. However, as DC and TS sample data at different points, the rolling windows observed under DC and TS are different.

[2] Each data point is a transaction, with transactions collated by the second. There may be no transactions in some seconds, and two adjacent transactions therefore do not have to be one second apart.
[3] For clarity, the fifth rolling window does not overlap with the first rolling window, as it starts (250,000x4) 1,000,000 data points after the start of the first rolling window.

4.6. Result

This section presents the results and findings. An overview of the whole period is offered, charting the SD, NDC and median $aTMV_{EXT}$ values measured in the rolling windows. Four individual sub-periods are then examined to highlight the usefulness of each of the three indicators.

4.6.1. Overview

As shown in Figure 4. 2 and Figure 4. 3, SD, median $aTMV_{EXT}$ and NDC report similar volatilities in some periods but not in others. Different indicators report variations in volatility, with high volatility sometimes coinciding under different indicators, while in other time periods, one indicator may show the high volatility while the others do not.

4.6.2. The results of the whole period P

Market tracking using the indicators in DC and TS series are shown in Figure 4. 2 and Figure 4. 3, respectively. The higher the value, the higher the volatility of the market according to the indicator.

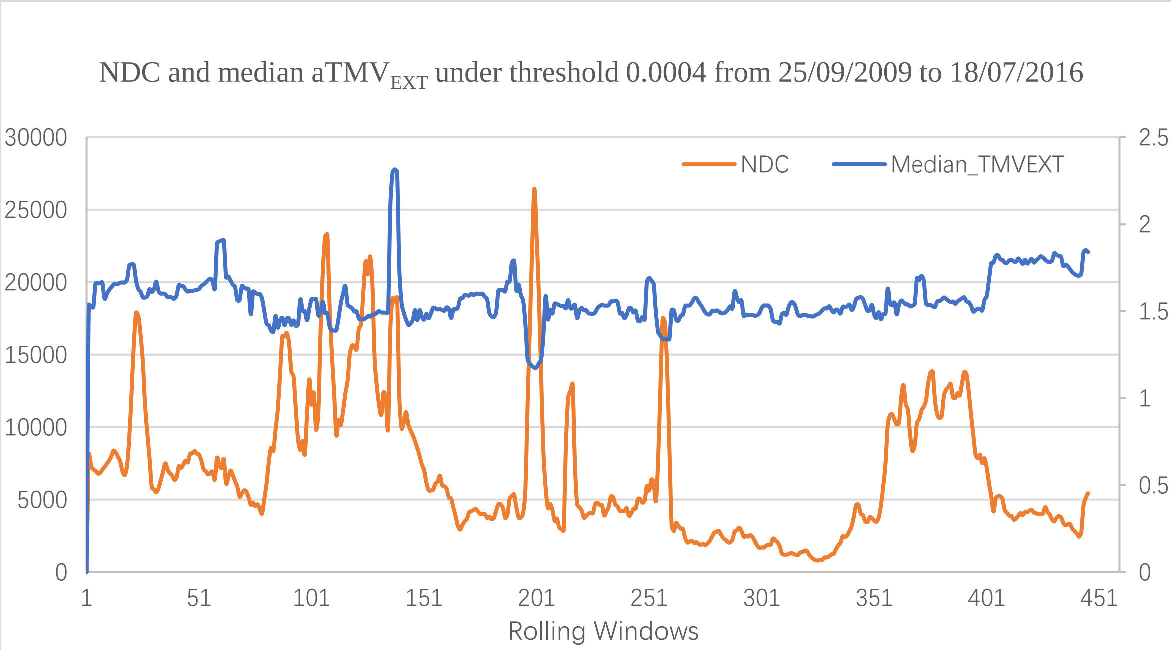

Figure 4. 2 NDC and median aTMV$_{EXT}$ for the rolling windows EUR/USD data from 25/09/2009 to 18/07/2016 under DC threshold 0.0004

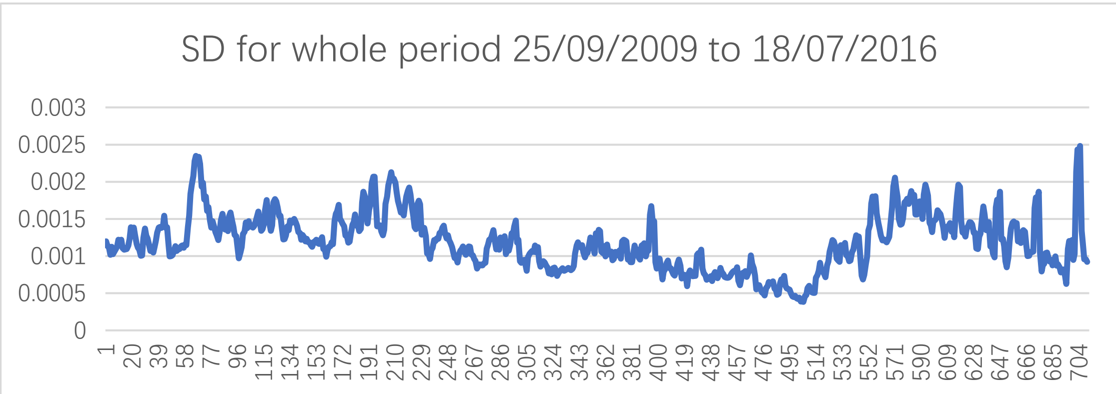

Figure 4. 3 SD (log return) for the rolling windows under TS for EUR/USD data from 25/09/2009 to 18/07/2016

It is important to point out that both x-axis of Figure 4. 2 and Figure 4. 3 represents the number of rolling windows. For example, when the x-axis is equal to 1, that means it's the value of the first window. The reason Figure 4. 2 and Figure 4. 3 have different number is that DC and TS use different rolling windows and rolling speeds which will lead the total number of windows different. This requires the values to be plotted on different figures. The rolling windows in TS are based on physical time, while the rolling windows in DC are data-driven, and thus the first 500 events could take a much longer (or shorter) physical time than the next 500 events.

4.6.3. Overview of the period P under investigation

Under Directional Change with a threshold equal to 0.0004, the median $aTMV_{EXT}$ from each rolling window's value ranges from 1.175165 to 2.315351. The largest median $aTMV_{EXT}$ value in any rolling window is equal to 2.315351, which occurred in the time period from 08:10:12 on 22/09/2011 to 02:11:17 on 11/11/2011. The smallest median $aTMV_{EXT}$ in the rolling window was equal to 1.175165, which occurred in the period from 10:55:41 on 01/10/2012 to 01:51:09 on 19/10/2012. The average value of median $aTMV_{EXT}$ for all rolling windows in the whole period P is 1.567286.

With a threshold is equal to 0.0004, the NDC for each rolling window's value ranged from 802 to 26,433. NDC= 802 occurred during the rolling window time period from 14:44:33 on 03/07/2014 to 02:31:01 on 25/07/2014, while NDC= 26,433 occurred during the rolling window time period from 01:51:09 on 01/10/2012 to 01:51:09 19/10/2012. The average value of NDC in the rolling windows in the whole period P is 7,011.

Under Time Series, the volatility (SD) was calculated using hourly log returns within each of the windows. By inspecting all windows, the maximum value of volatility was identified as 0.002481249, while the minimum was 0.000383666. The average value of volatility was 0.001187753, and the median volatility value of the time series was 0.001154151.

4.6.4. Sub-periods that we focus on

The rest of this chapter focuses on four sub-periods (P1, P2, P3, and P4), which are further defined below. These four periods are examined under both DC and TS, giving $P_{DC}1$ to $P_{DC}4$ and $P_{TS}1$ to $P_{TS}4$. Observations based on median $aTMV_{EXT}$, NDC, and SD are independently calculated, and the results are compared and contrasted to assess the usefulness of using multiple indicators to track the market.

As TS and DC have different sampling points, they do not use exactly the same physical timings for their time periods. Each of $P_{DC}1$ to $P_{DC}4$ and $P_{TS}1$

to $P_{TS}4$ thus includes all periods that cover the observations.

Each individual indicator, median $aTMV_{EXT}$, NDC, and SD, is assessed with specific reference to its usefulness in tracking. For this purpose, these indicators' potential roles in four periods are examined for both DC and TS.

As DC is data-driven, the data points chosen in DC and TS are very different. To make the results under DC and TS comparable, the DC and TS windows for roughly the same time period were selected. The time periods used are defined in Table 4. 1, with the periods used for collecting DC indicators, denoted as $P_{DC}1$, $P_{DC}2$, $P_{DC}3$ and $P_{DC}4$, and the periods in which TS indicators were collected denoted as $P_{TS}1$, $P_{TS}2$, $P_{TS}3$, and $P_{TS}4$.

Table 4. 1, shows the four periods under both DC and TS that have significant changes in one of them or both of them.

P1: February to July 2010
$P_{DC}1$: 07:06:32 9th February 2010 to 15:53:34 29th July 2010 (20 rolling windows) $P_{TS}1$ 17:00:16 3rd February 2010 to 17:00:06 28th July 2010 (26 rolling windows)
P2: September 2011 to January 2012
$P_{DC}2$: 08:10:12 22^{ed} September 2011 to 00:59:05 11th January 2012 (20 rolling windows) $P_{TS}2$ 17:00:06 09th November 2011 to 17:00:06 11th January 2012 (19 rolling windows)
P3: May to August 2013
$P_{DC}3$: 09:37:53 3rd May 2013 to 06:18:42 29th August 2013 (25 rolling windows) $P_{TS}3$ 17:00:04 3rd May to 17:00:07 18th August 2013 (32 rolling windows)
P4: April to June 2016
$P_{DC}4$: 22:39:23 24th April 2016 to 12:19:43 20th June 2016

(8 rolling windows) $P_{TS}4$ 17:00:04 28th April 2016 to 17:02:24 7th July 2016 (15 rolling windows)

Table 4. 1 The four periods under both DC and TS with significant changes in one of them or both of them

These four periods were picked because of the significant changes in volatility observed in the DC and Time series data sets. Every period has its own characteristics and values of each indicator under DC and TS, however, which are discussed in the following sections.

The following three sections thus focus on the observations made under SD, median $aTMV_{EXT}$ and NDC. For each of these indicators, the observations in the four sub-periods listed in Table 4. 1 will be summarized and discussed.

4.6.5. Observations by SD in the four different periods

This subsection focuses solely on the observations made under SD in each of the four periods $P_{TS}1$ to $P_{TS}4$

In this section, we found Observation 1: where we can found high historical volatility under the periods $P_{TS}1$ and $P_{TS}4$. In $P_{TS}2$ and $P_{TS}3$ we do not observe a significant signal for changes.

4.6.5.1.Observation by SD from $P_{TS}1$

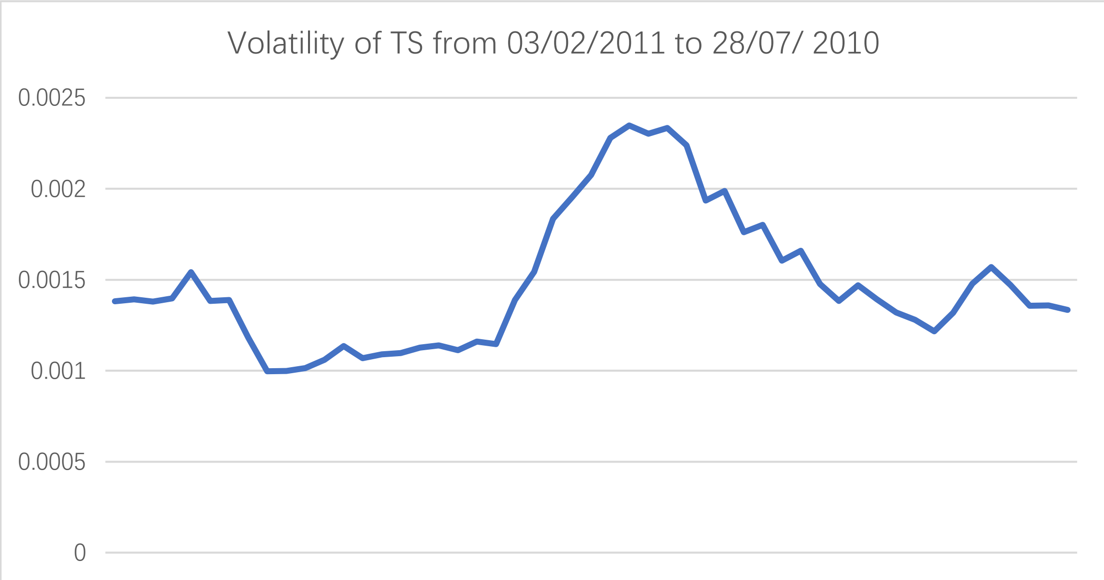

Figure 4. 4 SD observed under TS from 03/02/2011 to 28/07/ 2010 ($P_{TS}1$)

$P_{TS}1$ covers the period from 17:00:16 on 03/02/2010 to 17:00:06 on 28/07/2010. The average SD for this period is 0.001485; for reference, the average SD of P is 0.0012. In $P_{TS}1$, the average SD is thus more or less the same across P. However, a big jump can be observed from 0.00114 on 04/04/2010 April to the peak value of 0.00234 on 09/05/2010, with the latter value being 3.10 times the standard deviation away from the mean value of P. It is also the third highest SD value across P (as noted in Section 4.6.3):

Observation 1: Volatility observed based on SD in $P_{TS}1$ is one of the highest in the whole period (P); it is 3.1 standard deviation above the mean SD value of P, which is 99.903% higher than all SD values.

4.6.5.2.Observation by SD from $P_{TS}2$

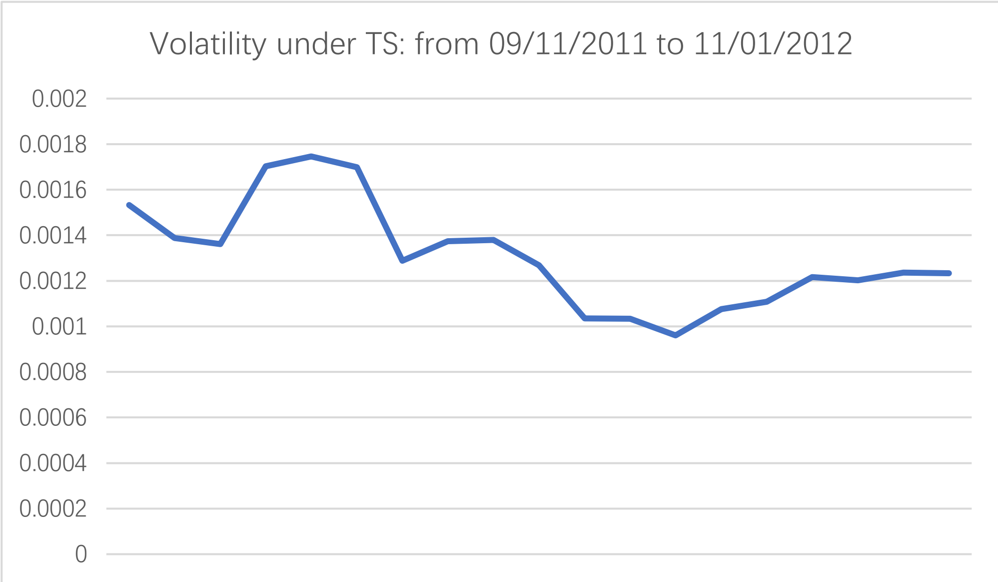

Figure 4. 5 SD observed under TS: from 09/11/2011 to 11/01/2012 ($P_{TS}2$)

From Figure 4. 5, from 17:00:06 on 09/11/2011 to 17:00:06 on 11/01/2012 the value of SD under time series lay between 0.00096053 and 0.001746078. The average value in this period was 0.001307427, just 0.32 standard deviations away from the population mean. Figure 4. 3 shows that the SD value during $P_{TS}2$ was relatively low in comparison with P as a whole, yet nothing special was observed in $P_{TS}2$.

4.6.5.3.Observation by SD from $P_{TS}3$

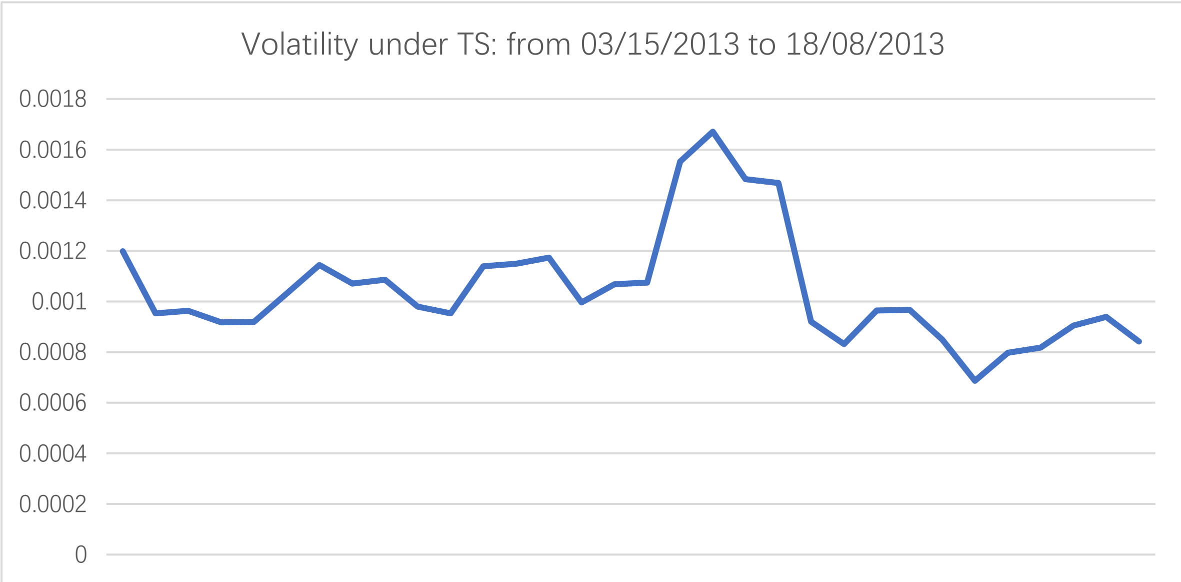

Figure 4. 6 SD observed under TS: from 03/15/2013 to 18/08/2013 ($P_{TS}3$)

Figure 4. 6, shows that, from 17:00:04 on 03/05/2013 to 17:00:07 on 18/08/2013, the value of SD under time series was between 0.00068 and 0.00167. The average value in this period was 0.001307427, just 0.32 standard deviations away from the population mean. From Figure 4. 3 we can see that the SD value in $P_{TS}3$ is relatively low within P. In other words, there are no particular leaps were observed in $P_{TS}3$.

4.6.5.4. Observation by SD from $P_{TS}4$

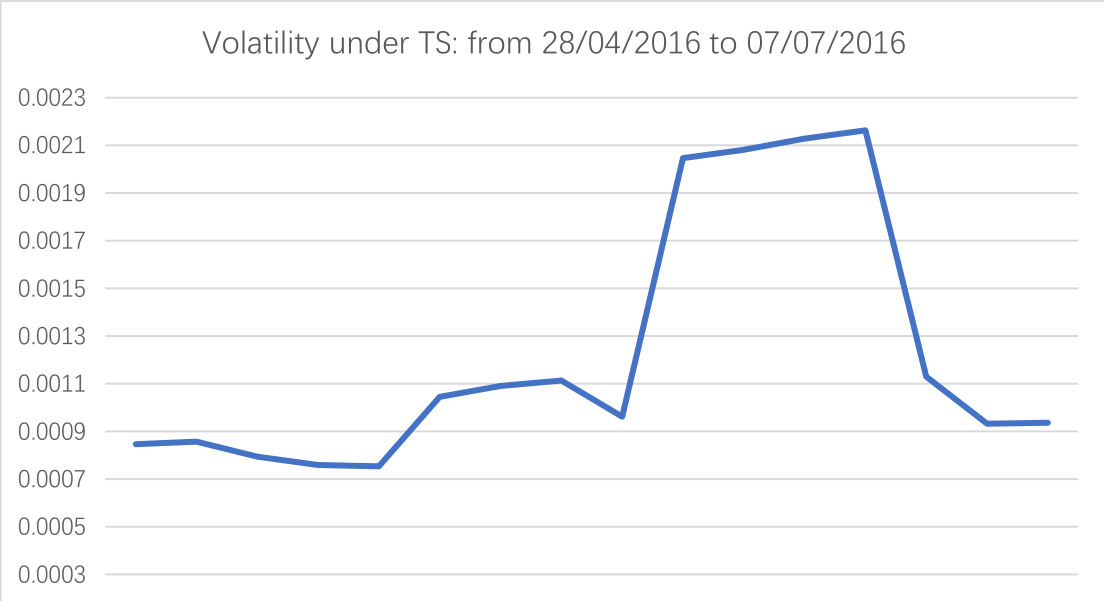

Figure 4. 7 SD observed under TS: from 28/04/2016 to 07/07/2016 ($P_{TS}4$)

From Figure 4. 7, we can see that from 17:00:04 28th April to 17:02:24 7th July 2016 the value of SD under time series is between 0.000753596 and 0.00216308. The average value in this period is 0.00122743. From Figure 4. 7 we can see that the SD value in $P_{TS}4$ started from around 0.00075 and ended in 0.000931. In between, there was a big increase from 0.000961 to 0.00204 and stay for one window before dropping back to 0.001129. At its peak, a 3.5 times standard deviation away from the mean value of $\mathbb{P}$ was thus observed.

Observation 2: The volatility observed in SD in $P_{TS}4$ was among the highest across the whole period ($\mathbb{P}$), being 3.5 standard deviations above the mean SD value in $\mathbb{P}$, making it higher than 99.976% of all SD values.

4.6.5.5. Conclusion on observation by SD

To conclude, the major observation under SD are:

Observation 1: The volatility observed in SD in $P_{TS}1$ is among the highest in the whole period (P), being 3.1 standard deviations above the mean SD value of P, making it higher than 99.903% of all SD values.

The volatility observed in SD in $P_{TS}4$ is among the highest across the whole period (P), being 3.5 standard deviations above the mean SD value of P, making it higher than 99.976% of all SD values.

To summarize, under TS, we can observe two changes, one is under $P_{TS}1$ and the other $P_{TS}4$.

4.6.6. Observation by median $aTMV_{EXT}$ under DC in the four different periods

This section focuses on the observations made using the DC indicator median $aTMV_{EXT}$ in each of the four periods $P_{DC}1$ to $P_{DC}4$.

4.6.6.1. Observations by median aTMV$_{EXT}$ from P$_{DC}$1

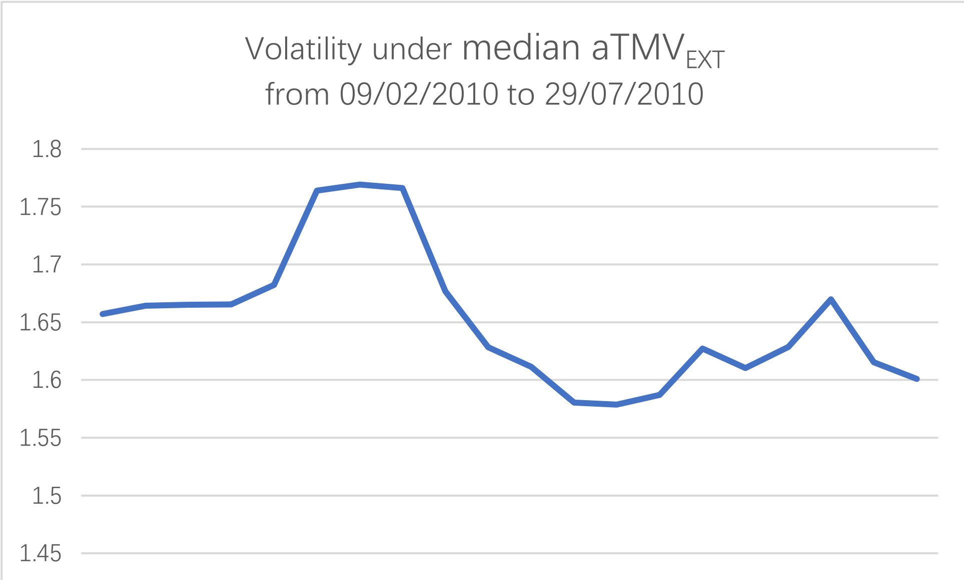

Figure 4. 8 Volatility under aTMV from 09/02/2010 to 29/07/2010 (P$_{DC}$1)

Figure 4. 8 shows the median aTMV$_{EXT}$ values from 07:06:32 on 09/02/2010 to 15:53:34 on 29/07/2010. The median aTMV$_{EXT}$ values ranged between 1.580392 and 1.76 in this period, with the average value being 1.65241, which is 0.62 standard deviations away from the population mean. This indicates that this period is slightly more volatile than average, but not by a margin; nothing special was observed in P$_{DC}$1.

4.6.6.2. Observations by median aTMV$_{EXT}$ from P$_{DC}$2

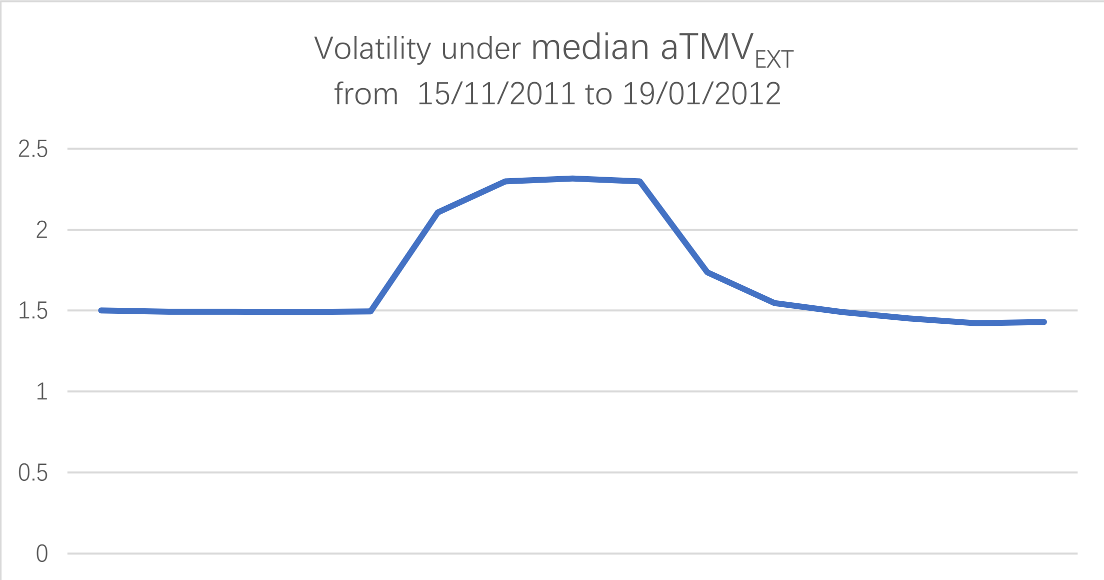

Figure 4. 9 Volatility under median aTMV$_{EXT}$ from 15/11/2011 to 19/01/2012 (P$_{DC}$2)

Figure 4. 9 shows that volatility went up dramatically between 27/11/2011 to 27/12/2011. On 20 November 2011, the median aTMV$_{EXT}$ was equal to 1.5, and it then rose to 2.1 on 27 December 2011. The median aTMV$_{EXT}$ reached a historic high of 2.31 in the period ending 16th November 2011, and the median aTMV$_{EXT}$ appears to have picked up the historic high volatility of this period, which was 5.46 times SD away from the mean of P.After reaching this historic high, the value of the median aTMV$_{EXT}$ decreased back to 1.49, much nearer the average value 1.56 for the whole period. This suggests there is something significant in this period.

Observation 3: Volatility observed via the median aTMV$_{EXT}$ value in DC in P$_{DC}$2 reached a historic high across the period (P), being

5.46 standard deviations higher than the median $aTMV_{EXT}$ value of P, placing it above 99.999% of all SD values.

4.6.6.3. Observations by aTMV from $P_{DC}3$

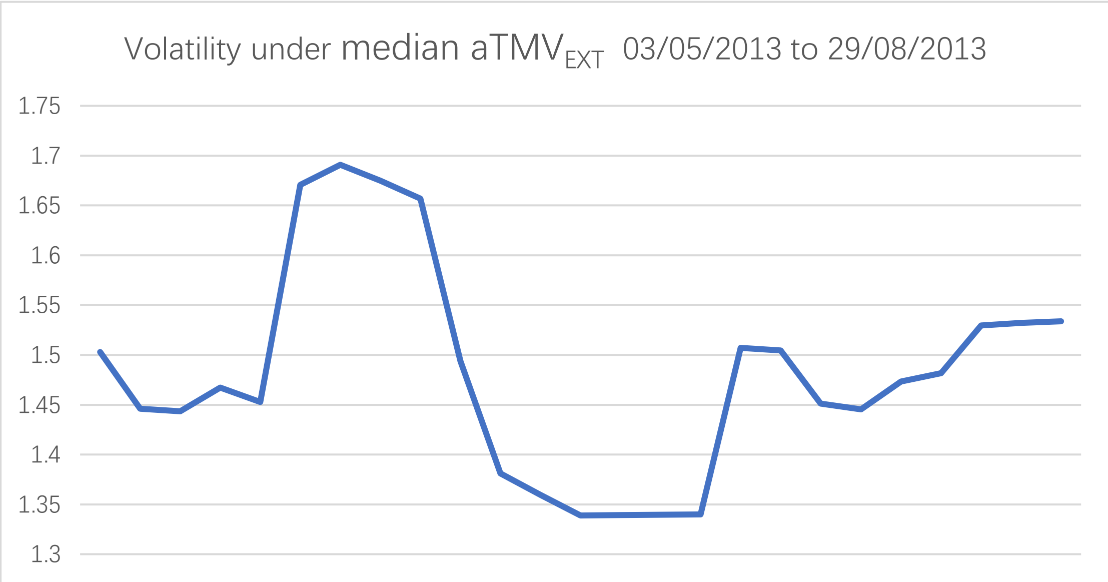

Figure 4. 10 Volatility under median aTMV$_{EXT}$
from 03/05/2013 to 29/08/2013 ($P_{DC}3$)

In this period (shown in Figure 4. 10), median aTMV$_{EXT}$ first went up from 1.45 to 1.69 and then decreased to 1.33. Later, the median aTMV$_{EXT}$ again increased, from 1.33 to 1.506. As noted in Section 4.6.2 and Section 4.6.3, the average value of median aTMV$_{EXT}$ for P is 1.56. Volatility in this period is therefore relatively low within P, and while there is fluctuation in the median aTMV$_{EXT}$ value, no special patterns can be observed in this period under median aTMV$_{EXT}$.

4.6.6.4.Observations by aTMV from $P_{DC}4$

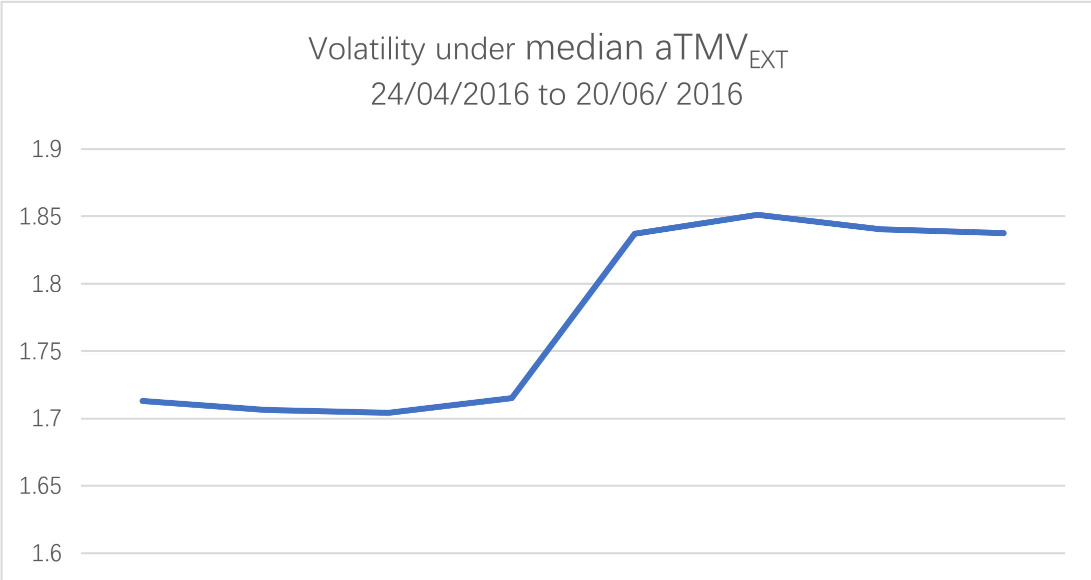

Figure 4. 11 Volatility under median $aTMV_{EXT}$
from 24/04/2016 to 20/06/2016 ($P_{DC}4$)

From 24/04/2016 to 20/06/2016, little change was observed in the median $aTMV_{EXT}$ value; this increased from 1.70 to 1.85, then remained above 1.8 until the end of the period, with nothing special observed.

4.6.6.5. Conclusions under observations by median $aTMV_{EXT}$

To conclude, the major observation under median $aTMV_{EXT}$ is:

Observation 3: Volatility observed by median $aTMV_{EXT}$ value in DC in $P_{DC}2$ reached a historic high across the period (P) at 5.46 standard deviation above the median $aTMV_{EXT}$ value of P, above 99.99% of all SD values.

Based on median $aTMV_{EXT}$ in DC, only one significant change, under $P_{DC}2$, was thus observed.

4.6.7. Observation by NDC under DC in the four different DC periods

This section shows the results obtained under NDC, with a focus on one period in each of the following sub-sections.

4.6.7.1. Observation by NDC from $P_{DC}1$

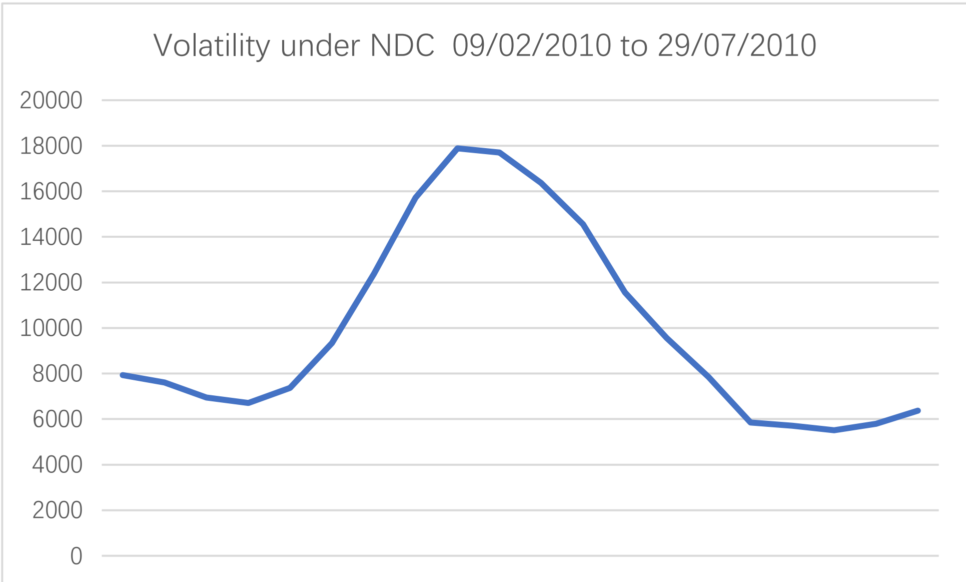

Figure 4. 12 Volatility under NDC 09/02/2010 to 29/07/2010 ($P_{DC}1$)

Figure 4. 12 shows the NDC increasing from 6,707 to 17,697 before dropping back to 5,845 throughout $P_{DC}1$. At the peak, when the NDC is equal to 17,866, it is 2.258 standard deviations away from the population mean, being higher than 99.802% of all measurements. Further, when the NDC is equal to 5,845 and 6,707, these are -0.24 and -0.06 standard deviations away from the mean, respectively. These figures suggest that in $P_{DC}1$, a significant change both up and down occurs.

Observation 4: Volatility observed in DC under NDC in $P_{DC}1$ is one of the highest in the whole period (P); it is 2.258 standard deviation above the mean NDC value of P, which is 98.802% above all SD values.

Also, in $P_{DC}1$, we observed one of the biggest jumps in NDC over the whole period P.

4.6.7.2. Observation by NDC from $P_{DC}2$

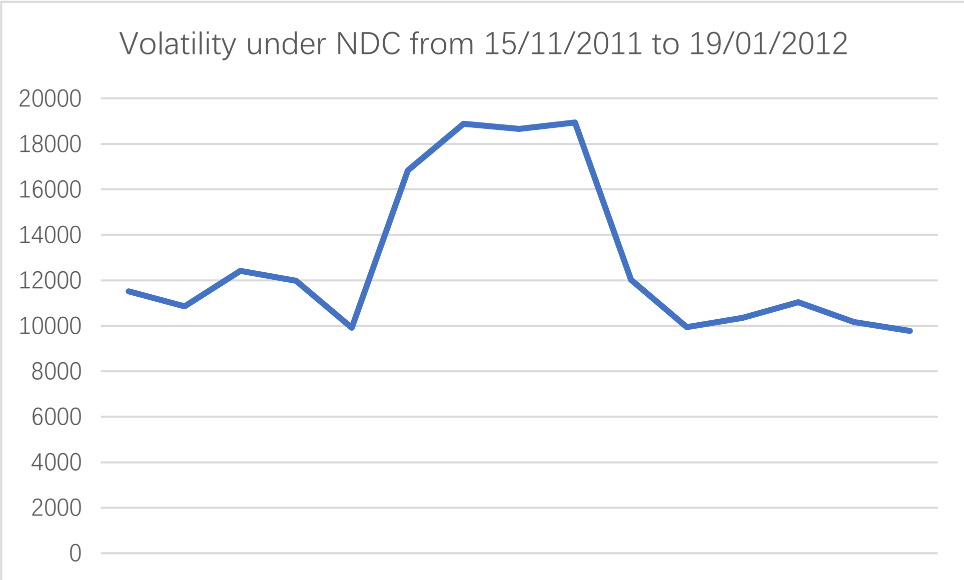

Figure 4. 13 Volatility under NDC 15/11/2011 to 19/01/2012 ($P_{DC}2$)

Figure 4. 13 shows that the average NDC in $P_{DC}2$ is 12,877 while the average NDC in P is 7,011. In $P_{DC}2$, the peak value of NDC is 18,891, 2.47 standard deviations away from the mean, while the smallest value of NDC is 9,938, 0.66 standard deviations away from the mean. Compared with other periods in P, something appears to occur in $P_{DC}2$, but it is not particularly significant.

4.6.7.3.Observation by NDC from $P_{DC}3$

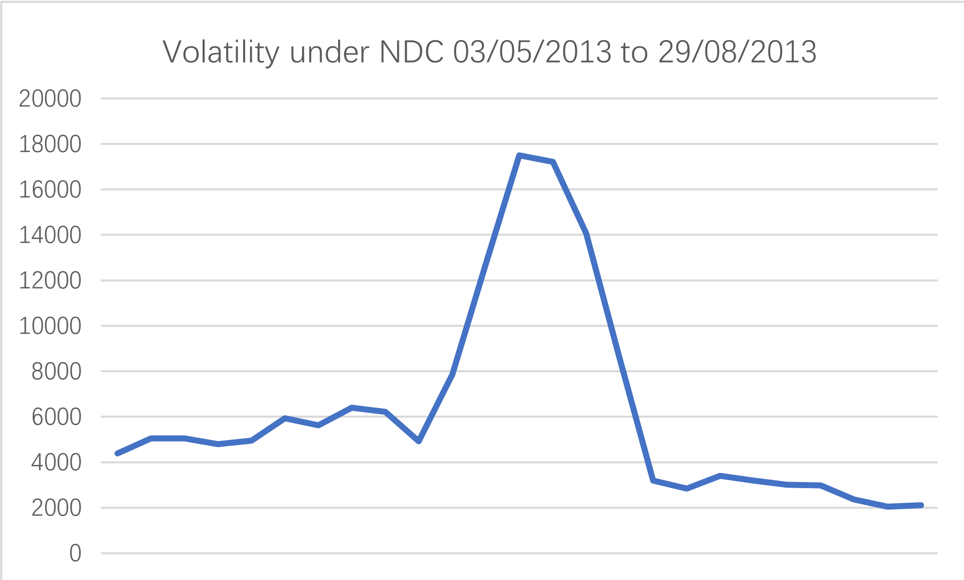

Figure 4. 14 Volatility under NDC 03/05/2013 to 29/08/2013 ($P_{DC}3$)

From Figure 4. 14, it is clear that NDC in $P_{DC}3$ went up dramatically between 16/06/2013 and 01/07/2013, from 4,920 to 17,496, more than triple the original value, then decreased quickly between 10/07/2013 and 02/08/2013, falling from 17,207 to 2,844, which is just one-sixth of the highest value. This clearly shows something significant in this period.

Observation 5: The volatility observed in DC under NDC in $P_{DC}3$ was among the highest in the whole period P (top 7%); in $P_{DC}3$, NDC also showed a bigger jump than in all other periods, increasing by three times and decreasing even further after the peak point.

4.6.7.4. Observation by NDC from $P_{DC}4$

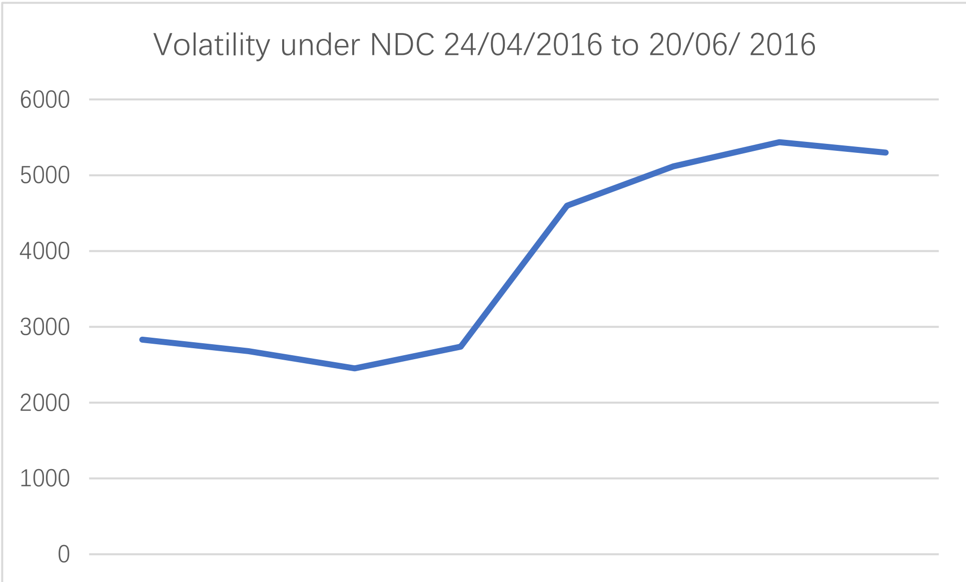

Figure 4. 15 Volatility under 24/04/2016 to 20/06/ 2016 ($P_{DC}4$)

Figure 4. 15 shows that, from 24/04/2016 to 20/06/2016, the NDC value increased from 2,452 to 5,436. As mentioned in Section 4.6.2 and Section 4.6.3, the average NDC was 7011; thus, when the NDC is equal to 5,436, this is -0.33 standard deviations away from the population average, while when NDC is equal to 2,452, it is -0.98 standard deviations away from the population average, which is not that far from the mean. Thus, nothing significant was observed in this period.

4.6.7.5. Conclusions under observation by NDC

To conclude, the major observations under NDC are:

Observation 4: Volatility observed in DC under NDC in $P_{DC}1$ is one of the highest in the whole period (P); it is 2.258 standard deviation above the mean NDC value of P, which is 98.802% above all SD values.

Observation 5: Volatility observed in DC under NDC in $P_{DC}3$ is one of the highest in the whole period P (Top 7%); In P3, NDC have a big jump that all other periods do not happen, its increase three times and decrease even deeper from the peak point.

To summarize, under NDC, we can observe two significant changes, one is under $P_{DC}1$ and the other $P_{DC}3$.

4.6.8. Summary of all the results

From the observations above, five observations on market volatility under both TS and DC emerged; these are rearranged below to better support the discussion in the next section:

1. Both NDC and median $aTMV_{EXT}$ observed high volatility in $P_{DC}1$. A big jump in SD is also observed in $P_{DC}1$'s corresponding sub-period $P_{TS}1$. (Observation 1: and Observation 4:).

2. A big jump in SD is observed in $P_{TS}4$ (Observation 2:).

3. A big jump in median $aTMV_{EXT}$ is observed in $P_{DC}2$ (Observation 3:).

4. A big jump in NDC is observed in $P_{DC}3$ (Observation 5:).

4.7. Discussion

The EUR/USD market (the most traded currency market) was tracked over a long period of time (June 2009 to July 2016) using SD (standard deviation in log returns) under Time Series (TS) and median $aTMV_{EXT}$ and NDC under DC. Here tracking means measuring SD, $aTMV_{EXT}$ and NDC rolling window by rolling window. Doing so allows us to assess the volatility of each rolling window as measured by the individual volatility indicators. It is worth reiterating that TS and DC use different rolling windows (see Sections 4.3 and 4.5.2).

The results show that volatility can be low under one indicator yet high under another; this indicates that all examined indicators are useful and cannot be entirely substituted for by the others. Table 4. 2 summarizes the main observations for all three indicators (SD, median $aTMV_{EXT,}$ and NDC) in the four periods (P1 to P4).

The subsections below will establish how the three volatility indicators complement each other:

i. We shall show that the three indicators agree with each other in some situations (Section 4.7.1); and

ii. We shall show that each indicator finds volatile periods that were not found by the other indicators (Sections 4.7.3, 4.7.4, and 4.7.2 respectively). This means no indicator can be replaced by the other two.

Table 4. 2 Summary of the findings for both TS and DC

Periods	P1 February 2010 to July 2010	P2 September 2011 to January 2012	P3 May 2013 to August 2013	P4 April 2016 to June 2016
Overall summary	Both SD and NDC observed high volatility	median $aTMV_{EXT}$ observed high volatility not observed by others	NDC observed high volatility not observed by others	SD observed high volatility not observed by DC
SD	High volatility was observed in $P_{ST}1$ (Observation 1:)	Nothing special was observed in $P_{ST}2$	Nothing special was observed in $P_{ST}3$	High volatility was observed in $P_{ST}4$ (Observation 2)
median $aTMV_{EXT}$	Nothing special was observed in $P_{DC}1$	median $aTMV_{EXT}$ has picked up historical high volatility in this $P_{DC}2$ (Observation 3:)	Nothing special was observed in $P_{DC}3$.	Nothing special was observed in $P_{DC}4$
NDC	High volatility was observed in $P_{DC}1$ (Observation 4:)	Nothing special was observed in $P_{DC}2$.	High volatility was observed in $P_{DC}3$ (Observation 5:)	Nothing special was observed in $P_{DC}4$.

Table 4. 2 Summary of the findings for both TS and DC

4.7.1. Information tracked by both TS and DC indicators

In some periods, both TS and DC found signals that supported each other. For example, referring to Section 4.6.5.1 and Section 4.6.7.1. It showed in P1, both TS and DC detected high volatility.

In Section 4.6.5.1, time series picked the third highest SD value (0.00234) in $P_{ST}1$, which is 3.101 away from the standard deviation of the whole population, while the average SD of this period is 0.001485, which is more or less the same as the average value of P (0.0012). This means SD had a significant change in this period. We can use this significant change to track market change.

Meanwhile, once we look back to Section 4.6.7.1, NDC increased from 5,584, which is lower than the average value (7,011) in P to 17,866, which is three times bigger than the start. Then, NDC goes down dramatically from 17,697 to 5,845 in $P_{DC}1$. This is where, 17,866 is 2.25, far away from the standard deviation of the whole population. In this case, we observed something significant in $P_{DC}1$.

This section has clarified how TS and DC can support each other in terms of discovering what is happening in financial markets. In P1, both TS and DC indicated significant changes, and this type of agreement can help track the market. Where a significant change is observed, it is possible to say that there are market changes in this period, or at least that *something* has happened in this period. However, some signals are picked up by one method that cannot be detected by the other methods, which is discussed further in sections 4.7.3 and 4.7.4 below.

4.7.2. Information tracked by SD in TS but not by DC indicators

In this section, we will show that TS can pick up signals that cannot be picked up by the other two indicators used in this chapter.

Under P4, volatility observed in SD in $P_{TS}4$ is one of the highest in the whole period (P); it is 3.5 standard deviation above the mean SD value of P, which is 99.976% above all SD values.

But, once we checked the change of aTMV and SD, we could observe nothing significant in $P_{DC}4$ (Section 4.6.6.4 and Section 4.6.7.4).

4.7.3. Information tracked by median $aTMV_{EXT}$ but not TS and NDC

In this section, we will show that median $aTMV_{EXT}$ can pick up signals that cannot be picked up by the other two indicators used in this paper.

Referring to Section 4.6.6.2, median $aTMV_{EXT}$ picked up historical high volatility in that period, stretching 5.46 times from the mean of the whole period. In $P_{DC}2$, median $aTMV_{EXT}$ went up dramatically between 27/11/2011 and 27/11/2011. On 20/11/2011, the median $aTMV_{EXT}$ was equal to 1.5, and it rose to 2.1 on 27/12/2011. The median $aTMV_{EXT}$ reached a historic high of 2.31 in the period ending 16/11/2011, after which the value of median $aTMV_{EXT}$ decreased back to 1.49, nearer the average value (1.56) of the whole period. In this period, however, the median $aTMV_{EXT}$ went up and down significantly, shown in huge market changes followed by a rapid reversion to normal.

At the same time, we do not observe any changes from TS and NDC, which can be checked from Section 4.6.5.2 and Section 4.6.7.2.

4.7.4. Information tracked by NDC but not TS and median $aTMV_{EXT}$

In this section, we will show that NDC can pick up signals that cannot be picked up by the other two indicators used in this paper.

Under P3, the NDC in $P_{DC}3$ went up dramatically between 16/06/2013 and 01/07/2013, from 4,920 to 17,496, with the latter being three times the original value; it then decreased quickly between 10/072013 and 02/08/2013, falling from 17,207 to 2,844, which is just one sixth of the value at the highest point. As the average value of NDC is 7,011, In this period, the NDC goes up from under average value to about 2.5 times average, offering a significant signal of market changes.

A further examination of the changes in SD and median $aTMV_{EXT}$ still made it possible to conclude that nothing significant occurred in P3

(Section 4.6.5.3 and Section4.6.6.3).

4.8. Experimental work on USD/JPY

4.8.1 Experimental work set up

4.8.1.1 Data used

USD/JPY data, collected by the second (some seconds showed no trading), from 17:00:40 on 27/09/2009 to 14:17:11 on 22/07/2015 was used; this set contains 84,991,649 data points (transactions) across the period and both under TS and DC were applied to the same raw data, so these data points were the same. For clarity of reference, this period is labelled $P_{USD/JPY}$.

4.8.1.2 Rolling window set up

In TS, SD is used to measure the volatility of the market. For DC the indicators median $aTMV_{EXT}$ and NDC are used to measure the volatility of the market.

DC rolling windows setup:

In DC, the rolling windows are defined by the number of data points[4]. Each DC rolling window was therefore comprised of 1,000,000 data points, with NDC and median $aTMV_{EXT}$ calculated for this set. The

[4] Each data point is a transaction, with transactions collated by the second. There may be no transactions in some seconds, and two adjacent transactions therefore do not have to be one second apart.

rolling speed was 250,000 data points, as each rolling window started 250,000 data points after the beginning of the previous rolling window[5].

TS rolling windows setup:

For the time series (TS) theory, 16 days was set as the rolling window size, and 4 days was set as the rolling speed. For each rolling window, the hourly log return was calculated as the SD.

The reason for using 1,000,000 data points per window under DC and 16 days under TS was to make the TS window size and DC window size approximately the same physical duration. However, as DC and TS sample data at different points, the rolling windows observed under DC and TS are different.

4.8.1.3 Threshold set up

Under DC we use a threshold value of 0.0004.

[5] For clarity, the fifth rolling window does not overlap with the first rolling window, as it starts (250,000x4) 1,000,000 data points after the start of the first rolling window.

4.8.2 the result of the whole period for both DC and TS

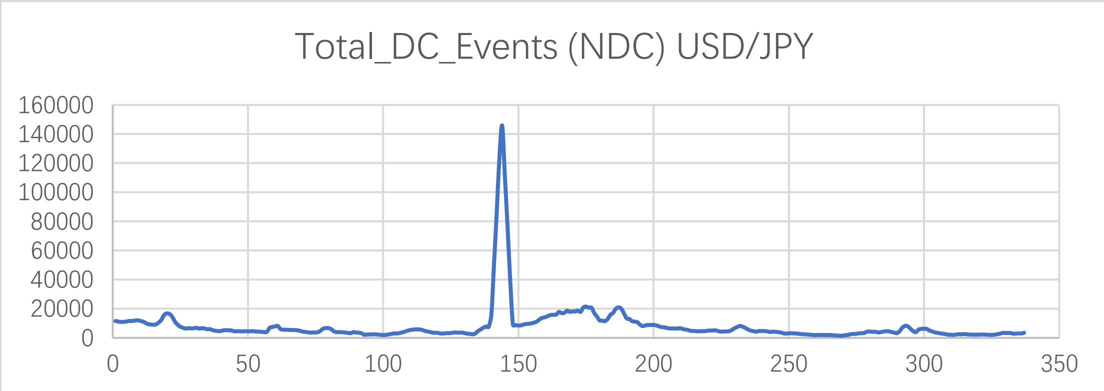

Figure 4. 16 NDC of rolling windows USD/JPY data from 27/09/2009 to 22/07/2015 under DC threshold 0.0004

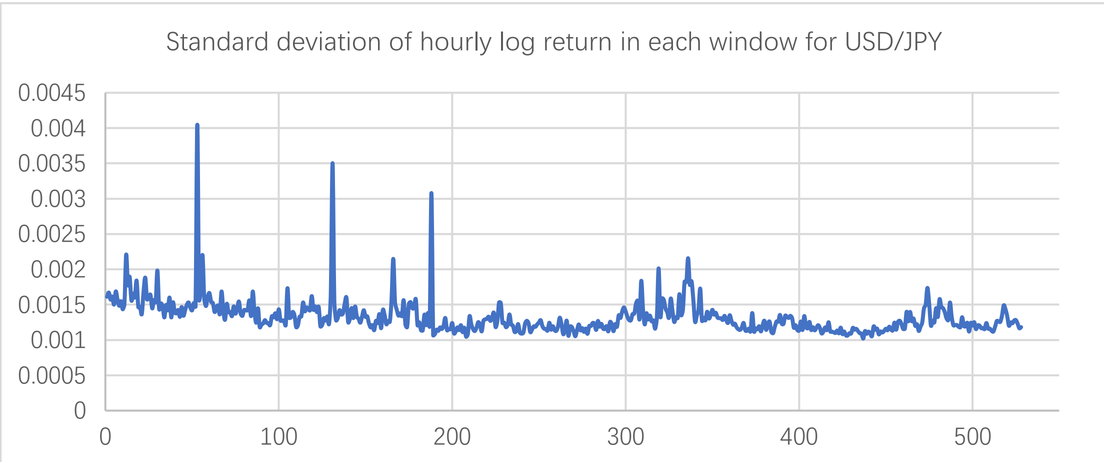

Figure 4. 17 SD (hourly log return) of rolling windows under TS for USD/JPY data from 27/09/2009 to 22/07/2015

Under Directional Change with a threshold equal to 0.0004, the NDC for each rolling window's value ranged from 1,437 to 145,784 NDC= 1437 occurred during the rolling window time period from 00:51:50 on 08/08/2014 to 09:17:32 on 01/09/2014, while NDC= 145,784 occurred during the rolling window time period from 17:20:08 on 01/10/2012 to 11:35:24 24/10/2012. The average value of NDC in the rolling windows in the whole period $P_{USD/JPY}$ is 8,190.

Under Time Series, the volatility (SD) was calculated using hourly log returns within each of the windows. By inspecting all windows, the maximum value of volatility was identified as 0.00404787, while the minimum was 0.00102048 The average value of volatility was 0.00132634, and the median volatility value of the time series was 0.00127099.

4.8.3 Observation by DC indicators

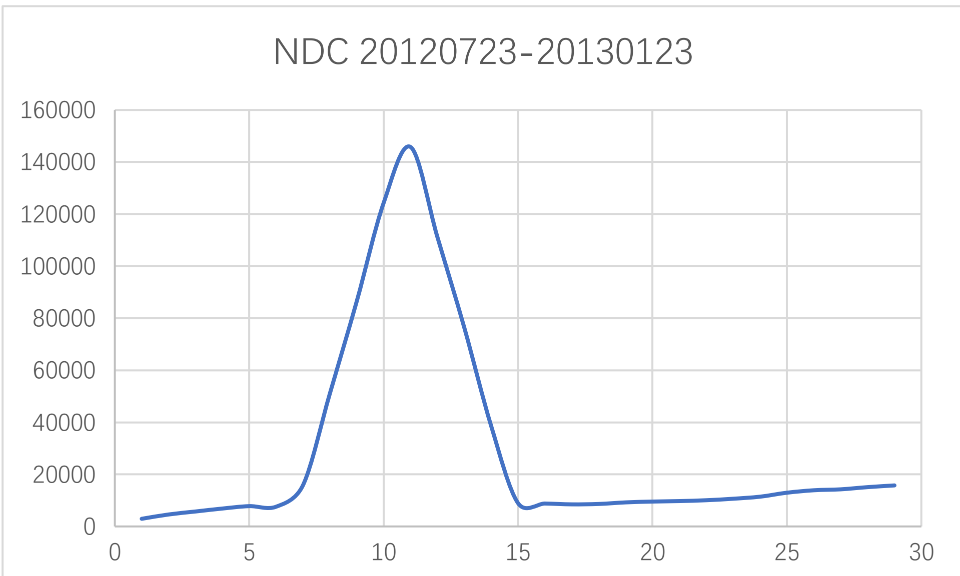

Figure 4. 18 Volatility under NDC 23/07/2012-23/01/2013 ($P_{USD/JPY}1$)

Figure 4. 18 shows the NDC increasing from 2,940 to 145,784 before dropping back to 8,913 throughout $P_{USD/JPY}1$. At the peak, when the NDC is equal to 145,784, it is 25.81 standard deviations away from the population mean, being higher than 99.99999% of all measurements. Further, when the NDC is equal to 8,913 and 2,940, these are 1.57 and -0.52 standard deviations away from the mean, respectively. These figures suggest that in $P_{ISD/JPY}1$, a significant change both up and down occurs.

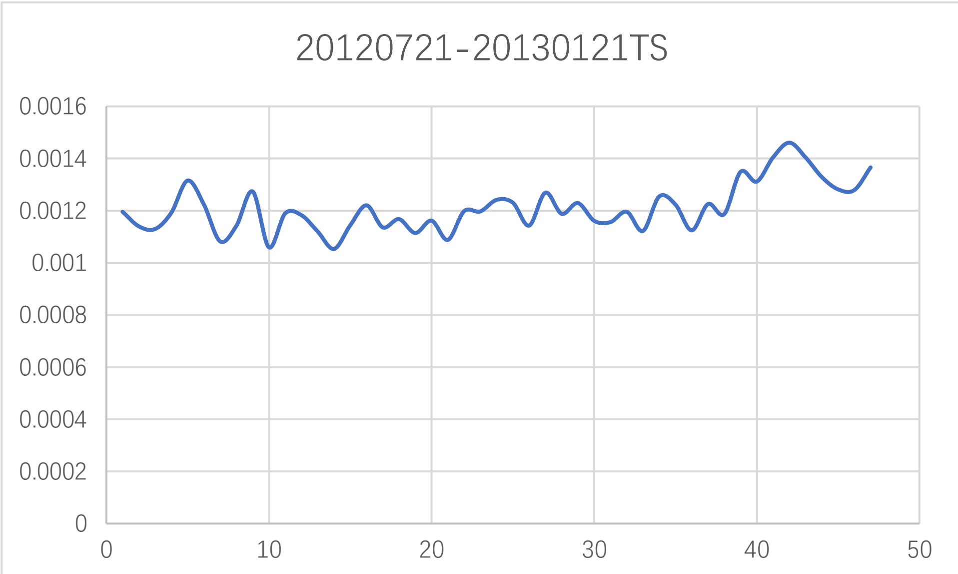

Figure 4. 19 Volatility under TS 07/04/2010 to 18/06/2010

Meanwhile, from Figure 4. 19 we can see under nearly the same time period, the volatility under TS is more or less the same where the average value is 0.00121(nearly the same as the average value of the whole period which is 0.0013) where the highest point is 0.00146 （1.35 standard deviations away from the mean of the whole population） and the lowest point is 0.00108(1.01 standard deviations away from the mean of the whole population). This means at the same period, TS does not have significant changes.

4.8.4 Observation by TS

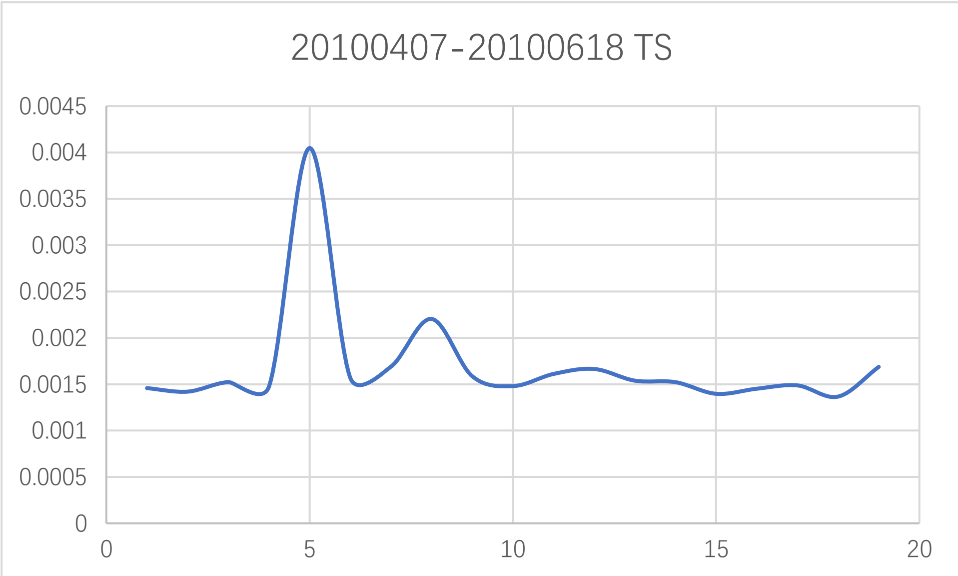

Figure 4. 20 Volatility under TS 07/04/2010 to 18/06/2010. ($P_{USD/JPY}2$)

$P_{USDJPY}2$ covers the period from 00:00:00 on 07/04/2010 to 00:00:00 on 18/06/2010. The average SD for this period is 0.00169369; for reference, the average SD of $P_{USD/JPY}$ is 0.00132634. In $P_{USDJPY}2$, the average SD is thus more or less the same as across $P_{USD/JPY}$. However, a big jump can be observed from 0.001458139 on 07/04/2010 to the peak value of 0.004047875 on 23/04/2010, which is also the highest SD value across $P_{USD/JPY}$.

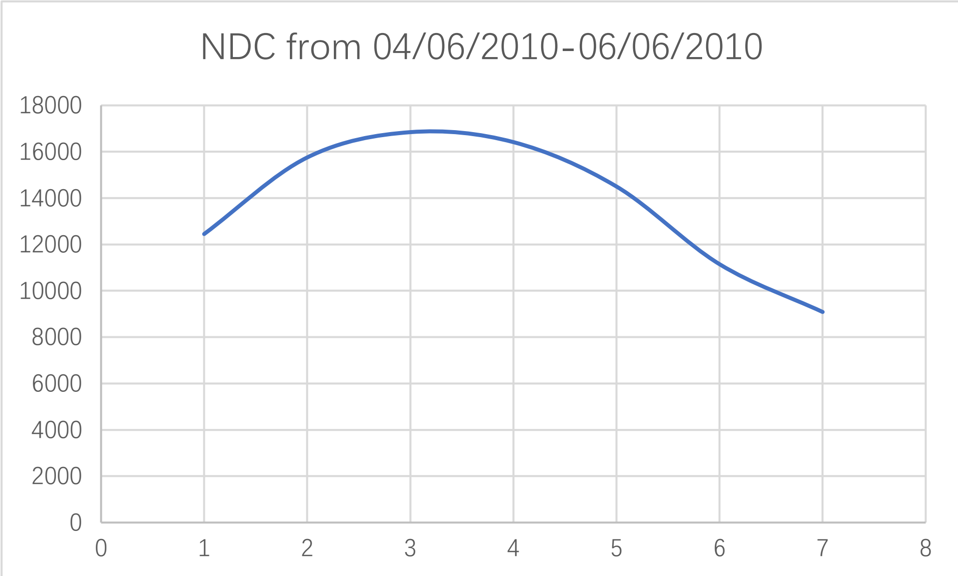

Figure 4. 21 NDC from 04/06/2010-06/06/2010

Meanwhile, from Figure 4. 21 we can see under nearly the same time period, the value of NDC under DC is more or less the same where the average value is 13,742 where the highest point is 16,845 and the lowest point is 9,087 (1.60 standard deviations away from the mean of the whole population). This means at the same period; DC does not have significant changes.

4.8.5 Observation by both DC and TS

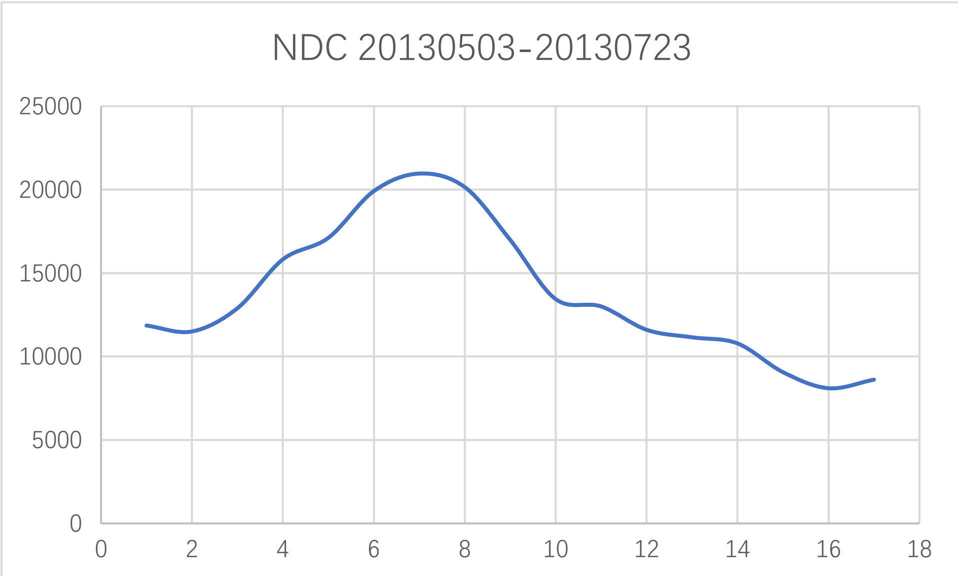

Figure 4. 22 Volatility under NDC 03/05/2013-23/07/2013 ($P_{USD/JPY}3$)

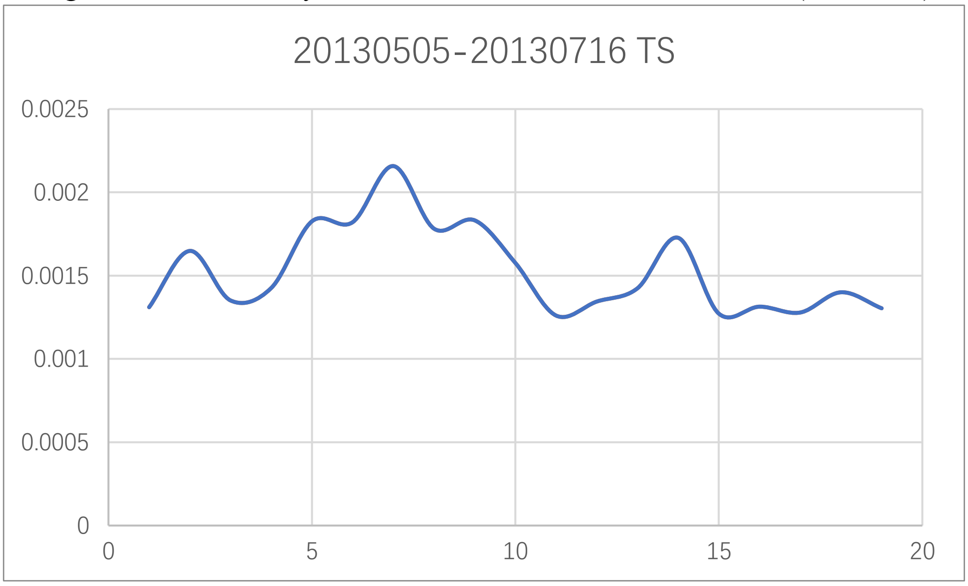

Figure 4. 23 Volatility under TS 05/05/2013 to 16/07/2013. ($P_{USD/JPY}4$)

Figure 4. 22 and Figure 4. 23 show that both DC indicator and TS SD shows a high value of volatility in both period $P_{USD/JPY}3$ and $P_{USD/JPY}4$, where $P_{USD/JPY}3$ and $P_{USD/JPY}4$ overlap with each other. In these two periods, both TS and DC found signals that supported each other.

In $P_{USD/JPY}4$, time series picked the sixth-highest SD value (0.002158) in

$P_{USD/JPY}$, which is 2.001 away from the standard deviation of the whole population, while the average SD of this period is 0.00152, which is more or less the same as the average value of $P_{USD/JPY}$ (0.00132). This means SD had a significant change in this period.

Meanwhile, once we look back to $P_{USD/JPY}3$, NDC increased from 11,851, then increase to 20,967, which is twice the start. Then, NDC goes down dramatically from 20,967 to 8,101 in $P_{USD/JPY}3$. This is where, 17,866 is 3.7128, far away from the standard deviation of the whole population which means that its less than 0.01% probability would be happened to reach this value in the whole $P_{USD/JPY}$. In this case, we observed a significant change in $P_{USD/JPY}3$.

4.8.6. Summary of all the USD/JPY results

From the observations above, three observations on market volatility under both TS and DC emerged

1. NDC observed high volatility in $P_{USD/JPY}1$
2. A big jump in SD is observed in $P_{USD/JPY}2$
3. Both DC and SD observed significant changes in $P_{USD/JPY}3$ and $P_{USD/JPY}4$ (where $P_{USD/JPY}3$ and $P_{USD/JPY}4$ are almost the same periods under DC and TS)

These results are consistent with our analysis in EUR/USD: As expected, volatility periods observed by NDC and SD overlapped with each other, as they both started with the same tick data. However, they each observe volatility periods not observed by the other indicator. This shows that neither NDC nor SD can be replaced by the other.

4.9. Experimental work on GBP/USD

4.9.1 Experimental work set up

In this section, we will use rolling windows, rolling speed and thresholds different from those used in Section 4.6 and Section 4.8. We do so to assess the generality of our results. In particular, we want to know whether the results are sensitive to the size of the rolling windows, rolling speed and thresholds used.

4.9.1.1 Data used

GBP/USD data, collected by the second (some seconds showed no trading), from 00:00:00 on 25/09/2009 to 14:00:05 on 22/07/2015 was used; this set contains 99,162,825 data points (transactions) across the period and both under TS and DC were applied to the same raw data, so these data points were the same. For clarity of reference, this period is labelled $P_{GBP/USD}$.

4.9.1.2 Rolling window set up

In TS, SD is used to measure the volatility of the market. For DC the indicators median $aTMV_{EXT}$ and NDC are used to measure the volatility of the market.

DC rolling windows setup:

In DC, the rolling windows are defined by the number of data points[6]. Each DC rolling window was therefore comprised of 500,000 data points, with median $aTMV_{EXT}$ calculated for this set. The rolling speed was 125,000 data points, as each rolling window started 125,000 data points after the beginning of the previous rolling window[7].

TS rolling windows setup:

For the time series (TS) theory, 8 days was set as the rolling window size, and 2 days was set as the rolling speed. For each rolling window, the hourly log return was calculated as the SD.

The reason for using 500,000 data points per window under DC and 8 days under TS was to make the TS window size and DC window size approximately the same physical duration. However, as DC and TS sample

---

[6] Each data point is a transaction, with transactions collated by the second. There may be no transactions in some seconds, and two adjacent transactions therefore do not have to be one second apart.
[7] For clarity, the fifth rolling window does not overlap with the first rolling window, as it starts (125,000x4) 500,000 data points after the start of the first rolling window.

data at different points, the rolling windows observed under DC and TS are different.

4.9.1.3 Threshold set up

Under DC we use the threshold value of 0.0016.

4.9.2 The result of the whole period

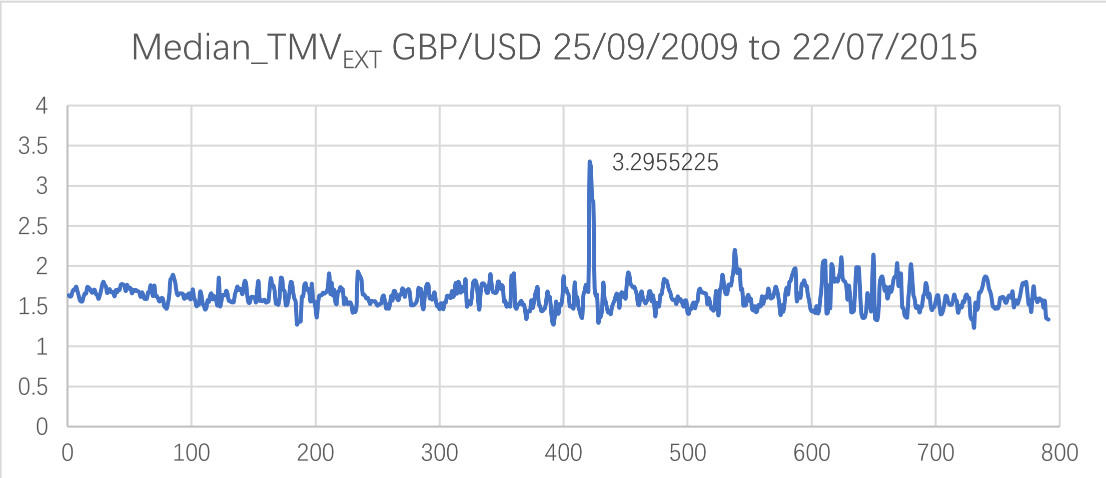

Figure 4. 24 Median_TMV$_{EXT}$ of rolling windows GBP/USD 25/09/2009 to 22/07/2015 under DC threshold 0.0016

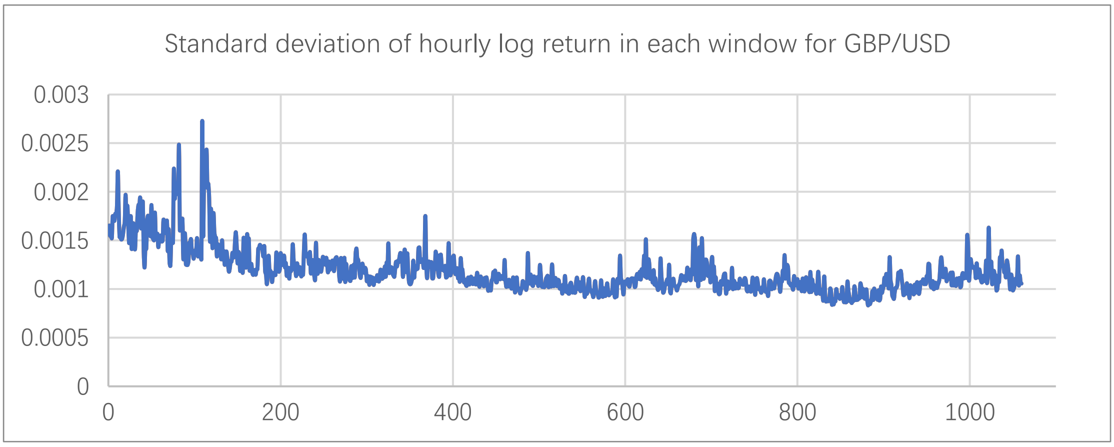

Figure 4. 25 SD (hourly log return) of rolling windows under TS for GBP/USD data from 25/09/2009 to 22/07/2015

Under Directional Change with a threshold equal to 0.0016, the $Median\_TMV_{EXT}$ for each rolling window's value ranged from 1.234461 to 3.2955225. $Median\_TMV_{EXT}$=3.2955225 occurred during the rolling window time period from 20:34:55 on 16/12/2012 to 02:46:38 on 27/12/2012, while $Median\_TMV_{EXT}$=1.234461 occurred during the rolling window time period from 13:58:02 on 24/02/2015 to 10:53:38 05/03/2015. The average value of $Median\_TMV_{EXT}$ in the rolling windows in the whole period $P_{GBP/USD}$ is 1.638966.

Under Time Series, volatility (SD) was calculated using hourly log returns within each of the windows. By inspecting all windows, the maximum value of volatility was identified as 0.00272915, while the minimum was 0.00083162 The average value of volatility was 0.00118171, and the median volatility value of the time series was 0.00112026

4.9.3 Observation by DC indicators

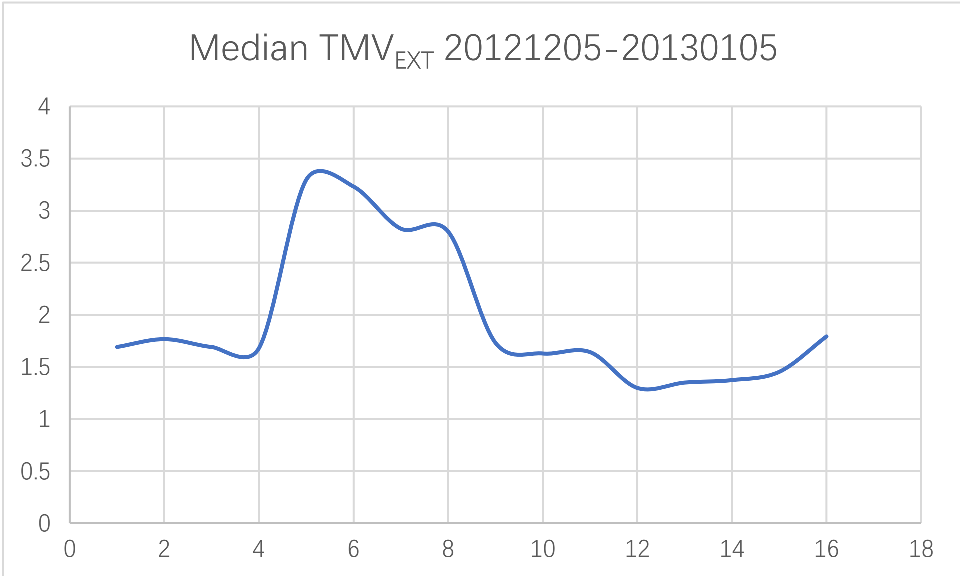

Figure 4. 26 Volatility under Median TMV$_{EXT}$ 20121205-20130105 (P$_{GBP/USD}$1)

Figure 4. 26 shows the Median_TMV$_{EXT}$ increasing from 1.6914 to 3.2955 than dropping back to 1.2984 throughout P$_{GBP/USD}$1. At the peak, when the Median_TMV$_{EXT}$ is equal to 3.2955 (the historical high point), it is 2.2 standard deviations away from the population mean, being higher than 99% of all historical measurements. Further, when the Median_TMV$_{EXT}$ is equal to 1.6914 and 1.1591, these are 1.57 and -0.88 standard deviations away from the mean, respectively. These figures suggest that in P$_{ISD/JPY}$1, a significant change both up and down occurs.

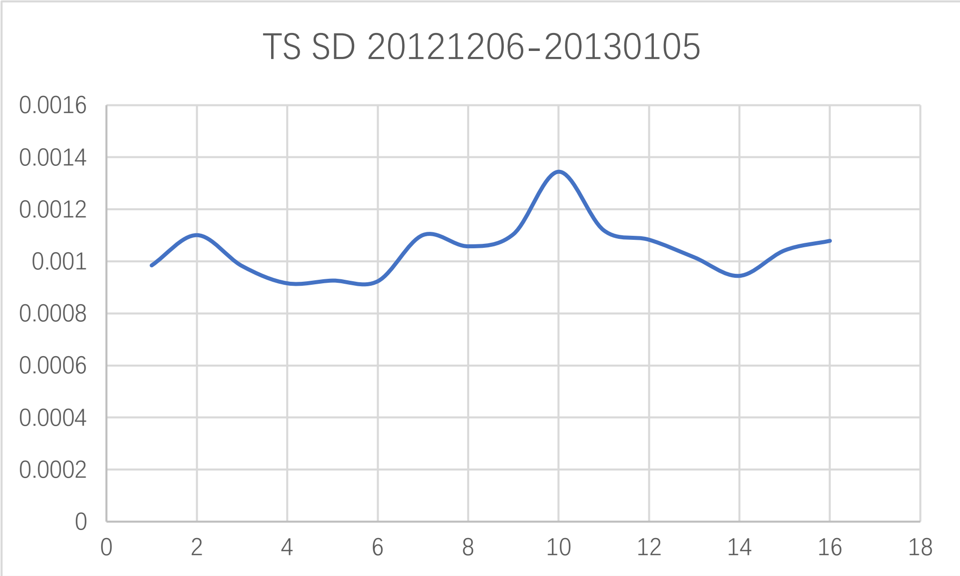

Figure 4. 27 Volatility under TS 06/12/2012 to 05/01/2013

In the meantime, under TS, in approximately the same time period, the average value of volatility from 06/12/2012 to 05/01/2013 is 0.001045 where is about the same as the average value of the whole period with the maximum value of 0.00134 (1.4 times away from the standard deviation) and minimum value of 0.00091 (-0.95 times away from the standard deviation).

4.9.4 Observation by TS

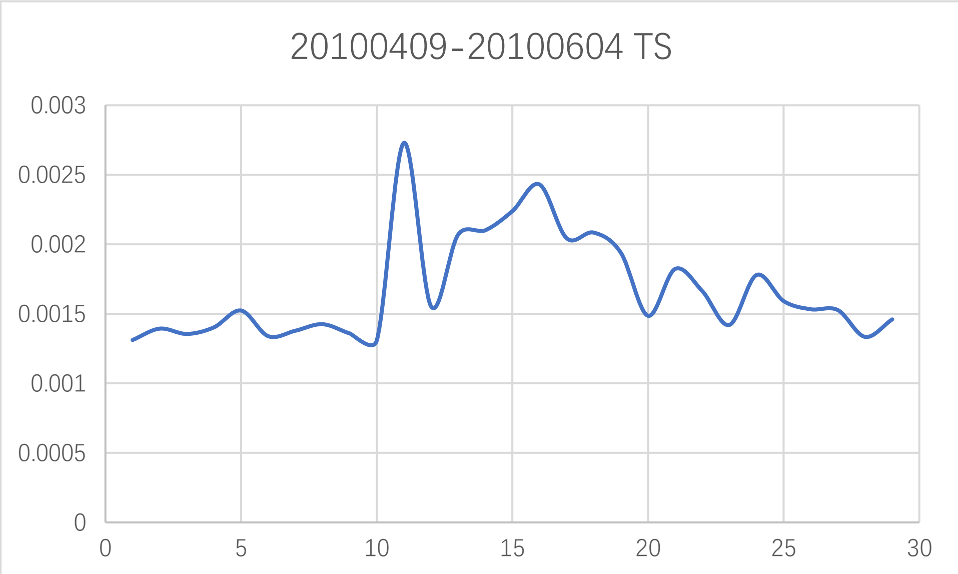

Figure 4. 28 Volatility under TS 09/04/2010 to 04/06/2010. ($P_{GBP/USD}2$)

Figure 4. 28 covers the period from 23:59:59 on 09/04/2010 to 23:59:59 on 04/06/2010. The average SD for this period is 0.00167612; for reference, the average SD of $P_{GBP/USD}$ is 0.001181. In $P_{GBP/USD}2$, the average SD is thus more or less the same as across $P_{USD/JPY}$. However, a big jump can be observed from 0.001307 on 27/04/2010 to the peak value of 0.002729 on 29/04/2010, with the latter value being 2.87 times the standard deviation away from the mean value of SD $P_{GBP/USD}$ which means this value is less than 1% chance of being observed through the whole period. It is also the highest SD value across $P_{USD/GBP}$.

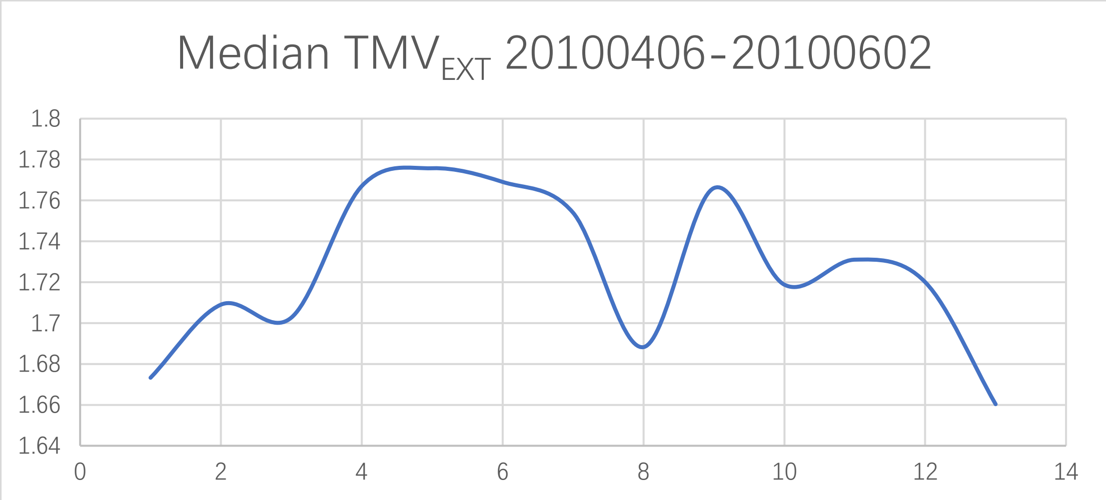

Figure 4. 29 Median TMV_{EXT} from 06/04/2010-02/06/2010

Meanwhile, from Figure 4. 21 we can see under nearly the same period as the TS time period above, the value of Median TMV_{EXT} under DC is more or less the same where the average value is 1.7241 where the highest point is 1.7757 and the lowest point is 1.6603. These suggest that at the same period, significant changes are not observed under DC.

4.9.5. Summary of all the GBP/USD results

From the observations above, three observations on market volatility under both TS and DC emerged

1. Median TMV_{EXT} observed high volatility in $P_{GBP/USD}1$

2. A big jump in SD is observed in $P_{GBP/USD}2$

These results are consistent with our analysis in EUR/USD: as expected, each of TMV_{EXT} and SD observed volatility periods not observed by the other indicator. This shows that neither TMV_{EXT} nor SD can be replaced by the other.

4.10. Conclusion

The DC indicators NDC and median $aTMV_{EXT}$ (Section 0) and the TS indicator SD (Section 4.4.2) measure different aspects of market volatility. In this chapter, all three indicators were used to inspect market volatility and to examine in detail the observations possible for historic data (Section 4.5.1). By comparing and contrasting these observations, this chapter demonstrates that each indicator could observe useful information that could not be observed by the other two indicators, and by using multiple indicators, a deeper understanding of a market's volatility can be developed.

Summarising the observations under SD (Section 4.6.5), median $aTMV_{EXT}$ (Section 4.6.6) and NDC (Section 4.6.7) over four periods (Table 4. 1, Section 4.6.4) showed that these indicators support each other in some cases (Section 4.7.1). However, each may pick up signals that are not picked up by the other indicators (see Table 4. 2 at the beginning of Section 4.7). For example, the significant volatility changes observed under DC's median $aTMV_{EXT}$ could not be observed under TS (Section 4.7.3). The DC NDC could also pick up significant volatility changes that were not observable under TS (Section 4.7.4), while TS could also pick up volatility changes that were not picked up by DC (Section 4.7.2).

In Section 4.8 and Section 4.9 we repeat our experimental work using the approach described in Section 4.4. Results from the observations in Section 4.8 and Section 4.9 are consistent with those in EUR/USD (Sections 4.6 and 4.7).

Based on these results, it is clear that TS and DC are useful for tracking changes in the market, and that they complement each other and can thus support each other (Section 4.7.1, Section 4.8.6 and Section 4.9.5), as each can find information that the other cannot (Sections 4.7.3, 4.7.4, 4.7.2, 4.8.3, 4.8.4, 4.9.3 and 4.9.4). Different indicators appear to pick up different volatility signals from the market. The findings in this chapter further demonstrate that it is beneficial to use multiple indicators to study market volatility. We believe that all three indicators should be used together.

5. Chapter five: Nowcasting new trends under DC

A Directional Change event can only be confirmed in hindsight. In this chapter, we propose a novel algorithm to nowcast whether a new DC trend has already begun before a formal confirmation point happened.

5.1. Background

Nowcasting refers to recognizing what the current state is. In our context, we know that in DC, trends can only be recognized in hindsight. As shown in Figure 5. 1, we only know that the downtrend from EXT1 to EXT2 has ended at the DC confirmation (DCC) point of the new trend, DCC2. The price at DCC2 is θ above the lowest point, EXT2, of the previous downtrend (which is also the beginning of the current uptrend). The question is: could one recognize at any point after EXT2 and before DCC2 that the new trend has started? That is the task of nowcasting.

The goal in this chapter is to develop a method to help us recognize that a new trend has already begun before the new conformation point. What indicators could we look at? What counts as a success in nowcasting? How to measure the performance of a nowcast?

The goal in this chapter is to recognize the end of a trend as soon as possible. In the example in Figure 5. 1, we want to recognize the end of the downtrend after t1 before t2.

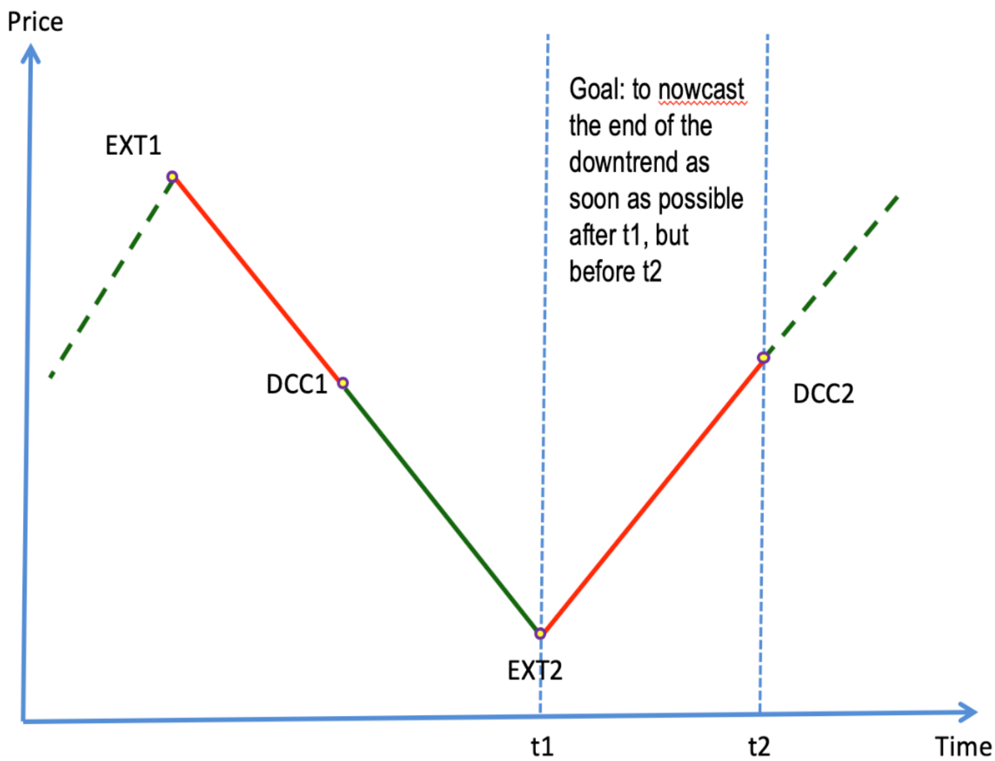

Figure 5. 1 The goal of this chapter in Figure

5.2. New indicators, aTMV, Max TMV and Below Max

In this section, we will introduce three new indicators that are used to nowcast under DC.

First of all, we will introduce some basic ideas, according to Chapters 2 and 3, we have already introduced what is TMV_{EXT} and DCC.

Tsang et al. (2017) defined that total price movements value at extreme points (TMV_{EXT}) measures the price distance between the extreme points that begin and end a trend, normalized by θ, where θ is the threshold used for generating the directional change summary. TMV_{EXT} is defined by:

$$\boldsymbol{TMV_{EXT_i}} = \frac{\boldsymbol{P_{EXT i+1} - P_{EXT\_i}}}{\boldsymbol{P_{EXT\_i} * \theta}} \tag{5.1}$$

Here P_{EXT_i} represents the price at the i-th directional change extreme point, P_{EXT_i+1} represents the price at the (i+1)-th directional change extreme point, θ is the threshold used

Directional Change Confirmation (DCC) Point is the point at which to confirm one DC event. We named P_{DCC} and T_{DCC} as the price and time at the DCC point.

5.2.1. Indicators for DC nowcasting

In this section, we introduce some indicators for nowcasting under DC.

First, we call the maximum price in an uptrend and the minimal price in a downtrend the Max Price (P_{Max}).

Definition 5. 1:P_{Max}

P_{Max} is the maximal transaction price in an uptrend and the minimal transaction price in a downtrend.

Definition 5. 2: Absolute TMV (aTMV)

The aTMV of a price P is the absolute TMV from the extreme point (P_{EXT_i}) to P.

$$aTMV(P) = \left|\frac{P-P_{EXT\_i}}{P_{EXT\_i}*\theta}\right| \quad (5.2)$$

Here, θ is the threshold used, and the current price (p) represents the price at this point. P_{EXT_i} represents the price at the i-th directional change extreme point.

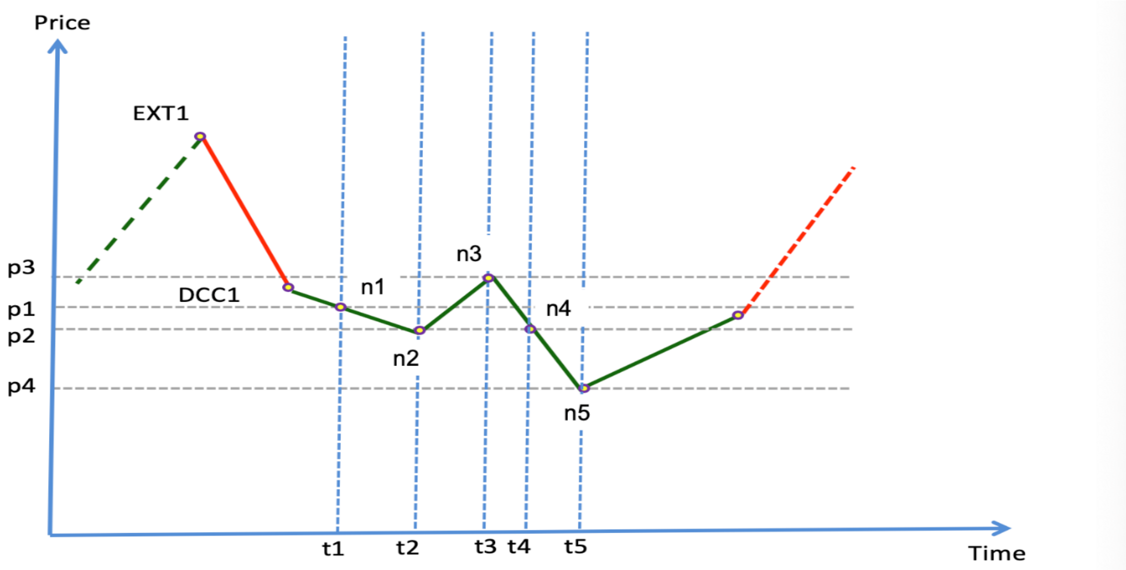

Figure 5. 2 Max

We calculate aTMV using 0In Figure 5. 2 every point after the Directional Change confirmation in a trend, we will be started calculating aTMV. At any point n, the aTMV is monitored. At any time during a trend, the point with the maximum aTMV is called Max. For example, n2 is a Max. From n2, prices bounced back to n3. Therefore, from t2 to t4, n2 will be recorded as the current Max. When the price went to n5, we record a new Max.

Figure 5. 4 provides an example of how to calculate aTMV in DC. Here, the P_{EXT1} is equal to 1.08727, and the current price at time 18:04:30 on 03/01/2016 is 1.08536. In this case, the aTMV value at the 1.08536 marks

is 1.097933, calculated as $\left|\frac{1.08727-1.08536}{1.08727*0.0016}\right|$.

5.2.2. Max

Max represents the max value of aTMV that happens within a trend. With the price moving in a trend, Max should change follow the price change.

Definition 5. 3: Max

Max is the aTMV of P_{Max}. In other words, Max is the maximum aTMV recorded so far in a trend.

$$Max = (|P_{Max} - P_{EXT}| / P_{EXT}) / \theta \tag{5.3}$$

Figure 5. 2 shows what Max is. Therefore, from t2 to t4, n2 will be recorded as the current Max. At t5, Max is updated to n5, as it is the lowest point in the downtrend so far.

In Figure 5. 4, at the point of 1.08536, while the aTMV value is 1.097933, this is the Max before the 1.08536. With the continuous monitor price movement, at 18:04:34 on 03/01/2016, the current price become 1.08523, while at this point, the aTMV value is 1.1726618. At this point, Max becomes 1.1726618 in this trend.

5.2.3. Below Max

Below Max (BM) is an indicator that measures how significant the price bounds back from the Max TMV in a trend. This is a dynamic indicator with the move of P_{MAX} change.

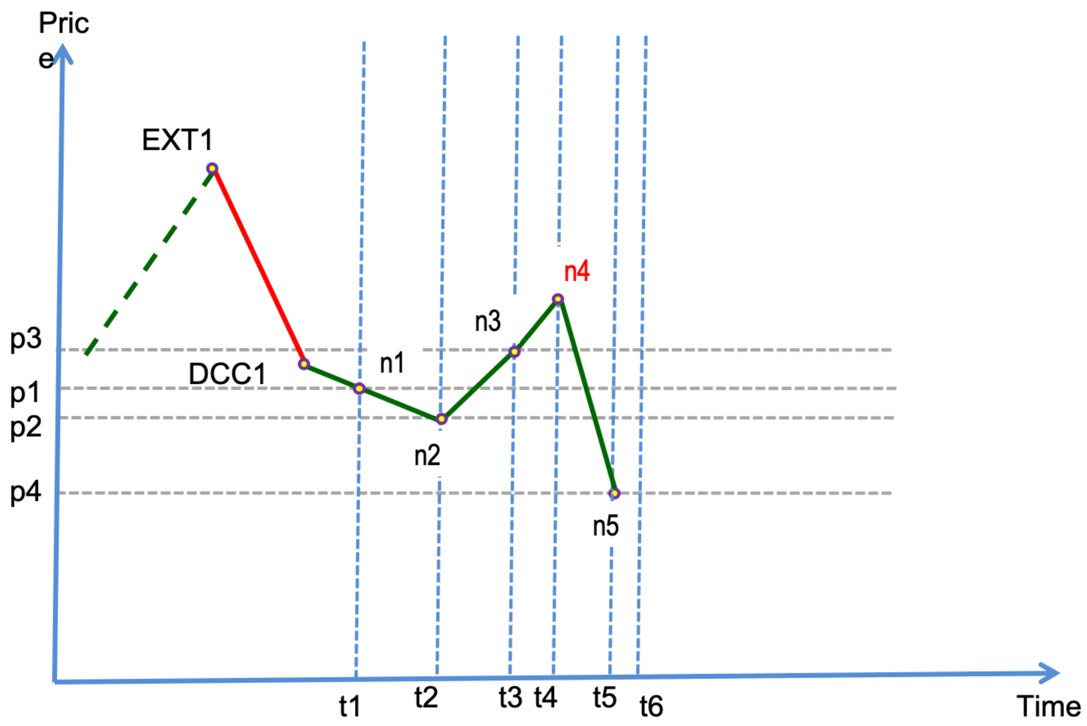

Figure 5. 3 Below Max and Maximum Below Max

Figure 5. 3, shows how Below Max and Maximum Below Max are defined: n3 bounces back from n2. We calculate the amount of bounce back according to formula 5.3. For example, n3 could be 0.25 below n2. We call it Below Max. Suppose the price bounces back to n4 before it turns the other direction to n5. We record the Maximum Below Max at n4. Below Max is show how significant the price bounds back in a trend. From t2 to t4, every point n will calculate a Below Max.

Definition 5. 4: Below Max (BM)[8]

[8] BM is called undershoot by Raju Chinthalapati and Han Ao (personal communication)

Below Max (BM) is the absolute contrarian price change percentage since Max, normalized by the threshold θ. For example, if P_{Max} records the minimal price in a downtrend, then BM is the normalized price increase above Max; if P_{Max} records the maximal price in an uptrend, then BM is the normalized price drop below Max.

$$BM = \left|\frac{P_{MAX} - Currentprice}{P_{MAX} * \theta}\right| \tag{5.4}$$

For example, from Figure 5. 4, as we mentioned in section 5.2, when Max TMV is 1.1726618, at that time, the P_{MAX} is 1.08523. At the price point of 1.08542, we can calculate the BM as the absolute value of $\left|\frac{1.08523-1.08542}{1.08523*0.0016}\right|$ which is equal to 0.011517.

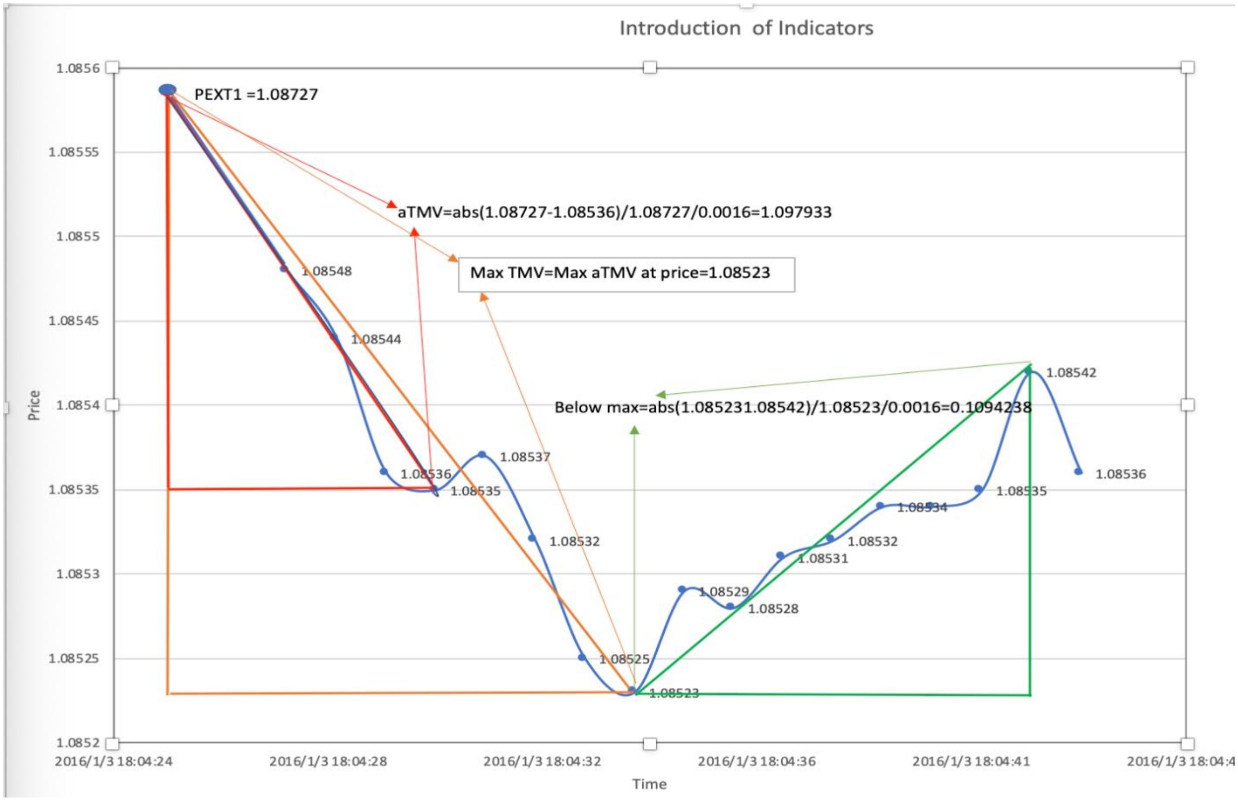

Figure 5. 4 The aTMV, Max and Below max values calculated by the EUR/USD data from 18:04:24 3/1/2016 to 18:04:45 3/1/2016 under threshold 0.0016.

Figure 5. 4 offers an example of how we calculate the aTMV, Max and Below max value. The calculation of aTMV follows 0where P is equal to 1.80727 and P_{EXT_i} is 1.08536. We use 0.0016 as our threshold here. The Max TMV, following the 0and we here use P_{EXT_i} is equal to 1.08523 while P is still 1.80727. We also follow 0to calculate the BM value. Here the current is 1.08542, and P_{MAX} is 1.08523 given the threshold of 0.0016. We calculate the BM value as 0.1094238.

5.3. Data Used and their profiles

To prepare for our nowcasting algorithm, we shall study the distributions of aTMV and introduce additional indicators. We use historical data to study these distributions. In this section, we introduce the data used.

Tick-to-tick EUR/USD rates are used throughout this chapter. Table 5. 1 summarizes the data used in this chapter. We separate the data into the Training and Nowcasting periods. Statistics in the former is used to determine the parameters used in our nowcasting algorithm (to be presented), which will be used for nowcasting in the Nowcasting period. To examine the generality of our nowcasting algorithm, we employ two thresholds, namely 0.0016 and 0.0032.

	Data in the Training Period		Data in the Nowcasting Period	
Threshold	0.0016	0.0032	0.0016	0.0032
Period	From: 00:00:10 25th September 2009 To: 20:08:53 31st December 2013		From 01:30:09 1st January 2014 To: 13:02:55 31st December 2015	
Number of transactions	72,629,464		33,041,496	
Number of DCs	9,552	2,052	4,644	1,185
Median aTMV	1.612047	1.686975	1.596844	1.436107
Median T	2,392	12,801	1,263	4,817

Table 5. 1 Tick-to-tick data used in this chapter, with their profiles compared

We call the period from 00:00:10 on 25/09/2009 to 20:08:53 on 31/12/2013 as the **Training Period**. We call the period from 01:06:38 on 01/01/2014 to 13:02:55 on 31/12/2015 the **Nowcasting Period**.

Let us compare and contrast the profiles of the two periods. From Table 5. 1. we can see that, under threshold 0.0016, the median aTMV values in the Training Period is 1.612047, whereas the median aTMV in the Nowcasting Period under the same threshold is 1.596844. Under threshold 0.0032, the median aTMV values in the Training and Nowcasting periods are 1.686975 and 1.436107. Under both thresholds, the median aTMV in the Training Period is slightly higher, but not significantly higher.

But when we look at the Median T values, we can see that under threshold 0.0016, a trend takes 2,392 seconds to finish a trend in the Training Period, but 1,263 seconds (nearly half of the time) to finish in the Nowcasting Period. Under threshold 0.0032, the median T values are 12,801 and 4,817

(38%) in the Training and Nowcasting Periods, respectively.

Both aTMV and T are indicators of volatility in a market-period: the former measures the magnitude of price changes and the latter measures the frequency of changes. The above comparisons suggest the following:

1. There is not much difference in the two periods' volatility as measured by median aTMV.

2. Measured by median T, the Nowcasting Period is much more volatile than that Training Period: direction changes much faster in the Nowcasting Period.

3. Observations under the two thresholds are consistent.

To summarize: the two periods are quite different in their volatility, though not in the magnitude of price changes in the trends (measured by aTMV), but in the frequency of directional changes (measured by T).

5.4. Distribution of TMV_{EXT}

In this section, we study the historical distribution of TMV_{EXT}. We shall show later how this could be used to support nowcasting.

As explained in Table 5.1, we use the EUR/USD tick data from 2009/09/25 00:00:10 to 2013/12/31 20:08:53. We named this period the Training Period. Also, we shall use EUR/USD tick data from 01:30:09 on

01/01/2014 to 13:02:55 on 31/12/2015, we named this period the Nowcasting period. We calculate TMV_{EXT} under thresholds 0.0016 and 0.0032.

We calculate the TMV_{EXT} under 0.0016 and 0.0032 for both Nowcasting and Training Periods. Then we collect it from every trend. In the Training period when the threshold is equal to 0.0016, we got 9,552 TMV_{EXT}, while when the threshold is equal to 0.0032 we got 2052 TMV_{EXT}. In the Nowcasting period, when the threshold is equal to 0.0016, we got 4,644 TMV_{EXT}, while when the threshold is equal to 0.0032 we got 1,185 TMV_{EXT}.

Once we collect all TMV_{EXT} for every period under both thresholds. We will start to summarize the TMV_{EXT}. At first, we will sort it from the smallest to the biggest and calculate all TMV_{EXT} to absolutely value. For instance, when the threshold is equal to 0.0016 and under the Training period, we know the biggest value is 10.575712 (10.575712 from Table 5. 2), then we minus one divided by 100. Which is (10.575712-1)/100, in this case, each bin value becomes 0.09575712. The reason we minus one, is that TMV_{EXT} only happened after the DCC point which the theoretical TMV is 1. And, we divided by 100, is that we want to get the result of distribution in the precision of 0.01. Observers can choose the desirable precision by themselves. Table 5. 2, Table 5. 3, Table 5. 4 and Table 5. 5

will show the Distribution of TMV_{EXT} under the Training period and the Nowcasting period under thresholds 0.0016 and 0.0032. All details probability table please see in Section 7 Appendix part.

Here we give the example to read Table 5. 2, Table 5. 3, Table 5. 4 and Table 5. 5. For example, when TMV_{EXT} is equal to 1.28727136, we can find in Table 5. 2 that historically, the probability for this value happened in the Training period is 7.63%, and the probability of reaching this TMV_{EXT} value is 71.45%. This means, that in the training period, there is a 71.45% probability (historically) that the TMV_{EXT} value happened is bigger than 1.28727136. We will use only the distributions of the TMV_{EXT} Training period under 0.0016 and 0.0032, the distribution of the TMV_{EXT} Nowcasting period will it be as an objective of reference to confirm the difference between the two data set.

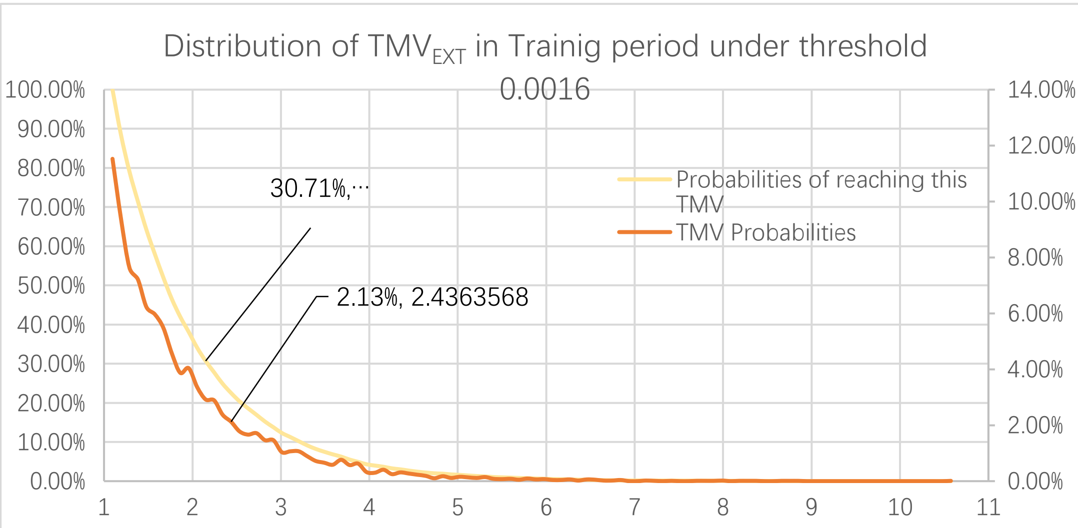

Figure 5. 5 Distribution of TMV_{EXT} in the Training period under threshold 0.0016

Figure 5. 5 show the graph of the probability of TMV_{EXT} happening in each

bin and the probability of reaching a certain number of TMV_{EXT} in the Training period when the threshold is equal to 0.0016. Here the x-axis represents the TMV_{EXT} value (the TMV value at the extreme point) and the y-axis represents the percentage of trends that end with a TMV_{EXT} value. For example, in Table 5. 2, we can read that when TMV_{EXT} is 2.4363568, we can find out that there is a 2.13% probability to reach this historical probability (orange line). At the same time, when we look at the yellow line, we can also find out that around 30.71% of the TMV_{EXT} are bigger than 2.14908544. on the other hand, we can say that 30.71% of the TMV_{EXT} do not reach the value of 2.14908544.

The of TMV_{EXT} in the Training period under threshold 0.0016 that we pick			
Bin	TMV_{EXT} Frequency	TMV_{EXT} Probability	Probabilities of reaching this TMV_{EXT}
1	-	-	100.00%
1.09575712	1101	11.53%	88.47%
……			
1.28727136	729	7.63%	71.45%
1.38302848	689	7.21%	64.24%
……			
1.67029984	524	5.49%	46.56%
1.68	**-**		**46.22%**
1.76605696	437	4.57%	41.98%
……			
10.575712	1	0.01%	

Table 5. 2 The TMV_{EXT} we pick to use under the distribution of TMV_{EXT} in the Training period under threshold 0.0016

When the threshold is equal to 0.0016, we will use the TMV_{EXT} value 1.68 which would historically have a probability of 46.22% happening in the Training period to nowcast the Training and Nowcasting period, which will be used in Section 5.8. This value is picked arbitrarily. No optimization is attempted in picking this value.

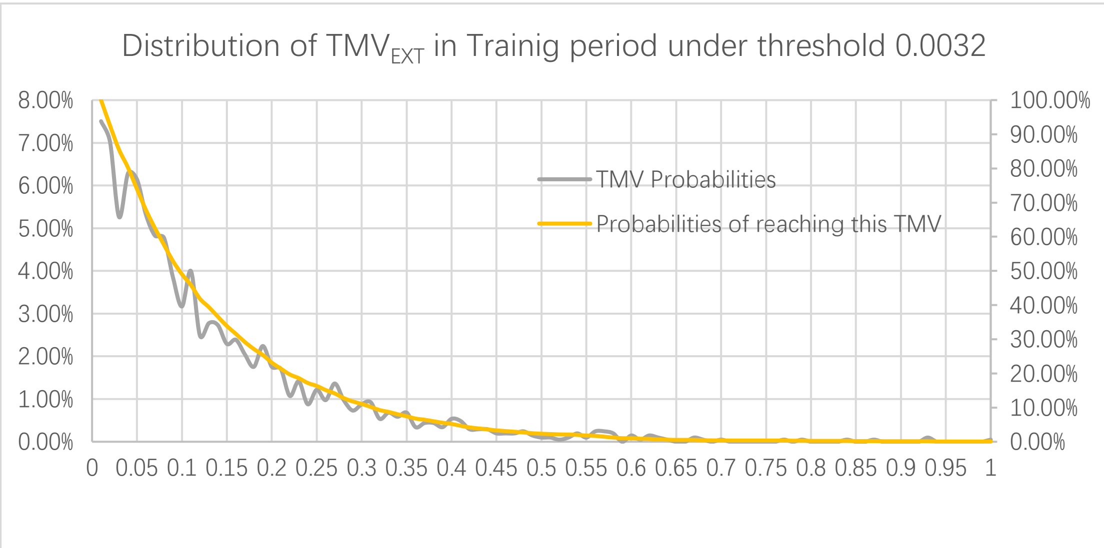

Figure 5. 6 Distribution of TMV_{EXT} in the Training period under threshold 0.0032

Figure 5. 6 show the graph of the probability of TMV_{EXT} happening in selected bins and the probability of reaching a certain number of TMV_{EXT} in the Training period when the threshold is equal to 0.0032

Distribution of TMV_{EXT} in the Training period under threshold 0.0032			
Bin	TMV_{EXT} Frequency	TMV_{EXT} Probabilities	Probabilities of reaching this TMV
1	-	-	100.00%
1.07790733	154	7.50%	92.50%
……			
1.54535131	99	4.82%	57.65%
1.60	**-**		**54.40%**
1.62325864	98	4.78%	52.88%
……			
8.790733	1	0.05%	

Table 5. 3 Distribution of TMV_{EXT} in Training period under threshold 0.0032.

When the threshold is equal to 0.0032, we will use the TMV_{EXT} value 1.60 to nowcast the Training and Nowcasting period as the parameter to do experimental work in Section 5.8. When the TMV_{EXT} value riches 1.60, these would have a 46.22% probability happening in the Training period. Again, this value is picked arbitrarily. No optimization is attempted.

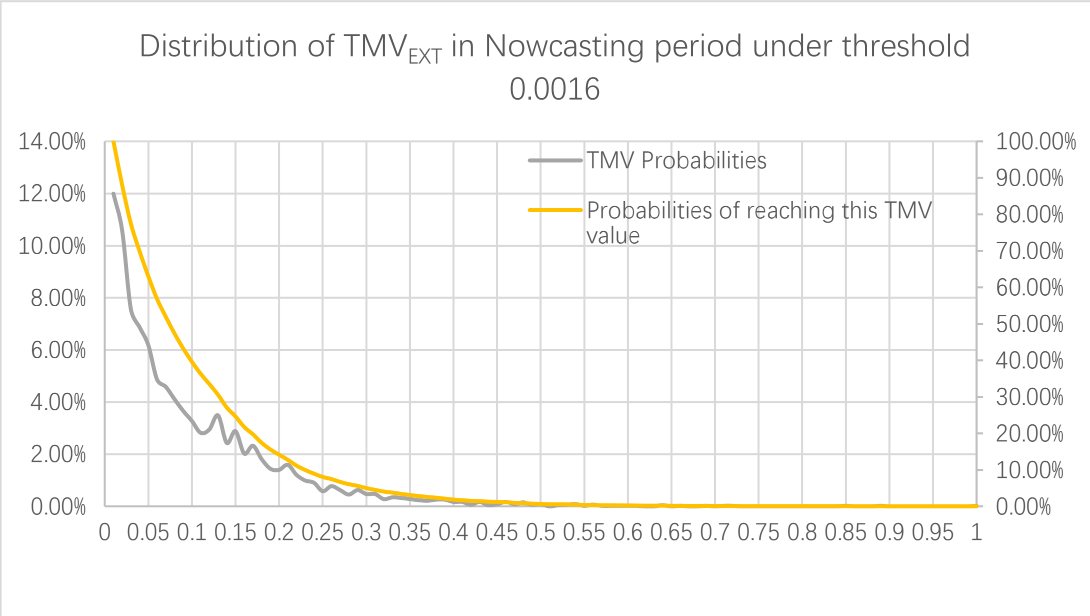

Figure 5. 7 Distribution of TMV_{EXT} in Nowcasting period under threshold 0.0016

Figure 5. 7 show the graph of the probability of TMV_{EXT} happening in each bin and the probability of reaching a certain number of TMV_{EXT} in the Nowcasting period when the threshold is equal to 0.0016.

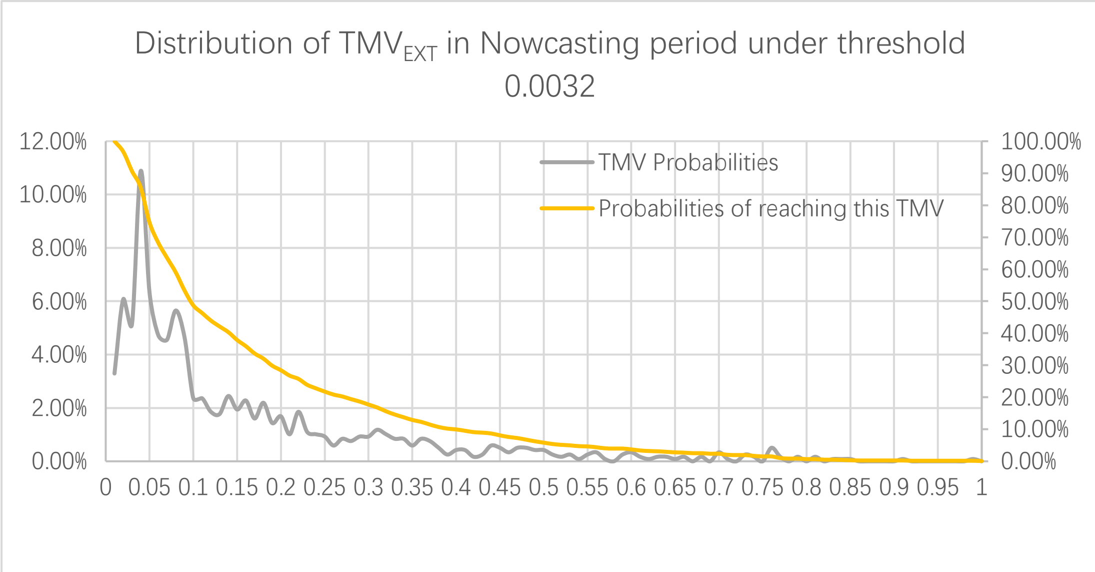

Figure 5. 8 Distribution of TMV$_{EXT}$ in Nowcasting period under threshold 0.0032

Figure 5. 8 show the graph of the probability of TMV$_{EXT}$ happening in each bin and the probability of reaching a certain number of TMV$_{EXT}$ in the Nowcasting period when the threshold is equal to 0.0032

The reason that we show Figure 5. 7 and Figure 5. 8 is we want to show the distribution graph about the absolute value of TMV$_{EXT}$ in the Nowcasting period. This would give the observer a direct vision of how the profile of the Nowcasting period is compared to that of the Training period.

5.5. Max Below Max (MBM) and its value distribution

In section 5.2, we provided the definition of BM. In this section, we will introduce a new indicator called Max Below Max (MBM). We shall also look at the historical distribution of MBM values. We shall show later how these could be useful for nowcasting.

5.5.1. Max Below Max (MBM)

Recall that BM is the normalised absolute contrarian price change percentage since Max, the lowest (highest) price in a downtrend (uptrend) (Definition 5. 3 and Definition 5. 4).

Definition 5. 5: Max Below Max (MBM)

The MBM is the maximum aTMV found below BM in the current trend.

In Figure 5. 3, From t2 to t4, every point n will calculate a Below Max. and n3, is the maximum value of Below max, we called Max below Max. In every trend, we will at least have one MBM. MBM shows what is the maximum bounce back value in a trend at the current price.

MBM is used to represent the level of price bounce back. The bigger the MBM is, the higher probability of DC will enter a new trend. We will show the distribution of MBM in the Training period and the Nowcasting period later. The same as the distribution of the absolute value of TMV_{EXT} in Section 5.4, the reason we show the distribution of MBM in the Nowcasting period will be as an objective of reference to confirm the difference between the two data set.

5.5.2. Distribution of Max Below Max (MBM)

We will use the Training period and the Nowcasting period to summarize the distribution of MBM under thresholds 0.0016 and 0.0032. In our experimental work, we only use the distribution of the Training period in both thresholds. the reason we have also shown the graph of the distribution of MBM in the Nowcasting period is that we want to give observers a directly view about how the graph would like to compare with the Training period. We got 563,499 MBM value under the Training period when the threshold is equal to 0.0016. When the threshold is equal to 0.0032, we got 23,445 MBM value. In the Nowcasting period, under threshold 0.0016, we got 220,300 MBM while under threshold 0.0032 we got 91,347 MBM value.

All MBM values are from 0 to 1, we put 0 to 1 into 100 bins which are 0.01 for each bin, we will have 100 bins from 0 to 0.01 to 0.99 to 1. Then we count for each bin, how many BMB values is in this bin. And how many BMB values are bigger or small than this bin's value. For instance, from Table 5. 4, we can see that nearly 1/4 (25.01%) MBM have happened from 0.04 to 0.05, and 10.86% of the MBM value are from 0.09 to 0.1. If an MBM value is 0.06, from Table 5. 4, we can find that there is a 64.26% chance that the MBM value is higher than 0.06.

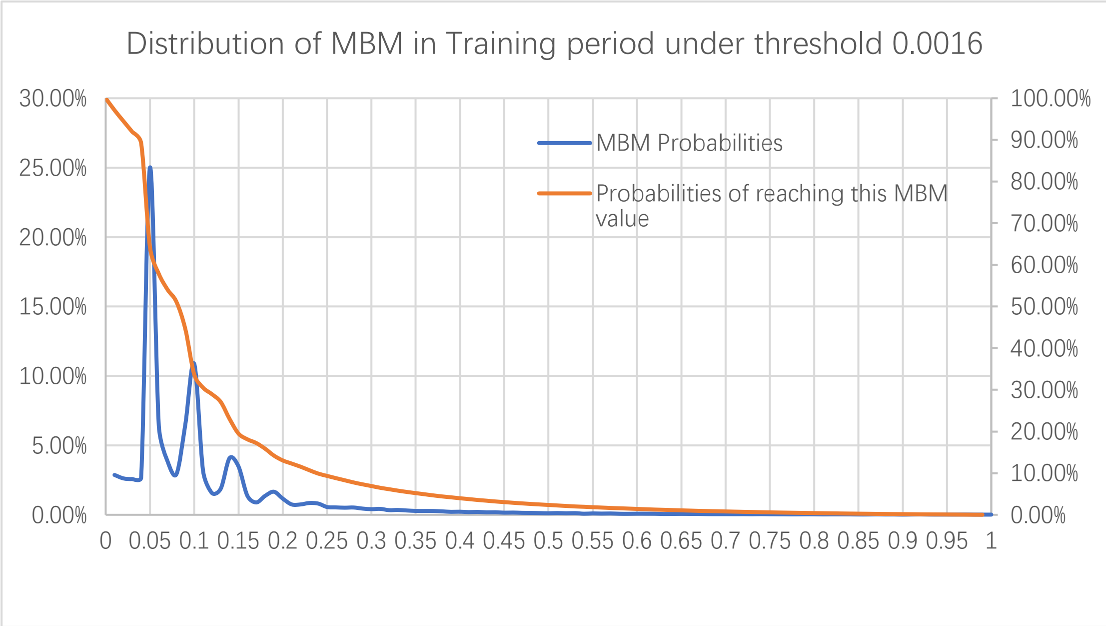

Figure 5. 9 Distribution of MBM in Training period under threshold 0.0016

Figure 5. 9 show the graph of the probability of MBM happening in each bin and the probability of reaching a certain number of MBM in the Training period when the threshold is equal to 0.0016.

Distribution of Max Below Max in Training period under Threshold 0.0016			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	16168	2.87%	100.00%
……			
0.67	339	0.06%	1.00%
0.68	**273**	**0.05%**	**0.95%**
……			
1	72	0.01%	0.01%

Table 5. 4 Distribution of Max Below Max in Training period under Threshold 0.0016

When the threshold is equal to 0.0016, we will choose 0.68 as the parameter to nowcast the Training period and the Nowcasting period (Section 5.8). This is because we only got a 0.95% probability that BM is bigger than 0.68. We will use this probability to nowcast both training and nowcasting data under 0.0016.

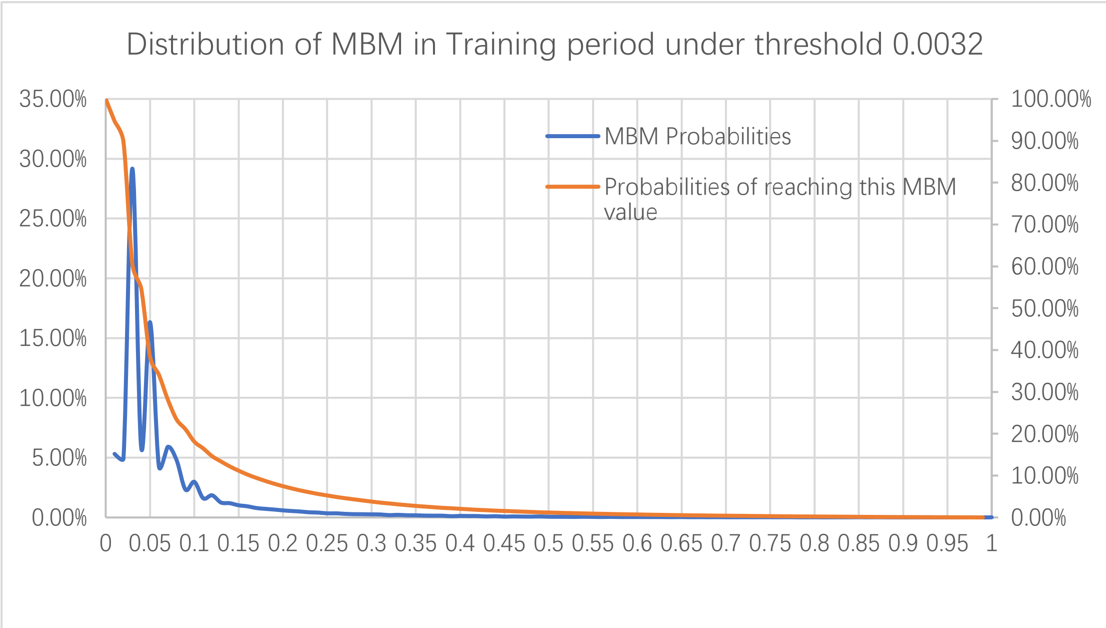

Figure 5. 10 Distribution of MBM in Training period under threshold 0.0032

Figure 5. 10 show the graph of the probability of MBM happening in each bin and the probability of reaching a certain number of MBM in the Training period when the threshold is equal to 0.0032.

Distribution of Max Below Max in Training period under Threshold 0.0032			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	12461	5.32%	100.00%
……			
0.6	87	0.04%	0.75%
0.61	**84**	**0.04%**	**0.71%**
0.62	84	0.04%	0.68%
0.63	78	0.03%	0.64%
……			
1	24	0.01%	0.01%

Table 5. 5 Distribution of Max Below Max in Training period under Threshold 0.0032

When the threshold is equal to 0.0032, we will choose 0.61 as the parameter to nowcast the Training period and the Nowcasting period. This is because, historically, we only have a 0.71% probability of BM bigger than 0.61. We will use these BM values to nowcast the training and nowcasting data under 0.0032 in section 5.8.

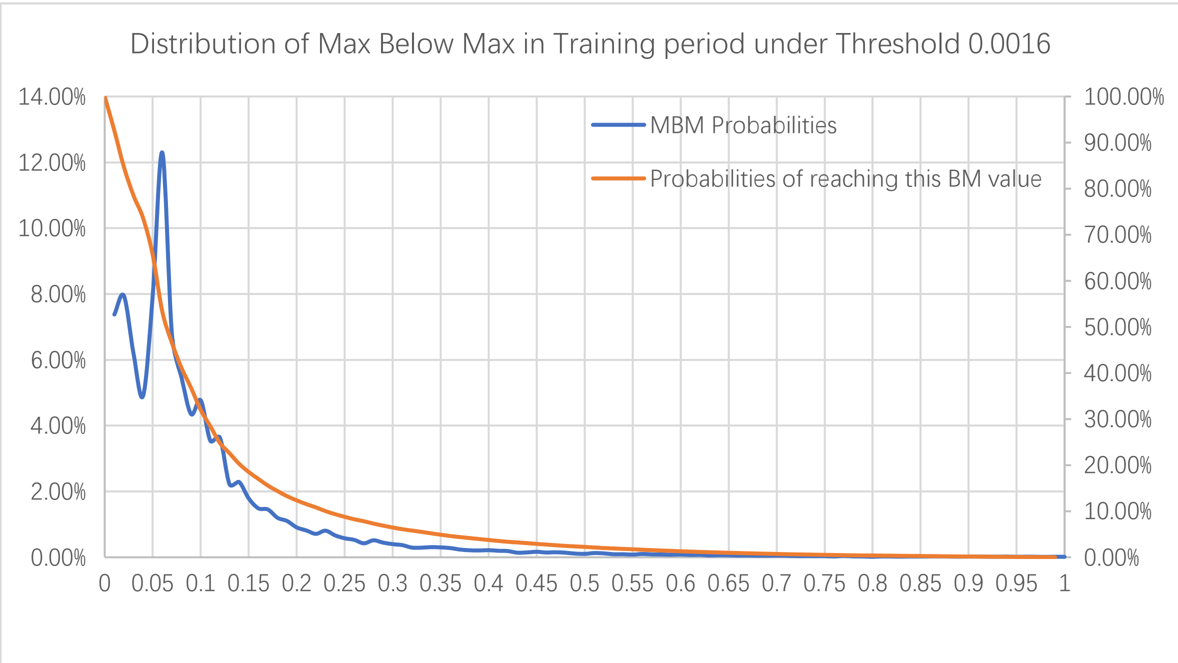

Figure 5. 11 Distribution of MBM in Nowcasting period under threshold 0.0016

Figure 5. 11 shows the probability of MBM occurring in each bin and the

probability of reaching a certain number of MBM in the Nowcasting period when the threshold is equal to 0.0016.

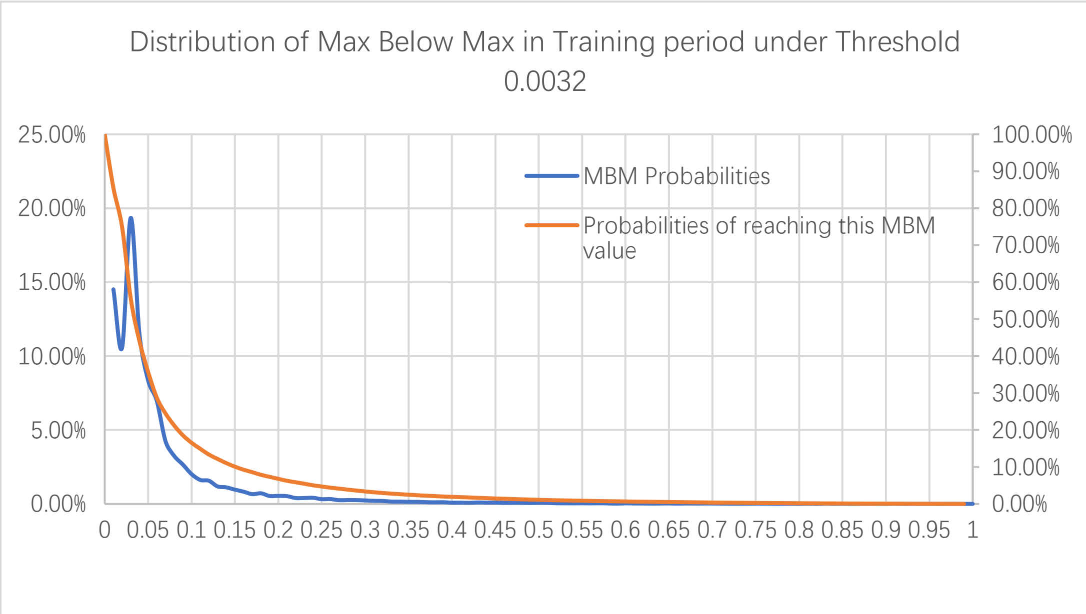

Figure 5. 12 Distribution of MBM in Nowcasting period under threshold 0.0032

Figure 5. 12 shows the probability of MBM occurring in each bin and the probability of reaching a certain number of MBM in the Nowcasting period when the threshold is equal to 0.0032.

The reason that we graph Figure 5. 11 and Figure 5. 12 is we want to show the distribution graph about MBM in the Nowcasting period. This would give the observer a direct vision of how the profile of the Nowcasting period is compared to that under the Training period.

5.6. Methodology

With the historical distribution of aTMV and MBM studied, we are now in a position to introduce an algorithm for nowcasting the beginning of a new trend. We shall also explain how we assess the performance of this algorithm.

5.6.1. Overview of Nowcasting

To recapitulate, a trend ends at an extreme point (EXT), but the end of a trend is only confirmed in hindsight when we reach the DC confirmation point (DCC) of the new trend. Our objective is to nowcast the end of the preceding trend as soon as possible after EXT but before the DCC of the new trend.

To nowcast, we monitor the market tick by tick. We calculate the aTMV (Section 5.4) and Below Max (Section 5.5) at every tick. We nowcast that we are in a new trend as soon as aTMV reaches P_{TMV} and BM reaches P_{BM}, where P_{TMV} and P_{BM} are parameters. We call this the Nowcast Constant Algorithm (NCA):

> **Nowcast Constant Algorithm (P_{TMV}, P_{BM}):**
>
> Given a current transaction *ct*, if and only if
>
> (1) aTMV(*ct*) $\geq P_{TMV}$ and
>
> (2) BM(*ct*) $\geq P_{BM}$,
>
> then we nowcast that *ct* is in a new trend.

Note that P_{BM} cannot be bigger than one, as once BM is bigger than one, a new trend has already begun.

Here the parameter P_{TMV} measures the distance of *ct* from the start point of the trend. The bigger the P_{TMV} is the bigger chance that we have already entered a new trend. P_{BM} measures how significant that ct has bounced back from the last P_{Max}. The bigger this value, the higher chance that we are already in a new trend. When both aTMV(*ct*) and BM(*ct*) are high, there is a high probability that we are already in a new trend.

Any P_{TMV} and P_{BM} values may be selected, subject to sensible ranges. But they should be within sensible ranges, as explained below: As we are trying to determine whether *ct* is in the current trend or a new trend, TMV(*ct*) must be greater than or equal to 1. This is because, by definition, the TMV of the theoretical DC confirmation point (DCC*) is 1. Therefore, $1 \leq P_{TMV}$. Theoretically, there is no upper limit in the value of TMV(*ct*). Therefore, the range of P_{TMV} is:

$$1 \leq P_{TMV} < \infty$$

By definition, BM at Max is 0; therefore $0 \leq P_{BM}$. When BM(*ct*) is 1, we have already confirmed that that latest Max was an extreme point (based on the definition of DC and TMV). In other words, we already know that ct is in a new trend. Therefore, a sensible range of P_{BM} is:

$$0 \leq P_{BM} < 1.$$

5.6.2. Criteria for assessing nowcasting performance

When we nowcast a new trend, we are suggesting that Max was an extreme point that ended the previous trend and started the current trend. In this section, we define criteria for assessing when a nowcast is good or no good.

A good nowcast is one that takes place after the end of the previous trend (i.e., Max was indeed a DC extreme point) and before the DC confirmation point.

Definition 5. 6: Correctness of a nowcast

A Nowcast is only *correct* if it happens after EXT2.

Definition 5. 7: Usefulness of a nowcast

A Nowcast is only *useful* if it is before the DCC of the new trend.

Definition 5. 8: Good and nogood nowcast

We define a nowcast to be "*good*" if it is correct and useful.

A nowcast is "*nogood*" if it is either incorrect or not useful.

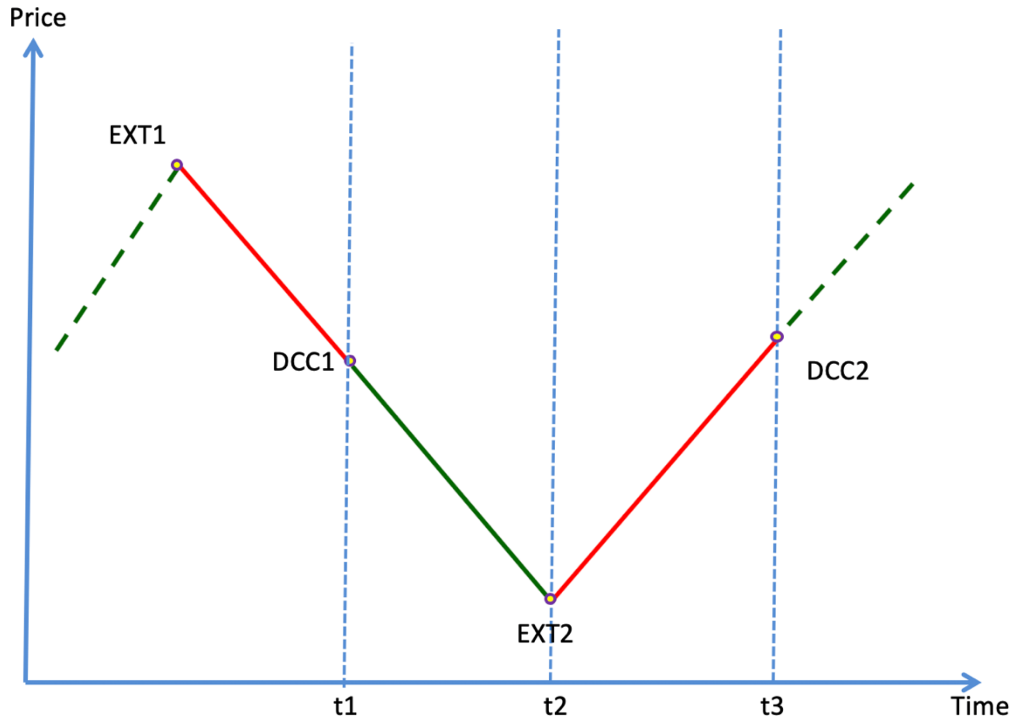

Figure 5. 13 Any nowcast from t2 to t3 is correct and useful (i.e., good). Any nowcast before t2 is incorrect. Any nowcast after t3 is not useful

Figure 5. 13 shows that if we nowcast a new trend before t2, then our nowcast is incorrect (because a new trend has not started before t2). If we nowcast a new trend after t2, then the nowcast is correct. However, if we nowcast after t3, then the nowcast is not useful (because the new trend is already confirmed by DCC2 at t3). A nowcast is good if it is correct and useful (Definition 5.8). Therefore, only nowcasts made between t2 and t3 are good.

Good nowcast (GN) is set as an indicator that represents the percentage of good nowcast trends compared with the total number of trends in the time period. It is an indicator that can measure how many of the trends we can nowcast correctly before the DCC point compares with the number of trends. The higher the GN value is, the better the performance for the nowcast in that period under the specified threshold and parameters we choose. This indicator is newly introduced in this section.

$$GN = \frac{Number\ of\ Good\ nowcast}{Number\ of\ DC} \tag{5.4}$$

The total number of trends (TnT) shows the total number of trends (NDC–1) in a time period under a certain threshold. Every trend may or may not have a nowcast, which depends on the parameter we choose for nowcast.

The number of nowcasts (Nc) counts how many nowcasts we have made in a certain period under a certain threshold with the parameters we choose. Once we have the Nc, we will want to summarize the Number of correct nowcasts (CN). Here, CN represents the number of correct nowcasts which is shown in Section 5.6. Once we know the number of correct nowcasts, we also know the incorrect nowcasts (IN).

Naturally, Nc = CN + IN.

Nowcast time (TimeNC) is an indicator that represents the time of the nowcast point in the trend.

TMV at Nowcast (TMV_{NC}) when we confirm a nowcast happened we calculate the TMV value at that point, we call it TMV_{NC}.

In the subsections below, we shall introduce criteria for measuring nowcast performance.

5.6.2.1 Precision

This section introduces a method to answer two questions. The first is what percentage of nowcasts are correct, where being correct means the nowcast was made after a new trend had started (Definition 5. 6). Of all the correct nowcasts, we want to know the percentage that is good (which means they are made before the DC confirmation point, Definition 5. 7).

$$Precision_{Correct} = CN \div Nc$$

$$Precision_{Good} = GN \div CN$$

False Positive in correctness:

$$FP_{Correct} = 1 - Precision_{Correct} = IN \div Nc$$

5.6.2.2 Timeliness

For those good nowcasts, we measure their performance by how close they are to EXT2, in terms of time and price. We use TMV_{DCC} to denote the TMV measured at the DC Confirmation point. By definition (Chen & Tsang 2020), TMV_{DCC} is normally close to 1. Following are measures of timeliness, in terms of time and price:

$$Timeliness_{Time} = DCC_{time} - Time_{NC}$$

$$\text{Timeliness}_{Price} = \text{TMV}_{DCC} - \text{TMV}_{NC}$$

In this Chapter, Timeliness_{Time} is measured in seconds. The bigger the Timeliness_{Time} value is, the sooner we nowcast correctly before DCC.

$\text{Timeliness}_{Price}$ is an indicator measuring how soon we nowcast before the DCC price. For example, suppose in an uptrend current price is 100, and we use a threshold of 10%. In this case the theoretical P_{DCC*} is 110 (100*(1+10%)) (Calculate follow the Section 2.3). Suppose the actual DCC price P_{DCC} is also 110 and we nowcast at the price of 106. TMV_{NC} = (((106-100)/100)/10%=) 0.6. That means $\text{Timeliness}_{Price}$ is equal to (1–0.6=) 0.4. This means we nowcast at 40% of the price before P_{DCC}.

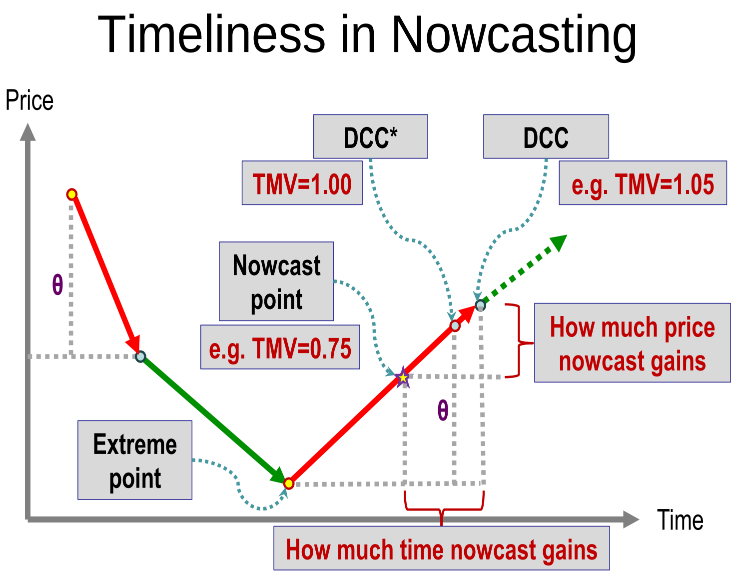

Figure 5. 14 Timeliness in Nowcasting: Timeliness_{time} measures how much time the nowcast gains before the DCC took place; $\text{Timeliness}_{price}$ measures how much price the nowcast gains before the DCC price.

5.6.2.3 Recall (for reference)

What percentage of the trends are nowcast correctly (Definition 5. 6)

Recall = CN / TnT

The recall is used as a reference. It is not a key performance measure because sometimes a good nowcast is not possible. This will be the case, for example, when the price reaches the next extreme point shortly after DCC (i.e., when overshoot is short). In this case, the direction is changed without much warning.

5.6.3. Recording nowcasting performance statistics

In this section, we need to find out all the results that are good. We monitor the market tick-by-tick. Once our program finds a tick that meets the two criteria of the Nowcast Constant Algorithm (Section 5.6.1), nowcast that a new trend has started. We record the time and price of that transaction.

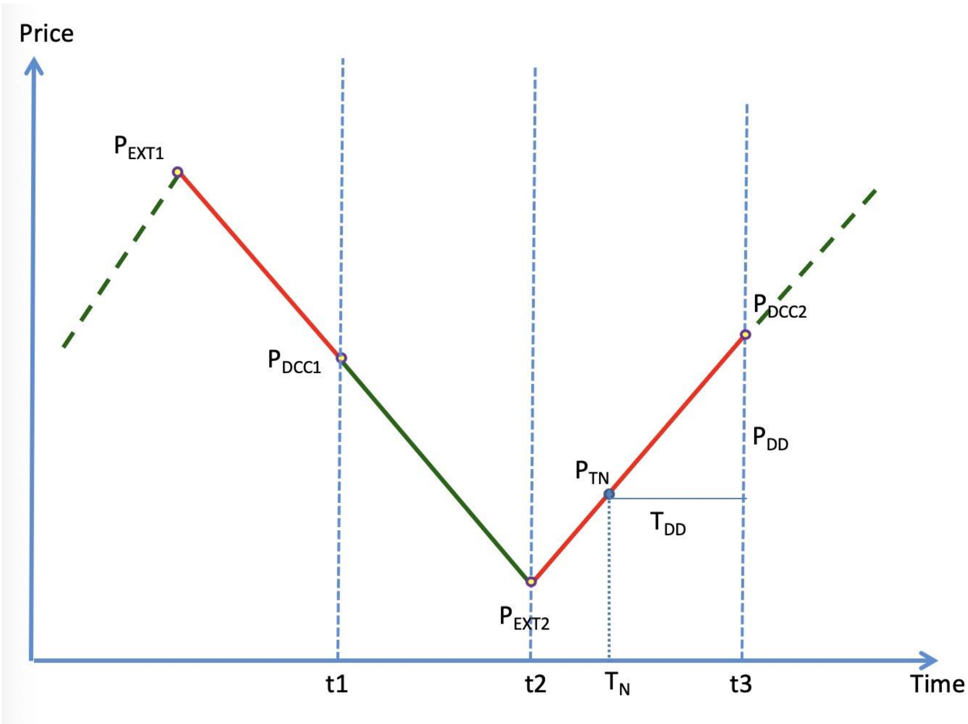

Figure 5. 15 Good nowcast point

Once we confirm a nowcast is a good nowcast, for example, T_N in Figure 5. 15. We can calculate the time difference between T_N to t3 and the price difference between the current price and P_{DCC2}.

How close are they to the DCC of the next trend, in terms of time and price?

Once we confirm a nowcast is a good nowcast, for example, T_N in Figure 5. 15. We can calculate the time difference between T_N to t3 and the price difference between the current price and P_{DCC}. We call the time different T_{DD} and price different P_{DD}. T_{DD} represent that once we know the time at the DCC point, we can calculate the time different from DCC point to nowcast point. The bigger the T_{DD} the sooner we know that we are in a new trend. If we record T_N as a good nowcast point then we can calculate the price difference between the current price and P_{DCC2}. Once we know the

P_{DD}, we can know how far or how significant the price changes from the nowcast point to the DCC point. The bigger the P_{DD} the better we can make more profit in the real treading world.

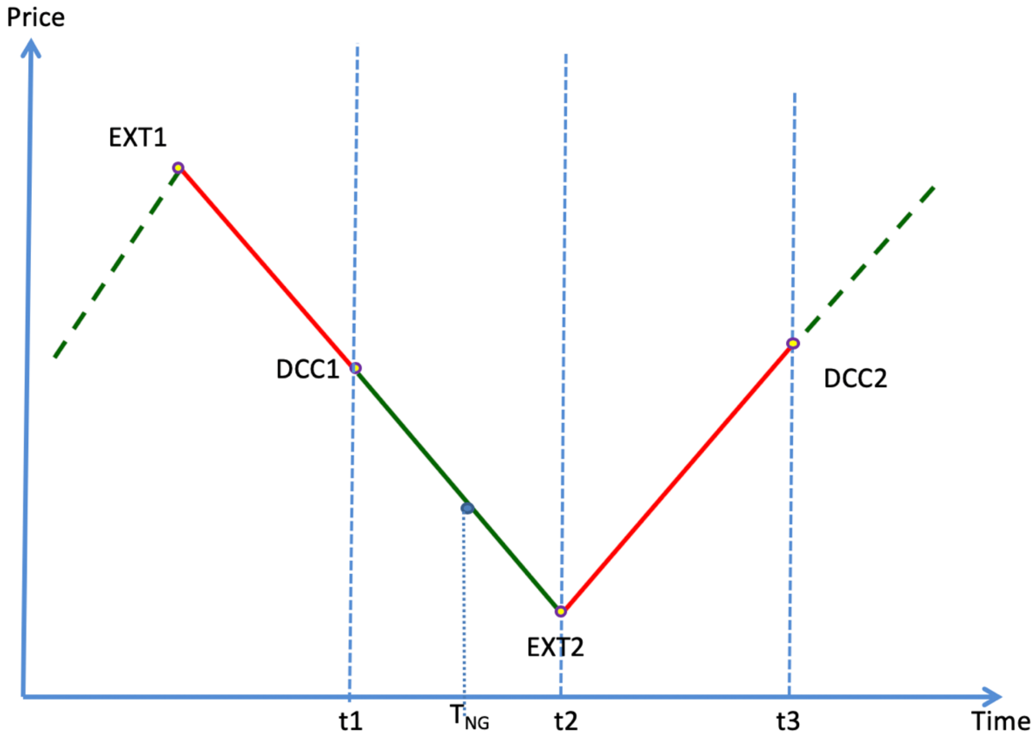

Figure 5. 16 Nogood nowcast point

For those nogood Nowcasts:

Nogood nowcast means the nowcast point program confirm before the EXT2 point. For example, if nowcast is made at T_{NG} in Figure 5. 16, then the nowcast is incorrect, hence a nogood nowcast.

5.6.4. Choice of parameters P_{TMV} and P_{BM}

How do we choose P_{TMV} and P_{BM}? We can base on the choice of these values on past distributions, as described in Sections 5.4 and 5.5. We can choose a P_{TMV} value that is, say, greater than 95% of historical TMV values; similarly, we could choose a P_{BM} value that is greater than 90% of historical BM values. Our choices of P_{TMV} and P_{BM} depend on the risk level that we want to take.

If we want to achieve higher precisions, we may use higher P_{TMV} and P_{BM} values. On the other hand, If we want more timely nowcasts, we should use lower P_{TMV} and P_{BM} values.

5.7. Experimentation

We have set up a set of experiments to assess the performance of the nowcasting algorithm proposed. In this section, we explain these experiments and our assessment.

5.7.1. Experimental setup

We use the distribution of MBM in the Training period (Section 5.5), and the distribution of TMV_{EXT} in the Training period under thresholds 0.0016 and 0.32 (Section 5.4) for nowcasting. We run the **Nowcast Constant Algorithm** (Section 5.6.1) on the Training Period for backtesting and the Nowcasting Period for assessment of the algorithm. We use the Nowcasting period (Section 5.3) in our experiments. We conduct the experiments on both DC thresholds: 0.0016 and 0.0032.

Let us recapitulate the NCA nowcasting rule: It takes two parameters, P_{TMV}

and P_{BM}. We track each new transaction as it emerges. If its aTMV is greater than P_{TMV}, then we check its BM value. If the BM value is greater than P_{BM}, then we nowcast that we are in a new trend.

The higher P_{TMV} and P_{BM} values one chooses, the more cautious one is in nowcasting new trends; that tends to favour precision. The lower P_{TMV} and P_{BM} values one chooses, one tends to favour recall.

We use the same P_{TMV} and P_{BM} to summarize the Training period and nowcasting. As we said before, we run NCA in the Training Period with the same P_{TMV} and P_{BM} that we used in Nowcasting. The reason for this backtesting is that we want to prove firstly, the P_{TMV} and P_{BM} we choose are working in the Training data as it is from the distribution of TMV and MBM in the Training period. Then, we use the same P_{TMV} and P_{BM} in the Nowcasting period to see whether the NCA still working under the nowcasting period by using the historical data. Finally, it's important to analyze the performance of the result under the Nowcasting period for both thresholds.

Under the Training period, we learned the P_{TMV} and P_{BM} values under DC thresholds 0.0016 and 0.0032. We use these learned P_{TMV} and P_{BM} values to nowcast in the Nowcasting period.

All the P_{TMV} and P_{BM} here can be chosen by the observer arbitrarily under the rule in Section 5.6.1. In this chapter, we will choose:

Under NCA in the Training period when the Threshold is equal to 0.0016 we choose P_{TMV}=1.68 and P_{BM}=0.68. There are 46.22% and 1% probability that the P_{TMV} and P_{BM} values would exceed 1.68 and 0.68 (Table 5. 2 and Table 5. 4)

In this case, when aTMV is bigger than 1.68 in a trend and MBM is bigger than 0.68, we will be recorded as a nowcast.

Under NCA in the Nowcasting period when the Threshold is equal to 0.0016 we choose P_{TMV}=1.68 and P_{BM}=0.68 as we introduced before.

We repeat the work as we do in the Training period, any of aTMV is bigger than 1.68 and MBM is bigger than 0.68 but less than 1, we recorded it as a nowcast.

Under NCA in Training period when the Threshold is equal to 0.0032 we choose P_{TMV}=1.60 and P_{BM} =0.61 There are 54.44% and 0.95% probability that the P_{TMV} and P_{BM} values would exceed 1.60 and 0.61 (Table 5. 3 and Table 5. 5)

Under NCA in the Nowcasting period when the Threshold is equal to 0.0032 we also choose P_{TMV}=1.60 and P_{BM} =0.61 as we define at the beginning of this section. That means once the aTMV value is bigger than 1.6 and the MBM value exceeds 0.61 but less than 1, we will record it as a nowcast.

5.8. Result

This section presents the results from running the Nowcasting Constant Algorithm (NCA, Section 5.6) on the data (Section 5.5) under the two selected thresholds of 0.0016 and 0.0032.

5.8.1. Result for the Training period

The Training period is used for learning the parameters for the Nowcasting Constant Algorithm (NCA). We backtest NCA in this period to assess its performance.

5.8.1.1 Result for Training period under threshold 0.0016

In the Training Period, when the threshold is 0.0016, we will have the result, under NCA.

Nowcast results in the Training period under threshold 0.0016			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	9,551	$\text{Timeliness}_{\text{Price}}$	$\text{Timeliness}_{\text{Time}}$ (seconds)
Nc: Number of nowcasts	8,447	Maximum	Maximum
GN: Number of good nowcast	5,389	1.706	211,408
CN: Number of correct nowcasts	5,420	Minimum	Minimum
IN: Number of incorrect nowcast	3,027	0.00	1.00
Nc = CN + IN	8,447	Median	Median
Recall = CN / TnT	56.7%	0.321	467
$\text{Precision}_{\text{Correct}}$ = CN / Nc	64.2%	Average	Average

$Precision_{Good}$ = GN / CN	99.4%	0.325	1,605
False Positive = IN / Nc	35.8%	Standard deviation	Standard deviation
		0.065	8,621

Table 5. 6 The result of NCA in the Training period under threshold 0.0016, with P_{TMV}=1.68, P_{BM}=0.68

Table 5. 6 shows the nowcasting results for NCA in the training period. In the first two columns, the values of the different performance indicators are recorded, showing that in the training period, when the threshold is 0.0016, 9,551 trends in total are identified, and the NCA nowcasts 8,447 times. Of these 8,447 nowcasts, 5,420 are correct, of which 5,389 are useful. 64.20% of nowcasts are good ($Precision_{Correct}$), while 99.40% of correct nowcasts are good nowcasts ($Precision_{Good}$). Using recall as a reference, of all existing trends, 56.70%, nearly two thirds, are the subject of good nowcasts.

Within those good nowcasts, we summarize the in the rightmost two columns in Table 5. 6 their timeliness. Under the $Timeliness_{Price}$ column, the maximum value of $Timeliness_{Price}$ is 1.706 and the minimum value is 0.00. Normally, $Timeliness_{Price}$ are smaller than one, but some nowcasting point would happen directly cross the DCC point, which made $Timeliness_{Price}$ bigger than 1, meanwhile, when the $Timeliness_{Price}$ is 0.00, these means that we nowcast at the DCC point. In the training period when the threshold is 0.0016, the average $timeliness_{Price}$ is 0.325 and the median is 0.321. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So, 0.325 means we manage to nowcast the new trend one third of the way to DCC.

$\text{Timeliness}_{\text{Time}}$ measures how early we manage to recognize a new trend before DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we nowcast the new trend started. In the training period, when the threshold is 0.0016 the maximum $\text{Timeliness}_{\text{Time}}$ is 211,408 seconds and the minimum is 1. The average value of $\text{Timeliness}_{\text{Time}}$ is 1,605 which means, on average NCA can get around 1,605 seconds ahead of the DCC point. The mean $\text{Timeliness}_{\text{Price}}$ is 0.325. That means for good nowcasts were made when the price reached just over one-third of P_{DCC}. Furthermore, when we look at the standard deviation of both $\text{Timeliness}_{\text{Price}}$ and $\text{Timeliness}_{\text{Time}}$. The standard deviation of $\text{Timeliness}_{\text{Price}}$ is 0.065, which is a very small number which means most of the nowcasting prices are reasonably close to the average $\text{Timeliness}_{\text{Price}}$ value. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 8,621, which is quite a big number. This would mean the $\text{Timeliness}_{\text{Time}}$ that we nowcast ahead of DCC point falls into a wide range, this would be a very big number and also would be a small number, which means n the good nowcast, we will get some of the new trends happened really early before DCC point, but some are not. As we can also see the median value of $\text{Timeliness}_{\text{Time}}$ is 467 while the average value is 1,605.

We can see that the standard deviation of $\text{Timeliness}_{\text{Time}}$ is high. But we only record $\text{Timeliness}_{\text{Time}}$ when the nowcast is correct and useful (i.e. before DC Confirmation). Hence, all that means is sometimes we manage to nowcast correctly really early, sometimes very late (but still before DC Confirmation). Since the lower bound of $\text{Timeliness}_{\text{Time}}$ is 0 (because only correct and useful nowcasts are counted), a big value in the standard deviation of $\text{Timeliness}_{\text{Time}}$ is a positive result.

5.8.1.2 Result for Training period under threshold 0.0032

Nowcast results in the Training period under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	2,051	$\text{Timeliness}_{\text{Price}}$	$\text{Timeliness}_{\text{Time}}$ (seconds)
Nc: Number of nowcasts	2,213	Max	Max
GN: Number of good nowcast	1,285	1.264	243,797
CN: Number of correct nowcasts	1,287	Min	Min
IN: Number of incorrect nowcast	926	0.118	1.00
Nc = CN + IN	2,213	Median	Median
Recall = CN / TnT	62.75%	0.391	2,550
$\text{Precision}_{\text{Correct}}$ = CN / Nc	58.16%	Average	Average
$\text{Precision}_{\text{Good}}$ = GN / CN	99.84%	0.393	9,828
False Positive = IN / Nc	41.84%	Standard deviation	Standard deviation
		0.041	29,126

Table 5. 7 The result of NCA nowcast in the Training period under threshold 0.0032, with P_{TMV}=1.60, P_{BM}=0.61

Table 5. 7 shows the nowcasting results for NCA in the training period when the threshold is 0.0032. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the training period, when the threshold is 0.0032, we have 2,051 trends in total. NCA nowcasts 2,213 times. The reason the total nowcast number is bigger than the total number of trends is we may nowcast more than once in a trend. For example, in one trend, the value of aTMV and MBM would exceed the value of P_{TMV} and P_{BM} more than once. In our NCA, every time the value of aTMV and MBM would exceed the value of P_{TMV} and P_{BM} we record as a nowcast. In these 2,213 nowcast, 1,287 are correct; among which 1,285 are useful. 58.16% of nowcast from the total nowcast are good

($\text{Precision}_{\text{Correct}}$), and 99.84% of correct nowcast are good nowcast ($\text{Precision}_{\text{Good}}$). We use recall as a reference: out of all trends, 62.75%, nearly two third are nowcasted good.

Within those good nowcasts, we summarize their timeliness in the rightmost two columns in Table 5. 7. Under the $\text{Timeliness}_{\text{Price}}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 1.264 and the minimum value is 0.118. Normally, $\text{Timeliness}_{\text{Price}}$ is below one, but some nowcasting points would happen directly across the DCC point, which made $\text{Timeliness}_{\text{Price}}$ bigger than 1. In the training period when the threshold is 0.0032, the average $\text{timeliness}_{\text{Price}}$ is 0.393 and the median is 0.391. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So 0.393 means we manage to nowcast the new trend nearly 40% of the way to DCC.

$\text{Timeliness}_{\text{Time}}$ measures how early we manage to recognize a new trend before DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we know the new trend started. In the training period, when the threshold is 0.0032 the maximum $\text{Timeliness}_{\text{Time}}$ is 243,797 seconds and the minimum is 1. The average value of $\text{Timeliness}_{\text{Time}}$ is 9,828 which means, on average NCA can get around 9,828 seconds ahead of the DCC point. Furthermore, when we look at the standard deviation of both $\text{Timeliness}_{\text{Price}}$ and $\text{Timeliness}_{\text{Time}}$. At first, when we get the result of the standard deviation of $\text{Timeliness}_{\text{Price}}$ is equal to 0.041, this is a very small number which means most of the nowcasting prices are reasonably close to the average value of $\text{Timeliness}_{\text{Price}}$. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 29,126, which is quite a big number. This means in the good nowcast, we will get some of the new trends early before the DCC point, but some are not. As we can also see the median value of $\text{Timeliness}_{\text{Time}}$ is 2,550 while the average value is 9,828.

5.8.2. Result for the Nowcasting period

The parameters used in the Nowcasting period by the NCA were selected based on the Training Period. The Nowcasting period is used to assess the performance of the Nowcast Constant Algorithm (NCA) in an out-of-sample period.

5.8.2.1 Result for Nowcasting period under threshold 0.0016

Nowcast results in the Nowcasting period under threshold 0.0016			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	4,645	$Timeliness_{Price}$	$Timeliness_{Time}$ (seconds)
Nc: Number of nowcasts	3,452	Max	Max
GN: Number of good nowcast	2,264	2.154	221,194
CN: Number of correct nowcasts	2,300	Min	Min
IN: Number of incorrect nowcast	1,152	0.0113	1.00
Nc = CN + IN	3,452	Median	Median
Recall = CN / TnT	49.52%	0.322	367.5
$Precision_{Correct}$ = CN / Nc	66.63%	Average	Average
$Precision_{Good}$ = GN / CN	98.43%	0.336	1,967
False Positive = IN / Nc	33.37%	Standard deviation	Standard deviation
		0.107	11,745

Table 5. 8 The result of NCA nowcast in the Nowcasting period under threshold 0.0016, with P_{TMV}=1.68, P_{BM}=0.68

Table 5. 8 shows the nowcasting results for NCA in the Nowcasting period when the threshold is 0.0016. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the nowcasting period, when the threshold is 0.0016, we have 4,645 trends in total. NCA nowcasts 3,452 times. In these 3,452 nowcasts, 2,300 are correct, among which 2,264 are useful (hence good). 66.63% of nowcast from the total nowcast are good ($Precision_{Correct}$), and 98.43% of correct nowcast are good nowcast ($Precision_{Good}$). We use recall as a reference: out of all trends, 49.52%, i.e., roughly half, are nowcasted good.

Within those good nowcasts, we summarize the in the rightmost two columns in Table 5. 8 their timeliness. Under the $Timeliness_{Price}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 2.154 and the minimum value is 0.011. In the nowcasting period when the threshold is 0.0016, the average $\text{timeliness}_{\text{Price}}$ is 0.336 and the median is 0.322. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So, 0.336 means we manage to nowcast the new trend one third of the way to DCC. The standard deviation of $\text{Timeliness}_{\text{Price}}$ is 0.107, which is reasonably small.

$\text{Timeliness}_{\text{Time}}$ measures how early the NCA manage to recognize a new trend before the DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we manage to recognize the start of the new trend. In the Nowcasting period, under threshold 0.0016, the median $\text{Timeliness}_{\text{Time}}$ is 367.5 while the average value is 1,967. The big difference is explained by extreme values: The maximum $\text{Timeliness}_{\text{Time}}$ is 221,194 seconds and the minimum is 1. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 11,745 seconds, which is quite a big number. This means amongst the good nowcast, some are early before the DCC point, but some are not.

Scenario Nowcast_1 (SNC1):

The maximum $\text{Timeliness}_{\text{Price}}$ is 2.154. Figure 5. 17 provides details of the nowcast that produced this $\text{Timeliness}_{\text{Price}}$.

Figure 5. 17 A nowcast on 22/01/2015 which produced the maximum $\text{Timeliness}_{\text{Price}}$ of 2.154

The DC confirmation point DCC1 (at the price of 1.15660) confirms that a downtrend has started from the extreme point EXT1 (at the price of 1.15904). Without the benefit of hindsight, all the transactions after DCC1 are considered to be part of the downtrend, until DCC2 (the highest point on the right, at 08:49:37, 1.16011) is encountered. At DCC2, one learns (in hindsight) that EXT2 (at 08:48:25, 1.15486) was, in fact, an extreme point.

Let us examine how NCA perform in this scenario: After EXT2 and before DCC2, NCA nowcasts at the point NCP (at 08:48:37, 1.15613) that it is in a new trend. This means NCA recognizes that EXT2 was, in fact, an extreme point. Nowcast is made at NCP because, at the price of 1.15486, EXT2 became the Max (i.e., the lowest point in the current downtrend from EXT1). The two conditions in NCP are met: EXT2 has an aTMV of (|1.5486–1.15904|/1.15904/0.0016=)2.254, which is bigger than the threshold P_{TMV}=1.68. At NCP, the BM value (Equation 5.4, Section 5.2) is (|1.15613-1.15486|/1.15486/0.0016=) 0.6873, which is bigger than the threshold P_{BM}=0.68. $Timeliness_{Time}$ of this nowcast is (08:49:37–08:48:37=) 60 seconds.

This scenario shows in a downtrend an extremely valuable nowcast because it recognizes a directional change way before the DC confirmation point, both price-wise and time-wise. Careful examination of the tick data reveals that this new trend was confirmed when the transaction price jumped from 1.1557 (which is not yet 0.16% above EXT2) to 1.15613 in one transaction. NCA nowcasted the new trend (2.154*0.0016=) 0.00346, or 0.3446%, in advance[9]. NCA nowcasts just 12 seconds after EXT2, 60 seconds ahead of DCC2. So, this nowcast could have gained a trader a lot

[9] Another way to calculate this percentage is to use the prices in Figure 5.17: The new trend is confirmed at DCC2, at the price of 1.16011.NCA nowcasted the new trend at the price 1.15613, which is (1.16011-1.15613=) 0.00298. From the extreme point EXT2 (at the price 1.15486), this represents an early warning of((0.00298/1.15486)x100%=) 0.3446%.

of price and time ahead of its competitors.

[End of Scenario SNC1]

Scenario Nowcast_2 (SNC2):

The maximum $\text{Timeliness}_{\text{Price}}$ is 2.154. Figure 5. 18 provides details of the nowcast that produced this $\text{Timeliness}_{\text{Price}}$. For simplicity, we show the seconds instead of actual dates and times in Figure 5. 18.

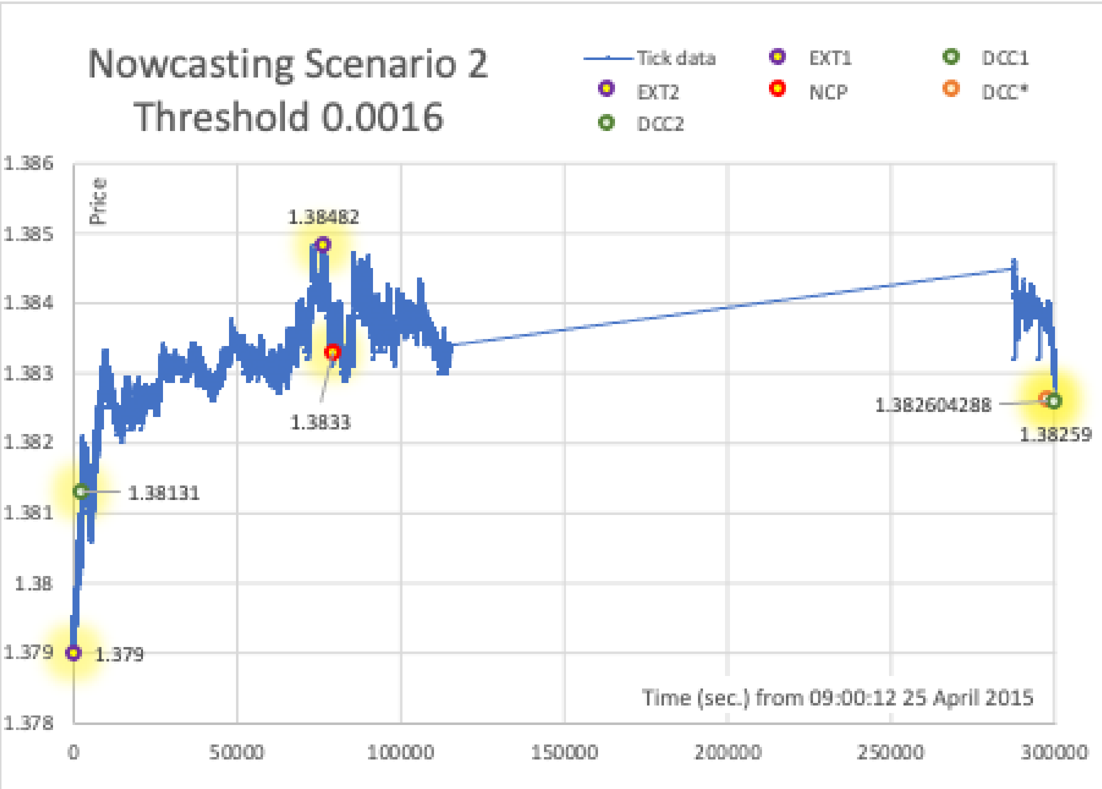

Figure 5. 18 A nowcast between 09:00:12 25 April 2015 to 20:27:19 27 April 2015 which produced the maximum $\text{Timeliness}_{\text{Time}}$ of 221,194 seconds; The x-axis counts the seconds from the first data point.

The DC confirmation point DCC1 (at the 2,726th second, price 1.138131) confirms that an uptrend has started from the extreme point EXT1 (at the 2nd second, price 1.379). Without the benefit of hindsight, all the transactions after DCC1 are considered to be part of the uptrend, until DCC2 (the lowest point on the right, at 300,426 1.38259) is encountered.

At DCC2, one learns (in hindsight) that EXT2 (at 76,512, 1.38482) was an extreme point.

Let us examine how NCA perform in this scenario: After EXT2 and before DCC2, NCA nowcasts at the point NCP (at 79,232, 1.38259) that a new trend has already started. This means NCA recognizes that EXT2 was, in fact, an extreme point. $\text{Timeliness}_{\text{Time}}$ of this nowcast is (300,426–79,232=) 221,194 seconds, or 2 days, 13 hours, 26 minutes and 34 seconds. Even if we discount the two days in which no trades took place, the nowcast is still nearly 13 and a half hours ahead of DCC. This is a huge success given that NCP is only (79,232-76,512=) 2,720 seconds after the new trend has started from EXT2.

[End of Scenario SNC2]

This is work in nowcasting new trends, we argue that these results are good. We have managed to nowcast correctly in about half of the trends (Recall=49.52%). Two third of the nowcasts are correct ($Precision_{Correct}$ = 66.63%). The vast majority of our nowcast is good ($Precision_{Good}$ = 98.43%). For those good nowcasts, we manage to nowcast on average one-third of the price (Average $Timeliness_{Price}$ = 0.336), and 1,967 seconds (Average $Timeliness_{Time}$) before NC confirmation. In the best case, nowcast recognized the new trend 2.154 times the threshold before the DCC transaction. We conclude that the parameters P_{TMV} and P_{BM} learned from historical distributions in the Training period are good for nowcasting in the Nowcasting period, despite the profile of the two periods being very different (Section 5.3).

5.8.2.2 Result for Nowcasting period under threshold 0.0032

Nowcast results in the Nowcasting period under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	1,184	$Timeliness_{Price}$	$Timeliness_{Time}$ (seconds)
Nc: Number of nowcasts	913	Max	Max
GN: Number of good nowcast	544	1.961	229,764
CN: Number of correct nowcasts	546	Min	Min
IN: Number of incorrect nowcast	367	0.00	1.00
Nc = CN + IN	913	Median	Median
Recall = CN / TnT	46.11%	0.392	1,623
$Precision_{Correct}$ = CN / Nc	59.80%	Average	Average
$Precision_{Good}$ = GN / CN	99.63%	0.399	8,789
False Positive = IN / Nc	40.20%	Standard deviation	Standard deviation
		0.098	24,796

Table 5. 9 The result of NCA nowcast in the Nowcasting period under threshold 0.0032, with P_{TMV}=1.61, P_{BM}=0.61

Nowcast results in the Nowcasting period under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	1,184	TimelinessPrice	TimelinessTime (seconds)
Nc: Number of nowcasts	913	Max	Max
GN: Number of good nowcast	544	1.961	229,764
CN: Number of correct nowcasts	546	Min	Min
IN: Number of incorrect nowcast	367	0.00	1.00
Nc = CN + IN	913	Median	Median
Recall = CN / TnT	46.11%	0.392	1,623
PrecisionCorrect = CN / Nc	59.80%	Average	Average
PrecisionGood = GN / CN	99.63%	0.399	8,789
False Positive = IN / Nc	40.20%	Standard deviation	Standard deviation
		0.098	24,796

Table 5. 9 shows the nowcasting results for NCA in the nowcasting period

when the threshold is 0.0032. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the nowcasting period, when the threshold is 0.0032, we have 1,184 trends in total. NCA nowcasts 913 times. In these 913 nowcasts, 546 are correct; among which 544 are useful (hence good). 59.80% of nowcast from the total nowcast are good ($\text{Precision}_{\text{Correct}}$), and 99.63% of correct nowcast are good nowcast ($\text{Precision}_{\text{Good}}$), which is excellent. We use recall as a reference: out of all trends, 46.11%, i.e., just less than half, are nowcasted good.

Within those good nowcasts, we summarize the in the rightmost two columns in

Nowcast results in the Nowcasting period under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	1,184	TimelinessPric e	TimelinessTime (seconds)
Nc: Number of nowcasts	913	Max	Max
GN: Number of good nowcast	544	1.961	229,764
CN: Number of correct nowcasts	546	Min	Min
IN: Number of incorrect nowcast	367	0.00	1.00
Nc = CN + IN	913	Median	Median
Recall = CN / TnT	46.11%	0.392	1,623
PrecisionCorrect = CN / Nc	59.80%	Average	Average
PrecisionGood = GN / CN	99.63%	0.399	8,789
False Positive = IN / Nc	40.20%	Standard deviation	Standard deviation
		0.098	24,796

Table 5. 9 their timeliness. Under the $\text{Timeliness}_{\text{Price}}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 1.961 and the minimum value is 0.0. Here the

0.00 means on one or more nowcast point, it has happened at the same point as DCC point. In the nowcasting period when the threshold is 0.0032, the average $timeliness_{Price}$ is 0.399 and the median is 0.392. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. Here 0.399 means we manage to nowcast the new trend nearly 40% of the threshold away from DCC (Section 5.6.2).

As we introduce in section 5.6.2.2, $Timeliness_{Time}$ measures how early we manage to recognize a new trend before DCC. The bigger the $Timeliness_{Time}$ is, the earlier we know the new trend started. In the nowcasting period, when the threshold is 0.0032 the maximum $Timeliness_{Time}$ is 229,764 seconds and the minimum is 1. The average value of $Timeliness_{Time}$ is 8,789 which means, on average NCA can get around 8,789 seconds ahead of the DCC point. Furthermore, when we look at the standard deviation of both $Timeliness_{Price}$ and $Timeliness_{Time}$. At first, when we get the result of the standard deviation of $Timeliness_{Price}$ which is 0.098, this means that the volatility of $Timeliness_{Price}$ is low, and most of the nowcasting price is reasonably close to the average value of $Timeliness_{Price}$. The standard deviation of $Timeliness_{Time}$ is 24,796 seconds, which is very a big number. This means that time we nowcast ahead of DCC point are in a very wide range. In the good nowcasts, we get some of the new trends very early before the DCC point, but some are not. The median value of $Timeliness_{Time}$ is 1,623 while the average value is 8,789, which is affected by extreme values, as explained in the wide range.

Scenario Nowcast_3 (SNC3):

The maximum Timeliness$_{\text{Price}}$ is 1.961. Figure 5. 19 provides details of the nowcast that produced this Timeliness$_{\text{Price}}$.

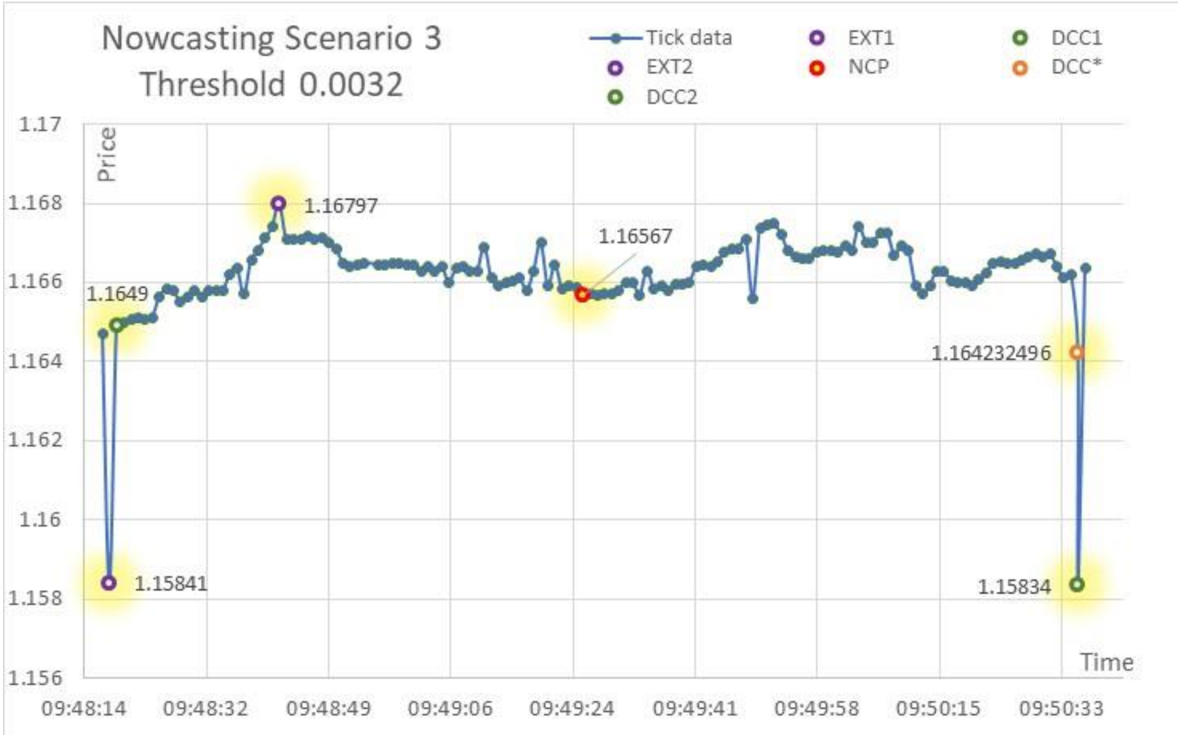

Figure 5. 19 A nowcast on 21/01/2015 which produced the maximum Timeliness$_{\text{Price}}$ of 1.961

The DC confirmation point DCC1 (at the price of 1.1649) confirms that an uptrend has started from the extreme point EXT1 (at the price of 1.15841). Without the benefit of hindsight, all the transactions after DCC1 are considered to be part of the uptrend, until DCC2 (the lowest point on the right, at 09:50:35, 1.15834) is encountered. At DCC2, one learns (in hindsight) that EXT2 (at 09:48:42, 1.167979) was, in fact, an extreme point.

Let us examine how NCA perform in this scenario: After EXT2 and before

DCC2, NCA nowcasts at the point NCP (at 09:49:25, 1.16567) that it is in a new trend. This means NCA recognizes that EXT2 was, in fact, an extreme point. Nowcast is made at NCP because, at the price of 1.16567, EXT2 became the Max (i.e., the highest point in the current uptrend from EXT1). The two conditions in NCP are met: EXT2 has an aTMV of (|1.16797–1.16567|/1.15834/0.0032=) 1.961, which is bigger than the parameter P_{TMV}=1.61. At NCP, the BM value (Equation 5.4, Section 5.2) is (|1.16567-1.16797|/1.16797/0.0032=) 0.61538, which is bigger than the threshold P_{BM}=0.61. Timeliness Time of this nowcast is (09:50:35–08:49:25=) 70 seconds.

This scenario shows in an extremely valuable nowcast an uptrend because it recognizes a directional change way before the DC confirmation point, both price-wise and time-wise. Careful examination of the tick data reveals that this new trend was confirmed when the transaction price jumped from 1.16586 (which is not yet 0.32% below EXT2) to 1.16567 in one transaction. NCA nowcasted the new trend (1.961*0.0032=) 0.006266, or 0.6266%, in advance. NCA nowcasts 43 seconds after EXT2, 70 seconds ahead of DCC2. So, this nowcast could have gained a trader a lot of price and time ahead of its competitors.

[End of Scenario SNC3]

In summary, under threshold 0.0032, we find that under a different threshold, NCA still managed to nowcast correctly in nearly half of the trends (Recall=46.11%). About 60% of the nowcasts are correct ($Precision_{Correct}$ = 59.80%). The vast majority of our nowcast is good ($Precision_{Good}$ = 99.63%); out of 546 correct nowcasts, only two nowcasts are correct but not good. For those good nowcasts, we manage to nowcast on average 40% of the price (Average $Timeliness_{Price}$ = 0.399), and 8,789 seconds (Average $Timeliness_{Time}$) before NC confirmation. We conclude that the parameters P_{TMV} and P_{BM} learned from historical distributions in the Training period are good for nowcasting in the Nowcasting period when the threshold is 0.0032 as well, despite the profile of the two periods being very different.

5.9. Discussion

In this section, we shall assess the performance of NCA based on the results presented so far. Key performance measures are shown in Table 5. 10 below. We shall comment on NCA's nowcasts' timeliness in terms of price and time. We shall also comment on the robustness of NCA.

Summary of key performance measures				
	Training Period		Nowcasting Period	
Threshold	0.0016	0.0032	0.0016	0.0032
$\text{Precision}_{\text{Correct}}$	64.16%	58.16%	66.63%	59.80%
$\text{Precission}_{\text{Good}}$	99.42%	99.84%	98.43%	99.63%
Recall	56.75%	62.75%	49.52%	46.11%
Average $\text{Timeliness}_{\text{Price}}$	0.325	0.393	0.336	0.399
Standard deviation of $\text{Timeliness}_{\text{Price}}$	0.065	0.041	0.107	0.098
Maximum $\text{Timeliness}_{\text{Price}}$	1.706	1.264	2.154	1.961
Average $\text{Timeliness}_{\text{Time}}$	1,605 sec.	9,828 sec.	1,967 sec.	8,789 sec.
Standard deviation of $\text{Timeliness}_{\text{Time}}$	8,621 sec.	29,126 sec.	11,745 sec.	24,796 sec.
Maximum $\text{Timeliness}_{\text{Time}}$	221,408 sec.	243,797 sec.	221,194 sec.	229,764 sec.

Table 5. 10 Summary of key performance measures in all the experiments

5.9.1. NCA produces timely nowcasts in terms of Price

In scenarios SNC1 and SNC3, we have shown two exceptionally valuable nowcasts as measured by $\text{Timeliness}_{\text{Price}}$. In scenario SNC1, $\text{Timeliness}_{\text{Price}}$ was 2.514. This is translated into nowcasting 0.3446% ahead of DC confirmation. This is especially significant when we consider the fact that the DC threshold was only 0.16%. That means DC is confirmed when the absolute price change is 0.16% or more. 0.3446% was only achieved because the market jumped in one transaction to the DC confirmation point[10]. What makes this nowcast even more valuable was that it nowcasts

[10] Close inspection reveals that price jumped from 1.1553 (which does not yet confirm directional change from the last extreme point at 1.115486) to 1.16011, a jump of (1.16011-1.1553)/1.115486=) 0.417%.

only 12 seconds after the new trend has started (which means that the market was particularly volatile at the time), 60 seconds ahead of the DCC point.

Likewise, in Scenario SNC3, NCA nowcasts 0.6266% ahead of NCC point ($Timeliness_{Price}$ 1.961) when the DC threshold was only 0.32%. Such a nowcast could be used to gain advantages in trading.

While not every good nowcast is as timely as SNC1 and SNC3 in terms of price, Table 5. 10 shows that the average $Timeliness_{Price}$ in both the Training and Nowcast periods under both thresholds range between 0.325 and 0.399. That means, on average, good nowcasts were made two-thirds of the price from the extreme point to the DCC point (see explanation in SNC1). The relatively small standard deviations of $Timeliness_{Price}$ across Table 5. 10 suggest that most of the nowcasts are not far from one-third from the extreme price to the DCC price.

5.9.2. NCA produces timely nowcasts in terms of Time

Scenario SNC2 shows that nowcasts could be days ahead of DC confirmation. Though not as significant as SNC2, $Timeliness_{Time}$ of NCA SNC1 and SNC3 were still impressive:

- In SNC1, NCA nowcasts only 12 seconds after EXT2, 60 seconds ahead of DCC2.

- In SNC3, NCA nowcasts 43 seconds after EXT2, 70 seconds ahead of DCC2.

Table 5. 10 provides evidence that NCA provides timely nowcasts in terms of time: the average Timeliness in both Training and Nowcasting Periods, under both thresholds, ranged from 1,605 to 9,828 seconds. The standard deviations of $\text{Timeliness}_{\text{Time}}$ are big compared to the averages. This suggests that thousands of seconds can be gained through good nowcasting by NCA.

5.9.3. NCA is robust across market profiles

In section 5.3, we showed that the Training Period and Nowcasting Period have very different DC profiles. The former is a lot less volatile than the latter in terms of DC frequency (as reflected by the indicator Median T, see Table 5. 1).

SNC1 and SNC3 took place in volatile markets where trends ended within one or two minutes. More importantly, big jumps were encountered in both scenarios. So these two trends were not only volatile in terms of T, they are volatile in terms of TMV[11]. SNC2 took place in a quiet market: the trend took days to finish. NCA has produced valuable nowcasts, in terms of

[11] Readers are reminded that a market's volatility can be measured by TMV and Time in DC, see (Tsang et al 2017)

Timeliness $_{Price}$, Timeliness $_{Time}$ or both.

The consistency of these results supports the claim that NCA is robust over different market conditions.

NCA is robust across thresholds

Experiments were run under thresholds, 0.0016 and 0.0032. These thresholds were chosen according to the criteria defined in Chapter 3. Results summarized in Table 5. 10 show NCA's performance under the two different thresholds is consistent: $\text{Precision}_{\text{Correct}}$ under the two thresholds are both around 60%. Results under 0.0016 are slightly better than results under 0.0032 because a smaller threshold is more sensitive to price changes in the market. $\text{Precision}_{\text{Good}}$ values under the two thresholds are both around 99%.

Note that discrepancy in Average $\text{Timeliness}_{\text{Time}}$ between the two thresholds is not a sign of inconsistency. By nature, as the DC threshold increases, trends take exponentially longer to complete[12]. As aTMV is normalized by the threshold, the Median aTMVs in Table 5.1 are consistent between the two thresholds: between 1.43 and 1.68. They also support the robustness of NCA across different DC thresholds.

[12] Readers may also consult Glattfelder [5] in their study of the scaling laws. Readings of Median T can be found in Table 5. 1

5.9.4. Rooms for improvement

Without fine-tuning the parameters (P_{TMV} and P_{BM}), NCA achieves around 60% in $\text{Precision}_{\text{Correct}}$ and $\text{Precision}_{\text{Good}}$. One could expect better results by fine-tuning these parameters, for example by adopting them to the market's volatility or dynamically changing the value of P_{BM} depending on P_{TMV}. These will be left for future research.

5.10.Backtesting with an extremely small threshold

In Chapter 3, we presented a guideline for limiting the thresholds to be used for DC. We suggested that if we use a threshold that is too small, DC summaries contain a lot of noise. In this section, we shall test this guideline. We shall nowcast with a threshold that is smaller than the lower bound in the guideline introduction in Section 3.4.3. We shall use this threshold to nowcast in two periods of EUR/USD data to see whether nowcasting performance is affected. Will the number of nowcasts be affected? Will the nowcast performance be affected?

5.10.1. Data we used

5.10.1.1 Training period and nowcasting period

In this section, we will use tick data in the EUR/USD exchange market from 21/04/2016 07:31:08 to 06/05/2016 14:31:21 (The same as Section 3.4.1) as the training period_T and the EUR/USD exchange market from

06/05/2016 14:32:16 to 25/05/2016 06:59:21 as the nowcasting period$_T$.

5.10.1.2 Threshold we used

In this section, we will use Threshold=0.00002 for both the training period and nowcasting period while in Section 3.4.3 we see any threshold below 0.00005 is unsuitable for profiling EUR/USD Period One

5.10.2. Result

5.10.2.1 Historical distribution of TMV_{EXT}

In this section, we study the historical distribution of TMV_{EXT}.

We use the EUR/USD tick data from 21/04/2016 07:31:08 to 06/05/2016 14:31:21. We named this period the Training Period$_T$. Also, we shall use EUR/USD tick data from 06/05/2016 14:32:16 to 25/05/2016 06:59:21, we named this period as the Nowcasting period$_T$. We calculate TMV_{EXT} under threshold 0.00002.

First, we calculate the TMV_{EXT} under 0.00002 for Training Period$_T$. Then we collect it from every trend. In the Training period$_T$ when the threshold is equal to 0.00002, we got 88,333 TMV_{EXT}.

Once we collect all TMV_{EXT} we will start to summarize the TMV_{EXT}. Table A. 9 in Section 7 appendix show the distribution of TMV_{EXT} under 0.00002 for Training Period$_T$

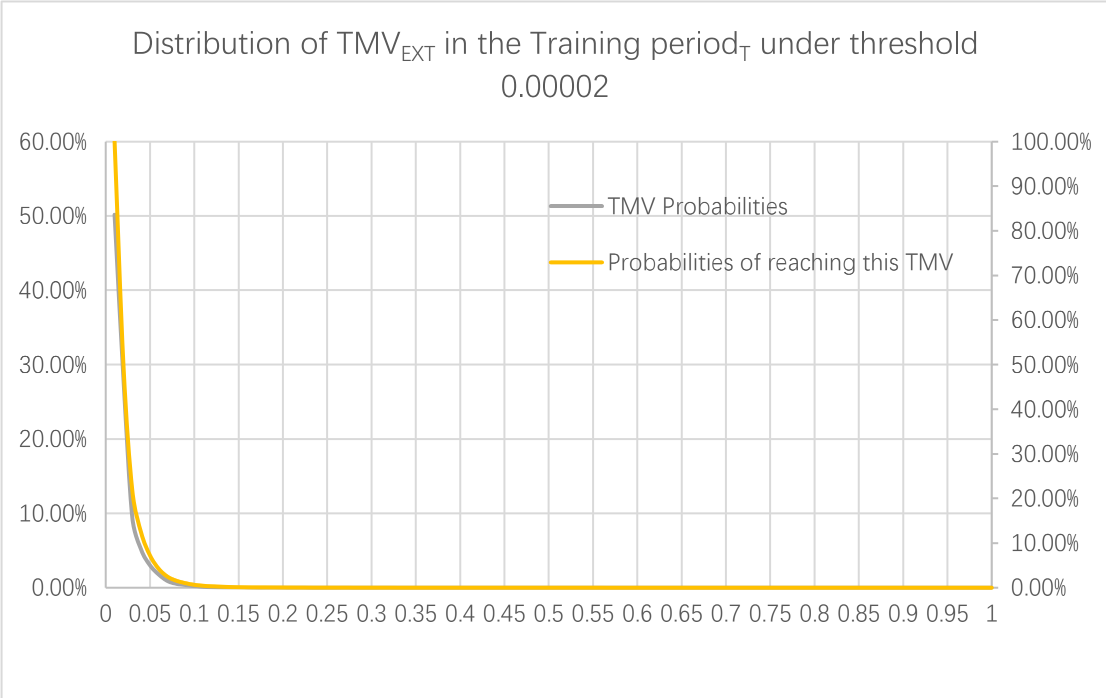

Figure 5. 20 Distribution of TMV$_{EXT}$ in the Training periodT under threshold 0.00002

Figure 5. 20 show the graph of the probability of TMV$_{EXT}$ happening in each bin and the probability of reaching a certain number of TMV$_{EXT}$ in the Training period$_T$ when the threshold is equal to 0.00002.

5.10.2.2 Historical distribution of Max below Max (MBM)

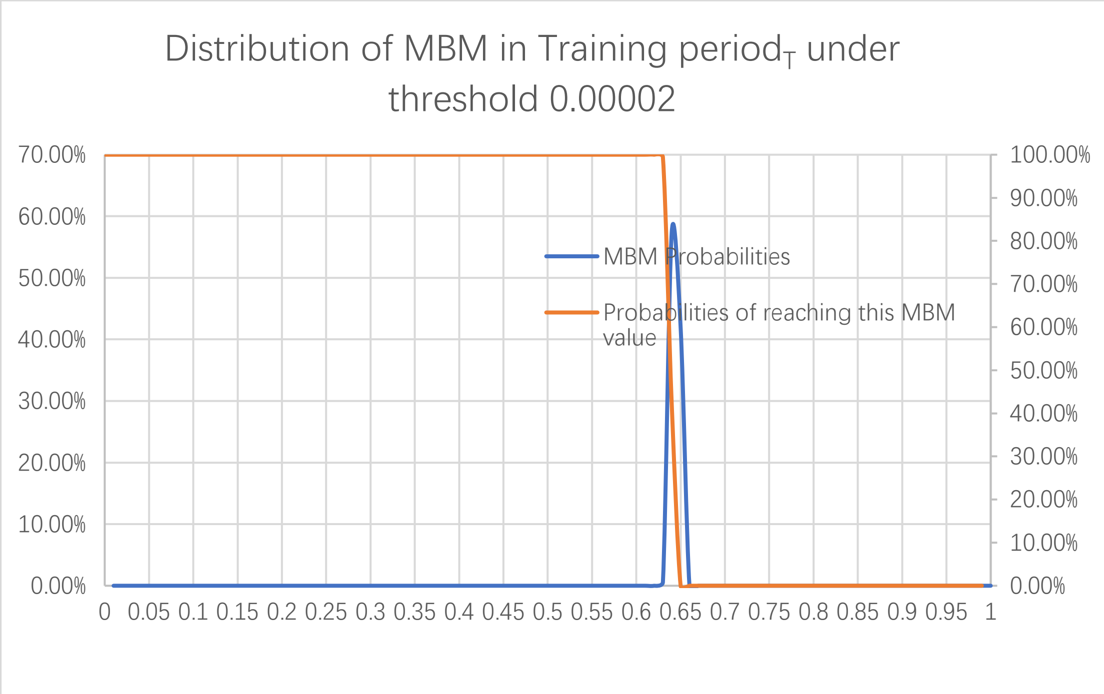

Figure 5. 21 Distribution of MBM in Training period$_T$ under threshold 0.00002

Figure 5. 21 show the graph of the probability of MBM happening in each bin and the probability of reaching a certain number of MBM in the Training period when the threshold is equal to 0.00002.

As the figure about shown, under threshold 0.00002, the MBM value is only from 0.63 to 0.65 which is strange (as we can see in Section 5.5, Section 5.11 and 5.12, all MBM values are subsection wildly from 0 to 1), normally it would happen in most of the bin from 0 to 1.

5.10.2.3 Choice of the Parameter

The MBM value in the Training period$_T$ under threshold 0.00002 that we pick			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
1	-	-	100.00%

0.63	183	0.58%	100.00%
……			
0.64	18133	57.43%	99.42%
0.6457	**-**		**99.00%**
0.65	13256	41.99%	41.99%
……			
1	0	0.0%	

Table 5. 11 The MBM we pick to use under the distribution of MBM in the Training $period_T$ under threshold 0.00002

The of TMV_{EXT} in the Training $period_T$ under threshold 0.00002 that we pick			
Bin	TMV_{EXT} Frequency	TMV_{EXT} Probabilities	Probabilities of reaching this TMV_{EXT} value
3.08433502	44311	50.16%	100.00%
……			
17.67468016	464	0.53%	1.55%
19.57			1.00%
19.75901518	330	0.37%	1.03%
……			
209.433502	1	0.00%	0.00%

Table 5. 12 The TMV_{EXT} we pick to use under the distribution of TMV_{EXT} in the Training $period_T$ under threshold 0.00002

Under the Training $period_T$, we learned the P_{TMV} and P_{BM} values under DC thresholds 0.00002. We use these learned P_{TMV} and P_{BM} values to nowcast in the Training $period_T$ to test the performance of nowcasting.

All the P_{TMV} and P_{BM} here can be chosen by the observer arbitrarily under the rule in Section 5.6.1.

In this section, we will choose:

Under NCA in Training $period_T$ when Threshold is equal to 0.00002 we choose P_{TMV}=19.57 and P_{BM}=0.6457. There are 1.00% and 1.00% probability that the P_{TMV} and P_{BM} values would exceed 19.57 and 0.6457

(Table 5. 11 and Table 5. 12)

5.10.2.4 Result of Nowcasting in Training period$_T$

In the Training Period$_T$, when the threshold is 0.00002, we will have the NCA results shown in Table 5. 13.

Nowcast results in the Training period$_T$ under threshold 0.00002			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	88,332	Timeliness$_{Price}$	Timeliness$_{Time}$ (seconds)
Nc: Number of nowcasts	0	Maximum	Maximum
GN: Number of good nowcasts	0	NIL	NIL
CN: Number of correct nowcasts	0	Minimum	Minimum
IN: Number of incorrect nowcast	0	NIL	NIL
Nc = CN + IN	0	Median	Median
Recall = CN / TnT	0	NIL	NIL
Precision$_{Correct}$ = CN / Nc	0	Average	Average
Precision$_{Good}$ = GN / CN	0	NIL	NIL
False Positive = IN / Nc	0	Standard deviation	Standard deviation
		NIL	NIL

Table 5. 13 The result of NCA in the Training period under threshold 0.00002, with P_{TMV}=19.57, P_{BM}=0.6457; as no nowcast was made, no statistical results are available

5.10.3. Discussion

From Table 5. 13 we can see that no nowcasts were made under the Training period$_T$; that means we can nowcast nothing under this extremely small threshold. This happens because, when the threshold is very small, many price changes against the current trend are directional changes. When overshoot does not exist, there is no MBM. When overshoot just takes one or two ticks, no MBM will be observed, thus NCA will make no nowcast. Thus, results in this experiment support the use of the DC threshold guidelines presented in Chapter 3.

5.11.Experimental work for USD/JPY

So far, NCA has been tested on EUR/USD. To check whether the results can be reproduced in other markets, we test NCA on USD/JPY in this section.

5.11.1 Experimental work set up

5.11.1.1 Data Used and their profiles

To prepare for our nowcasting algorithm, as in section 5.3, we will use Tick-to-tick USD/JPY rates are used throughout this section. Table 5. 14 summarizes the data used in this chapter. We separate the data into the Training and Nowcasting periods. Statistics in the former is used to determine the parameters used in our nowcasting algorithm, which will be used for nowcasting in the Nowcasting period. In this section, we will use thresholds=0.0016 which will namely 0.0016.

	Data in the Training Period$_{USD/JPY}$	Data in the Nowcasting Period$_{USD/JPY}$
Threshold	0.0016	0.0016
Period	From: 00:00:10 27th September 2009 To: 23:59:59 31st December 2013	From 00:00:09 1st January 2014 To: 23:50:59 31st December 2015
Number of transactions	56,820,049	28,171,600
Number of DCs	17,588	3,892
Median aTMV	1.688934	1.70645
Median T	2,882	4,350

Table 5. 14 Tick-to-tick data used in this section, with their profiles compared

We call the period from 00:00:10 on 27/09/2009 to 23:59:59 on 31/12/2013 as the **Training Period$_{USD/JPY}$**. We call the period from 00:00:09 on 01/01/2014 to 23:50:59 on 31/12/2015 the **Nowcasting Period$_{USD/JPY}$**.

Let us compare and contrast the profiles of the two periods. From **Table 5. 14**. we can see that, under threshold 0.0016, the median aTMV values in the Training Period$_{USD/JPY}$ is 1.688394, whereas the median aTMV in the Nowcasting Period$_{USD/JPY}$ under the same threshold is 1.70645. Under these thresholds, the median aTMV in the Training Period$_{USD/JPY}$ is more or less the same as Nowcasting Period$_{USD/JPY}$.

But when we look at the Median T values, we can see that under threshold 0.0016, a trend takes 2,882 seconds to finish a trend in the Training Period$_{USD/JPY}$, but 4,350 seconds (nearly 1.3 times of the time) to finish in the Nowcasting Period$_{USD/JPY}$.

Both aTMV and T are indicators of volatility in a market-period: the former measures the magnitude of price changes and the latter measures the frequency of changes. The above comparisons suggest the following:

1. There is nearly no difference in the two periods' volatility as measured by median aTMV.

2. Measured by median T, the Nowcasting Period is much more volatile than that Training Period: direction changes happened slower in the Nowcasting Period.

To summarize: the two periods are quite different in their volatility, though not in the magnitude of price changes in the trends (measured by aTMV), but in the frequency of directional changes (measured by T).

5.11.1.2 Distribution of TMV_{EXT} in Training period$_{USD/JPY}$ under threshold 0.0016

In this section, we study the historical distribution of TMV_{EXT}. We shall show later how this could be used to support nowcasting.

As explained in Table 5. 14 We call the period from 00:00:10 on 27/09/2009 to 23:59:59 on 31/12/2013 as the Training Period$_{USD/JPY}$. We call the period from 00:00:09 on 01/01/2014 to 23:50:59 on 31/12/2015 the Nowcasting Period$_{USD/JPY}$. We calculate TMV_{EXT} under threshold 0.0016 of Training Period$_{USD/JPY}$.

We calculate the TMV_{EXT} under 0.0016 for Training $Period_{USD/JPY}$. Then we collect it from every trend. In the Training $Period_{USD/JPY}$ when the threshold is equal to 0.0016, we got 17,588 TMV_{EXT}.

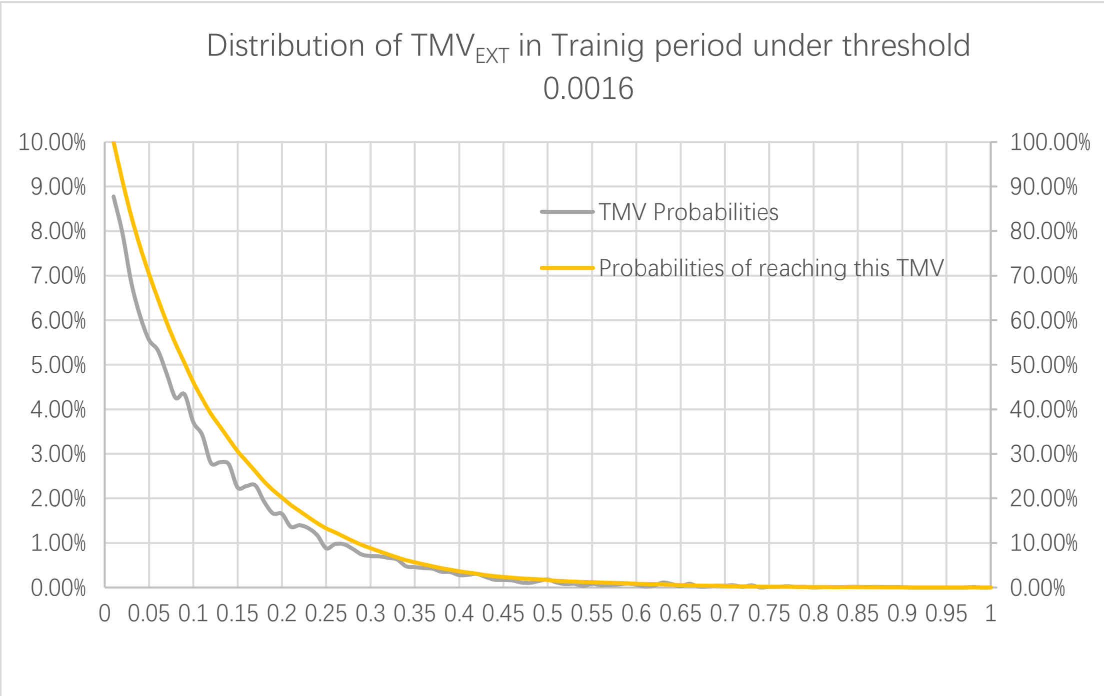

Figure 5. 22 Distribution of TMV_{EXT} in the Training $period_{USD/JPY}$ under threshold 0.0016

Figure 5. 22 show the graph of the probability of TMV_{EXT} happening in each bin and the probability of reaching a certain number of TMV_{EXT} in the Training $period_{USD/JPY}$ when the threshold is equal to 0.0016

The of TMV_{EXT} in the Training $period_{USD/JPY}$ under threshold 0.0016 that we pick			
Bin	TMV_{EXT} Frequency	TMV_{EXT} Probability	Probabilities of reaching this TMV_{EXT}
1	0		
1.07904397	1268	8.78%	100.01%
……			
5.50550629	8	0.06%	1.06%
5.58			1.00%
5.66359423	13	0.09%	0.93%
……			
8.904397	1	0.00%	0.01%

Table 5. 15 The TMV_{EXT} we pick to use under the distribution of TMV_{EXT} in the Training $period_{USD/JPY}$ under threshold 0.0016

When the threshold is equal to 0.0016, we will use the TMV_{EXT} value 5.58 which would have a probability of 1.00% happening in the Training $period_{USD/JPY}$ to nowcast the Training $period_{USD/JPY}$ and Nowcasting $period_{USD/JPY}$, which will be used in Section 5.11.2. In this section, we pick up the TMV_{EXT} only 1.00% historical probability happened in the Training $period_{USD/JPY}$

5.11.1.3 Distribution of MBM in Training $period_{USD/JPY}$ under threshold 0.0016

We use the distribution of MBM in the Training $period_{USD/JPY}$ under threshold 0.0016 for nowcasting. We run the **Nowcast Constant Algorithm** (Section 5.6.1) on the Training $period_{USD/JPY}$ for backtesting and the Nowcasting $Period_{USD/JPY}$ for assessment of the algorithm. We use the Nowcasting $period_{USD/JPY}$ in our experiments. We conduct the experiments on DC thresholds 0.0016. We use the same methodology as in

Section 5.4 to summarize the distribution of MBM for USD/JPY.

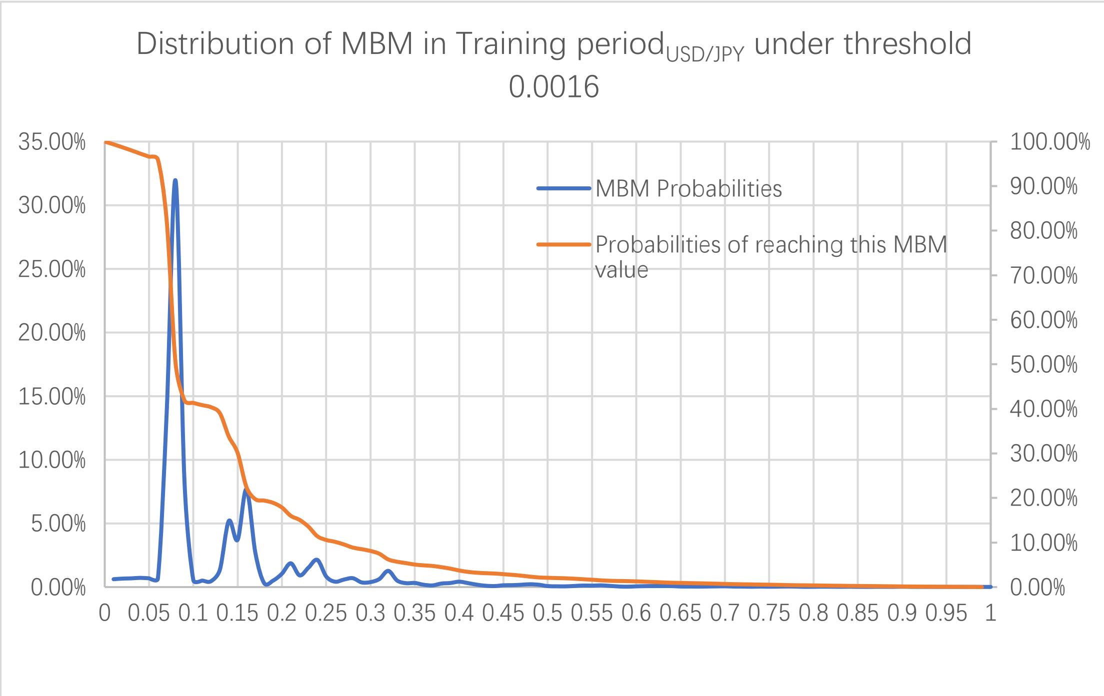

Figure 5. 23 Distribution of MBM in Training period$_{USD/JPY}$ under threshold 0.0016

Figure 5. 23 show the graph of the probability of MBM happening in each bin and the probability of reaching a certain number of MBM in the Training period$_{USD/JPY}$ when the threshold is equal to 0.0016.

Distribution of Max Below Max in Training period$_{USD/JPY}$ under Threshold 0.0016			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	3192	0.62%	100.00%
……			
0.64	401	0.08%	1.05%
0.641			1.00%
0.65	224	0.04%	0.97%
……			
1	56	0.01%	0.01%

Table 5. 16 Distribution of Max Below Max in Training period under Threshold 0.0016

When the threshold is equal to 0.0016, we will choose 0.641 as the parameter to nowcast the Training period$_{USD/JPY}$ and Nowcasting period$_{USD/JPY}$. This is because we only got a 1.00% historical probability that BM is bigger than 0.641. We will use this probability to nowcast both training and nowcasting data under 0.0016.

Now we follow the NCA nowcasting rule in Section 5.6. In this section, we use the same P_{TMV} and P_{BM} to summarize the Training period and nowcasting.

In this section, we will choose:

Under NCA in Training period$_{USD/JPY}$ when the Threshold is equal to 0.0016 we choose P_{TMV}=5.58 and P_{BM}=0.641. There is a 1% historical probability that P_{TMV} and P_{BM} values would simultaneously exceed 5.58 and 0.64, respectively.

In this case, when aTMV is bigger than 5.58 in a trend and MBM is bigger than 0.641, we will be recorded as a nowcast.

Under NCA in the Nowcasting period$_{USD/JPY}$ when the Threshold is equal to 0.0016 we choose P_{TMV}=5.58 and P_{BM}=0.641 as we introduced before.

We repeat the work as we do in the Training period$_{USD/JPY}$, any of aTMV is bigger than 5.58 and MBM is bigger than 0.641 but less than 1, we recorded it as a nowcast.

5.11.2 Result

Nowcast results in the Training period$_{USD/JPY}$ under threshold 0.0016			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	17,588	Timeliness$_{Price}$	Timeliness$_{Time}$ (seconds)
Nc: Number of nowcasts	202	Maximum	Maximum
GN: Number of good nowcasts	166	0.15445	13622
CN: Number of correct nowcasts	166	Minimum	Minimum
IN: Number of incorrect nowcasts	36	0.00	1
Nc = CN + IN	202	Median	Median
Recall = CN / TnT	0.94%	0.1	63
Precision$_{Correct}$ = CN / Nc	82.17%	Average	Average
Precision$_{Good}$ = GN / CN	100%	0.084	364
False Positive = IN / Nc	17.83%	Standard deviation	Standard deviation
		0.04	1178
Nowcast results in the Nowcasting period$_{USD/JPY}$ under threshold 0.0016			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	3,892	Timeliness$_{Price}$	Timeliness$_{Time}$ (seconds)
Nc: Number of nowcasts	43	Maximum	Maximum
GN: Number of good nowcasts	35	0.1845	5006
CN: Number of correct nowcasts	35	Minimum	Minimum
IN: Number of incorrect nowcasts	8	0.1055	2

Nc = CN + IN	43	Median	Median
Recall = CN / TnT	0.9%	0.124	185
$\text{Precision}_{\text{Correct}}$ = CN / Nc	81.39%	Average	Average
$\text{Precision}_{\text{Good}}$ = GN / CN	100%	0.1255	779
False Positive = IN / Nc	18.61%	Standard deviation	Standard deviation
		0.017	1208

Table 5. 17 Both Training $\text{period}_{\text{USD/JPY}}$ and Nowcasting $\text{period}_{\text{USD/JPY}}$ nowcasting result

Table 5. 17 Shows the result for both the Training $\text{period}_{\text{USD/JPY}}$ and the Nowcasting $\text{period}_{\text{USD/JPY}}$ nowcasting result.

Under the Training $\text{period}_{\text{USD/JPY}}$ when the threshold is 0.0016. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the Training $\text{period}_{\text{USD/JPY}}$, when the threshold is 0.0016, we have 17,588 trends in total. NCA nowcasts 202 times. In these 202 nowcasts, 166 are correct; among which all are useful. 82.17% of nowcast from the total nowcast are good ($\text{Precision}_{\text{Correct}}$), and 100.00% of correct nowcast are good nowcast ($\text{Precision}_{\text{Good}}$). We use recall as a reference: out of all trends, 0.9% are nowcasted good.

Within those good nowcasts, we summarize their timeliness in the rightmost two columns in Table 5. 17. Under the $\text{Timeliness}_{\text{Price}}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 0.15445 and the minimum value is 0. In the Training $\text{period}_{\text{USD/JPY}}$ when the threshold is 0.0016, the average $\text{timeliness}_{\text{Price}}$ is 0.084 and the median is 0.1. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So 0.084 means we manage to nowcast the new trend nearly 8.4% of the way to DCC.

$Timeliness_{Time}$ measures how early we manage to recognize a new trend before DCC. The bigger the $Timeliness_{Time}$ is, the earlier we know the new trend started. In the Training $period_{USD/JPY}$, when the threshold is 0.0016 the maximum $Timeliness_{Time}$ is 13,622 seconds and the minimum is 1. The average value of $Timeliness_{Time}$ is 364 which means, on average NCA can get around 364 seconds ahead of the DCC point. Furthermore, when we look at the standard deviation of both $Timeliness_{Price}$ and $Timeliness_{Time}$. At first, when we get the result of the standard deviation of $Timeliness_{Price}$ is equal to 0.04, this is a very small number which means most of the nowcasting prices are reasonably close to the average value of $Timeliness_{Price}$. The standard deviation of $Timeliness_{Time}$ is 1,178, which is quite a big number. This means in the good nowcast, we will get some of the new trends early before the DCC point, but some are not. As we can also see the median value of $Timeliness_{Time}$ is 63 while the average value is 364.

Table 5. 17 also shows the nowcasting results for NCA in the Nowcasting $period_{USD/JPY}$ when the threshold is 0.0016. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the nowcasting period, when the threshold is 0.0016, we have 3,892 trends in total. NCA nowcasts 43 times. In these 43 nowcasts, 35 are correct, among which 35 are useful (hence good). 81.39% of nowcast from the total nowcast are good ($Precision_{Correct}$), and 100% of

correct nowcast are good nowcast ($\text{Precision}_{\text{Good}}$). We use recall as a reference: out of all trends, 0.9%.

Within those good nowcasts, we summarize the in the rightmost two columns in Table 5. 17 their timeliness. Under the $\text{Timeliness}_{\text{Price}}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 0.1845 and the minimum value is 0.1055. In the nowcasting period when the threshold is 0.0016, the average $\text{timeliness}_{\text{Price}}$ is 0.1255 and the median is 0.124. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So, 0.12 means we manage to nowcast the new trend 12% ahead of the way to DCC. The standard deviation of $\text{Timeliness}_{\text{Price}}$ is 0.16, which is reasonably small.

$\text{Timeliness}_{\text{Time}}$ measures how early the NCA manage to recognize a new trend before the DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we manage to recognize the start of the new trend. In the Nowcasting period, under threshold 0.0016, the median $\text{Timeliness}_{\text{Time}}$ is 185 while the average value is 779. The big difference is explained by extreme values: The maximum $\text{Timeliness}_{\text{Time}}$ is 5,006 seconds and the minimum is 2. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 1,208 seconds, which is quite a big number. This means amongst the good nowcast, some are early before the DCC point, but some are not.

5.11.3 Discussion

In this section, we shall assess the performance of NCA based on the results presented so far. Key performance measures are shown below. We shall comment on NCA's nowcasts' timeliness in terms of price and time. We shall also comment on the robustness of NCA.

Summary of key performance measures under USD/JPY		
	Training Period$_{USD/JPY}$	Nowcasting Period$_{USD/JPY}$
Threshold	0.0016	0.0016
Precision$_{Correct}$	82.17%	81.39%
Precission$_{Good}$	100.00%	100.00%
Recall	0.94%	0.9%
Average Timeliness$_{Price}$	0.084	0.125
Standard deviation of Timeliness$_{Price}$	0.04	0.168
Maximum Timeliness$_{Price}$	0.1545	0.1854
Average Timeliness$_{Time}$	364 sec.	779 sec.
Standard deviation of Timeliness$_{Time}$	1,178 sec.	1,208 sec.
Maximum Timeliness$_{Time}$	13,622 sec.	5,006 sec.

Table 5. 18 Summary of key performance measures of nowcasting under USD/JPY

In summary, Table 5. 18 shows the average Timeliness$_{Price}$ in both the Training and Nowcast periods for USD/JPY under thresholds 0.0016 between 0.084 and 0.125. That means, on average, good nowcasts were made around 10% of the price from the extreme point to the DCC point. The relatively small standard deviations of Timeliness$_{Price}$ across Table 5. 18 suggest that most of the nowcasts are not far from around 10% from the extreme price to the DCC price.

Table 5. 18 also provides evidence that NCA provides timely nowcasts in terms of time: the average Timeliness in both Training and Nowcasting Periods under USD/JPY, under threshold 0.0016, ranged from 364 to 779 seconds. The standard deviations of Timeliness$_{Time}$ are big compared to the averages. This suggests that thousands of seconds can be gained through good nowcasting by NCA.

5.12.Experimental work for GBP/USD

In the previous section, we tested NCA on USD/JPY to see whether results are any different from our results on EUR/USD. In this section, we shall further examine the consistency of NCA by applying it to GBP/USD.

5.12.1 Experimental work set up

5.12.1.1 Data Used and their profiles

To prepare for our nowcasting algorithm, as in section 5.3, we will use Tick-to-tick GBP/USD rates are used throughout this section. Table 5. 19 summarizes the data used in this chapter. We separate the data into the Training and Nowcasting periods. Statistics in the former is used to determine the parameters used in our nowcasting algorithm, which will be used for nowcasting in the Nowcasting period. In this section, we will use thresholds=0.0032 which will namely 0.0032.

	Data in the Training Period$_{GBP/USD}$	Data in the Nowcasting Period$_{GBP/USD}$
Threshold	0.0032	0.0032
Period	From: 00:00:01 27th September 2009 To: 23:59:59 31st December 2013	From 00:00:01 1st January 2014 To: 23:59:59 31st December 2015
Number of transactions	71,136,369	28,026,456
Number of DCs	3,354	747
Median aTMV	1.702382	1.699942
Median T	15,813	25,116

Table 5. 19 Tick-to-tick data used in this section, with their profiles compared

We call the period from 00:00:01 on 27/09/2009 to 23:59:59 on 31/12/2013 as the **Training Period$_{GBP/USD}$**. We call the period from 00:00:01 on 01/01/2014 to 23:59:59 on 31/12/2015 the **Nowcasting Period$_{GBP/USD}$**.

Let us compare and contrast the profiles of the two periods. From the above, we can see that, under threshold 0.0032, the median aTMV values in the Training Period$_{GBP/USD}$ is 1.702382, whereas the median aTMV in the Nowcasting Period$_{GBP/USD}$ under the same threshold is 1.699942. Under these thresholds, the median aTMV in the Training Period$_{GBP/USD}$ is more or less the same as Nowcasting Period$_{GBP/USD}$.

But when we look at the Median T values, we can see that under threshold 0.0032, a trend takes 15,813 seconds to finish a trend in the Training Period$_{GBP/USD}$, but 25,116 seconds (nearly 1.7 times of the time) to finish in the Nowcasting Period$_{GBP/USD}$.

Both aTMV and T are indicators of volatility in a market-period: the former measures the magnitude of price changes and the latter measures the frequency of changes. The above comparisons suggest the following:

1. There is nearly no difference in the two periods' volatility as measured by median aTMV.
2. Measured by median T, the Nowcasting Period is much more volatile than that Training Period: direction changes happened slower in the Nowcasting Period.

To summarize: the two periods are quite different in their volatility, though not in the magnitude of price changes in the trends (measured by aTMV), but in the frequency of directional changes (measured by T).

5.12.1.2 Distribution of TMV_{EXT} in Training period$_{GBP/USD}$ under threshold 0.0032

In this section, we study the historical distribution of TMV_{EXT}. We shall show later how this could be used to support nowcasting.

As explained in Table 5. 19 We call the period from 00:00:01 on 27/09/2009 to 23:59:59 on 31/12/2013 as the Training Period$_{GBP/USD}$. We call the period from 00:00:01 on 01/01/2014 to 23:59:59 on 31/12/2015 the Nowcasting Period$_{GBP/USD}$. We calculate TMV_{EXT} under threshold 0.0032 of Training Period$_{GBP/USD}$.

We calculate the TMV_{EXT} under 0.0032 for Training Period$_{GBP/USD}$. Then we collect it from every trend. In the Training Period$_{GBP/USD}$ when the threshold is equal to 0.0032, we got 3,354 TMV_{EXT}.

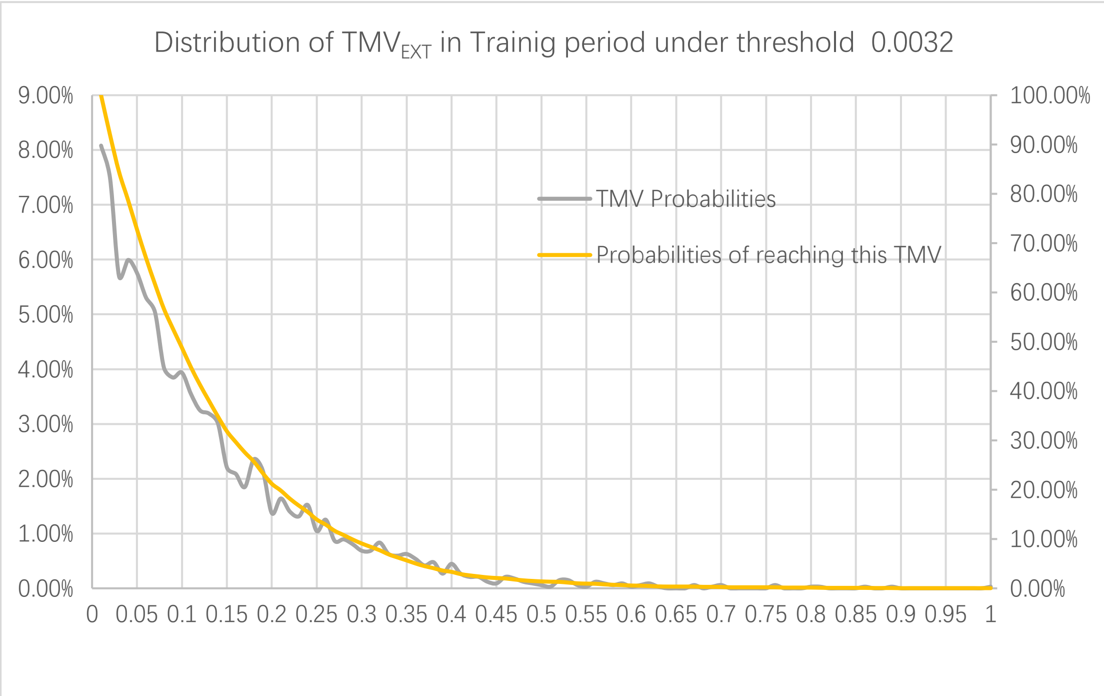

Figure 5. 24 Distribution of TMV_{EXT} in the Training period$_{GBP/USD}$ under threshold 0.0032

Figure 5. 24 show the graph of the probability of TMV_{EXT} happening in each bin and the probability of reaching a certain number of TMV_{EXT} in the Training period$_{GBP/USD}$ when the threshold is equal to 0.0032

The of TMV_{EXT} in the Training period$_{GBP/USD}$ under threshold 0.0016 that we pick			
Bin	TMV_{EXT} Frequency	TMV_{EXT} Probability	Probabilities of reaching this TMV_{EXT}
1	0		
1.08110049	271	8.08%	100.00%
……			
5.37942646	2	0.06%	1.01%
5.38			1.00%
5.46052695	1	0.03%	0.95%
……			
9.110049	1	0.03%	0.03%

Table 5. 20 The TMV_{EXT} we pick to use under the distribution of TMV_{EXT} in the Training period$_{GBP/USD}$ under threshold 0.0032

When the threshold is equal to 0.0032, we will use the TMV_{EXT} value 5.38 which would have a probability of 1.00% happening in the Training period$_{GBP/USD}$ to nowcast the Training period$_{GBP/USD}$ and Nowcasting period$_{GBP/USD}$, which will be used in Section 5.12.2. In this section, we pick up the TMV_{EXT} only 1.00% historical probability happened in the Training period$_{GBP/USD}$.

5.12.1.3 Distribution of MBM in Training period$_{GBP/USD}$ under threshold 0.0032

We use the distribution of MBM in the Training period$_{GBP/USD}$ under threshold 0.0032 for nowcasting. We run the **Nowcast Constant Algorithm** (Section 5.6.1) on the Training period$_{GBP/USD}$ for backtesting and the Nowcasting Period$_{GBP/USD}$ for assessment of the algorithm. We use the Nowcasting period$_{GBP/USD}$ in our experiments. We conduct the experiments on DC thresholds 0.0032. We use the same methodology as in

Section 5.4 to summarize the distribution of MBM for GBP/USD.

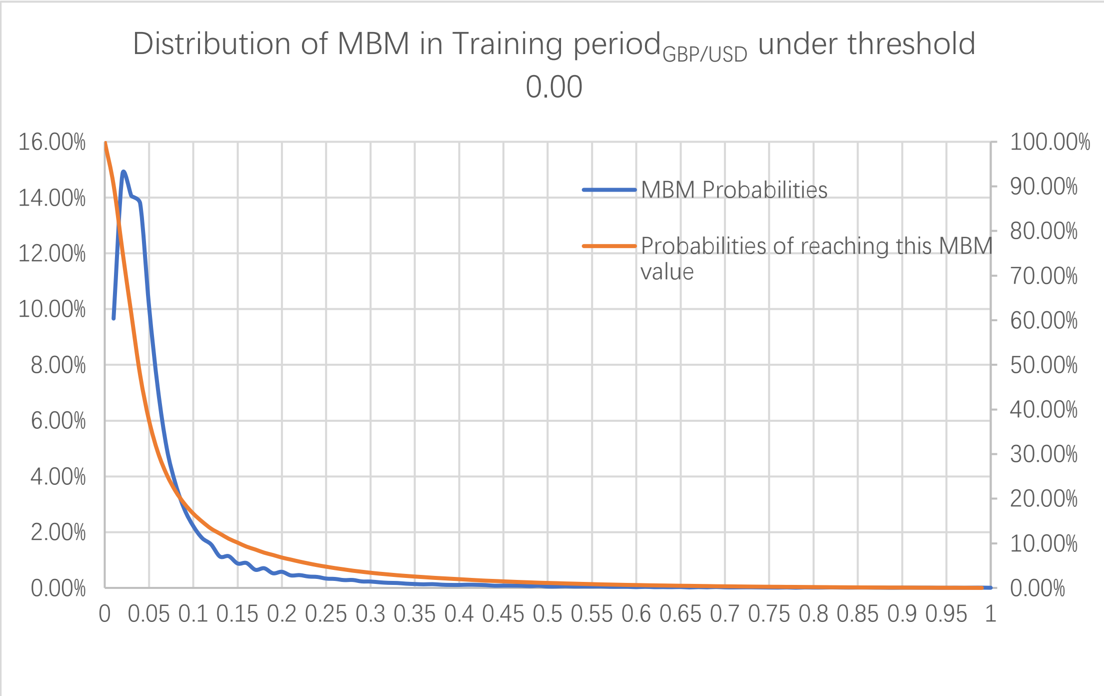

Figure 5. 25 Distribution of MBM in Training period$_{GBP/USD}$ under threshold 0.0032

Figure 5. 25 show the graph of the probability of MBM happening in each bin and the probability of reaching a certain number of MBM in the Training period$_{GBP/USD}$ when the threshold is equal to 0.0032.

Distribution of Max Below Max in Training period$_{GBP/USD}$ under Threshold 0.0032			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	18239	9.66%	100.00%
……			
0.52	115	0.06%	1.06%
0.53			1.00%
0.54	100	0.05%	0.95%
……			
1	10	0.01%	0.01%

Table 5. 21 Distribution of Max Below Max in Training period$_{GBP/USD}$ under Threshold 0.0032

When the threshold is equal to 0.0016, we will choose 0.53 as the parameter to nowcast the Training period$_{GBP/USD}$ and Nowcasting period$_{GBP/USD}$. We only got a 1.00% historical probability that BM is bigger than 0.53. We will use this probability to nowcast both training and nowcasting data under 0.0032.

Now we follow the NCA nowcasting rule in Section 5.6, and in this section, we use the same P_{TMV} and P_{BM} to summarize the Training period and nowcasting.

In this section, we will choose:

Under NCA in Training period$_{GBP/USD}$ when Threshold is equal to 0.0032 we choose P_{TMV}=5.38 and P_{BM}=0.53. There is a 1% historical probability that the P_{TMV} and P_{BM} values would simultaneously exceed 5.38 and 0.53, respectively.

In this case, when aTMV is bigger than 5.38 in a trend and MBM is bigger than 0.53, we will be recorded as a nowcast.

Under NCA in Nowcasting period$_{GBP/USD}$ when Threshold is equal to 0.0032 we choose P_{TMV}=5.38 and P_{BM}=0.53 as we introduced before.

We repeat the work as we do in the Training period$_{GBP/USD}$, any of aTMV is bigger than 5.38 and MBM is bigger than 0.53 but less than 1, we recorded it as a nowcast.

5.12.2 Result

Nowcast results in the Training period$_{GBP/USD}$ under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	3,354	Timeliness$_{Price}$	Timeliness$_{Time}$ (seconds)
Nc: Number of nowcasts	45	Maximum	Maximum
GN: Number of good nowcast	32	0.0352	222,260
CN: Number of correct nowcasts	32	Minimum	Minimum
IN: Number of incorrect nowcasts	13	0.0025	8
Nc = CN + IN	45	Median	Median
Recall = CN / TnT	0.95%	0.0027	1908
Precision$_{Correct}$ = CN / Nc	71.11%	Average	Average
Precision$_{Good}$ = GN / CN	100%	0.027	13,335
False Positive = IN / Nc	28.89%	Standard deviation	Standard deviation
		0.00020	39,546
Nowcast results in the Nowcasting period$_{GBP/USD}$ under threshold 0.0032			
Nowcast Indicator	Value	Timeliness analysis for Good nowcast	
TnT: Total number of trends	747	Timeliness$_{Price}$	Timeliness$_{Time}$ (seconds)
Nc: Number of nowcasts	9	Maximum	Maximum
GN: Number of good nowcast	7	0.00428	13,559
CN: Number of correct	7	Minimum	Minimum

nowcasts			
IN: Number of incorrect nowcasts	2	0.00251	11
Nc = CN + IN	9	Median	Median
Recall = CN / TnT	0.93%	0.00271	2,607
$Precision_{Correct}$ = CN / Nc	77.78%	Average	Average
$Precision_{Good}$ = GN / CN	100%	0.00292	5,189
False Positive = IN / Nc	22.22%	Standard deviation	Standard deviation
		0.0006	5,461

Table 5. 22 Both Training $period_{GBP/USD}$ and Nowcasting $period_{GBP/USD}$ nowcasting result

Table 5. 22 Shows the result for both Training $period_{GBP/USD}$ and Nowcasting $period_{GBP/USD}$ nowcasting result.

Under Training $period_{GBP/USD}$ when the threshold is 0.0032. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the Training $period_{GBP/USD}$, when the threshold is 0.0032, we have 3,354 trends in total. NCA nowcasts 45 times. In these 45 nowcasts, 32 are correct; among which all are useful. 71.11% of nowcast from the total nowcast are good ($Precision_{Correct}$), and 100.00% of correct nowcast are good nowcast ($Precision_{Good}$). We use recall as a reference: out of all trends, 0.95% are nowcasted good.

Within those good nowcasts, we summarize their timeliness in the rightmost two columns in Table 5. 22. Under the $Timeliness_{Price}$ column, the maximum value of $Timeliness_{Price}$ is 0.0352 and the minimum value is 0.0025. In the Training $period_{GBP/USD}$ when the threshold is 0.0032, the average $timeliness_{Price}$ is 0.0027 and the median is 0.0027. Readers are reminded that the theoretical TMV at DC confirmation is 1 by definition. So 0.0027 means we manage to nowcast the new trend nearly 0.27% of the way to DCC.

$\text{Timeliness}_{\text{Time}}$ measures how early we manage to recognize a new trend before DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we know the new trend started. In the Training $\text{period}_{\text{GBP/USD}}$, when the threshold is 0.0032 the maximum $\text{Timeliness}_{\text{Time}}$ is 222,260 seconds and the minimum is 8. The average value of $\text{Timeliness}_{\text{Time}}$ is 13,335 which means, on average NCA can get around 13,335 seconds ahead of the DCC point. Furthermore, when we look at the standard deviation of both $\text{Timeliness}_{\text{Price}}$ and $\text{Timeliness}_{\text{Time}}$. At first, when we get the result of the standard deviation of $\text{Timeliness}_{\text{Price}}$ is equal to 0.0002, this is a very small number which means most of the nowcasted prices are reasonably close to the average value of $\text{Timeliness}_{\text{Price}}$. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 39,546, which is quite a big number. This means in the good nowcast, we will get some of the new trends early before the DCC point, but some are not.

The table above also shows the nowcasting results for NCA in the Nowcasting $\text{period}_{\text{GBP/USD}}$ when the threshold is 0.0032. In the first two columns, we record the values of different performance indicators. In these two columns, we can see that in the nowcasting period, when the threshold is 0.0032, we have 747 trends in total. NCA nowcasts 9 times. In these 9 nowcasts, 7 are correct, among which 7 are useful (hence good). 77.78% of nowcast from the total nowcast are good ($\text{Precision}_{\text{Correct}}$), and 100% of correct nowcast are good nowcast ($\text{Precision}_{\text{Good}}$). We use recall as a reference: out of all trends, 0.9%.

Within those good nowcasts, we summarize the in the rightmost two columns in Table 5. 22 their timeliness. Under the $\text{Timeliness}_{\text{Price}}$ column, the maximum value of $\text{Timeliness}_{\text{Price}}$ is 0.00428 and the minimum value is 0.00251. In the nowcasting period when the threshold is 0.0032, the average $\text{timeliness}_{\text{Price}}$ is 0.0029 and the median is 0.0027. The standard deviation of $\text{Timeliness}_{\text{Price}}$ is 0.32, which is reasonably small(0.0006）.

$\text{Timeliness}_{\text{Time}}$ measures how early the NCA manage to recognize a new trend before the DCC. The bigger the $\text{Timeliness}_{\text{Time}}$ is, the earlier we manage to recognize the start of the new trend. In the Nowcasting $\text{period}_{\text{GBP/USD}}$, under threshold 0.0032, the median $\text{Timeliness}_{\text{Time}}$ is 2,607 while the average value is 5,189. The big difference is explained by extreme values: The maximum $\text{Timeliness}_{\text{Time}}$ is 13,559 seconds and the minimum is 11. The standard deviation of $\text{Timeliness}_{\text{Time}}$ is 5,461 seconds, which is quite a big number. This means amongst the good nowcast, some are early before the DCC point, but some are not.

5.12.3 Discussion

In this section, we shall assess the performance of NCA based on the results presented so far. Key performance measures are shown below. We shall comment on NCA's nowcasts' timeliness in terms of price and time. We shall also comment on the robustness of NCA.

Summary of key performance measures under GBP/USD		
	Training Period$_{GBP/USD}$	Nowcasting Period$_{GBP/USD}$
Threshold	0.0032	0.0032
Precision$_{Correct}$	71.11%	77.78%
Precission$_{Good}$	100.00%	100.00%
Recall	0.95%	0.9%
Average Timeliness$_{Price}$	0.0027	0.0029
Standard deviation of Timeliness$_{Price}$	0.0002	0.0006
Maximum Timeliness$_{Price}$	0.1545	0.00428
Average Timeliness$_{Time}$	13,335 sec.	5,189 sec.
Standard deviation of Timeliness$_{Time}$	39,546 sec.	5,461 sec.
Maximum Timeliness$_{Time}$	222,260 sec.	13,559 sec.

Table 5. 23 Summary of key performance measures of nowcasting under GBP/USD

In summary, Table 5. 23 shows that the average Timeliness$_{Price}$ in both the Training and Nowcast periods for USD/JPY under thresholds 0.0032 between 0.0027 and 0.0029. That means, on average, good nowcasts were made around 2.7% of the price from the extreme point to the DCC point. The relatively small standard deviations of Timeliness$_{Price}$ across Table 5. 23 suggest that most of the nowcasts are not far from around 10% from the extreme price to the DCC price.

Table 5. 23 also provides evidence that NCA provides timely nowcasts in terms of time: the average Timeliness in both Training and Nowcasting Periods under GBP/USD, under threshold 0.0032, ranging from 39,546 to 5,461 seconds. The standard deviations of $\text{Timeliness}_{\text{Time}}$ are big compared to the averages. This suggests that thousands of seconds can be gained through good nowcasting by NCA.

5.13. Conclusion on Nowcasting

In summary, this chapter presents a nowcasting problem in DC: the problem of recognizing a new trend before reaching the DC confirmation (DCC) point (Section 5.1). We have introduced a new algorithm, namely NCA (Section 5.6.1), for this nowcasting problem. NCA makes use of two indicators in DC: aTMV (05.2) and BM (0Section 5.2). The principle of the method is: as a trend reaches a higher aTMV, the probability of directional change happening increases (see Section 5.4). The higher the price bounces back from the maximum aTMV, the more likely that the direction has already changed (Section 5.5). Nowcast in NCA is based on the simultaneous happening of these two events with a reasonably big aTMV and BM thresholds (Section 5.6.4).

Experiments were set up (Section 5.7) to run NCA on a substantial amount of high-frequency data (Section 5.3). Results (Section 5.8) suggest that NCA is effective, in the sense that it can nowcast new trends way before

DCC points, both in terms of both price ($\text{Timeliness}_{\text{Price}}$, Section 5.9.1) and time ($\text{Timeliness}_{\text{Time}}$, Section 5.9.2). We have also shown that NCA is robust across market conditions (Section 5.9.3) and thresholds (Section 0). As parameters used by NCA were picked arbitrarily, there is still plenty of room for improvement (Section 5.9.4).

In Chapter 3, we presented a guideline for limiting the thresholds to be used for DC. We suggested that if we use a threshold that is too small, DC summaries contain a lot of noise. In section 5.10. we run experiments to test our guideline this guideline. The results in section 5.10.2 show that if we use a small threshold that is out of our guideline range, no nowcasting could be observed under the training period because most overshoots are too small (see discussion in Section 5.10.3). In conclusion, this set of experiments supports the usefulness of our guideline in Chapter 3.

NCA is proposed as a proof of concept. Parameters for NCA were picked arbitrarily. If NCA were to be developed to be operational, then the parameters should be fine-tuned. This will be left for future research.

More experiments have been conducted to test the robustness of NCA. For experiments with USD/JPY (Section 5.11) and GBP/USD (Section 5.12), we picked different PBM and PTMV parameters (we used 1% historical probability for both). The results show that the performance of NCA is insensitive to the parameters picked. In other words, preliminary research suggests that NCA is robust.

More experimental results (on USD/JPY and GBP/USD) were presented in Sections 5.11 and 5.12. The results show that NCA's performance was robust across markets. They also show that NCA's performance is not sensitive to the NCA parameters picked.

6. Chapter Six: Conclusion

6.1. Summary

This section summarizes the work done in this thesis. This research is built under the framework of Directional Change (DC). It builds on profiling research conducted by Ran Tao (Tao PhD 2018, Tsang et al 2017), which defined several indicators.

6.1.1. The usable range of thresholds for profiling

We have proposed two guidelines for assessing the threshold that we have chosen is too small or too big. This is a data-driven approach that data will tell us whether a threshold is too small or too big. It is based on Olsen's empirical research (Glattfelder, 2011) and profiling concepts proposed by Tsang et al (2017) and Tao (2018).

6.1.2. Monitor markets with multiple indicators

We have tracked the EUR/USD market (the most traded currency market) over a long period of time (June 2009 to July 2016) using standard deviation in returns under Time Series (TS) and median $aTMV_{EXT}$ and NDC under DC. We show that while volatility could be high under multiple measures, it could be low under one indicator, but high under another. This suggests that all of the above indicators are useful.

6.1.3. Nowcasting with multiple indicators

By definition, that DC is only confirmed in hindsight. A trend may have reversed, but it is not recognized before we reach the next DC confirmation point. We have created an algorithm called Nowcasting Constant Algorithm (NCA) with new indicators and demonstrated how, together with established DC indicators, it could help us nowcast whether a new trend has begun in DC. NDC has been tested across markets with different parameters. Results suggest that the performance of NDC is consistent across markets and parameters.

6.2. Contributions

This thesis contributes to the continuing research on the use of DC in the financial market:

Contribution 1: Usable range of thresholds for profiling

> This is the first piece of work that studies what is the range of usable thresholds for profiling (Section 3.6).

This is important as when thresholds which are either too small or too big are used, the profiles observed are distorted by noise. By proposing a data-driven approach for determining the useful range of thresholds, we contribute to scientific research in DC.

Contribution 2: Tracking and monitoring markets with multiple indicators

We have confirmed the hypothesis that each volatility indicator provides a partial view of volatility in the market. We have demonstrated how DC indicators could complement TS in tracking the market for volatility information (Section 4.6). The new indicators we set up in this section lay the foundation for the next Contribution.

Contribution 3: Formulation of nowcasting problem in DC

We have formulated a nowcasting problem in DC: the problem is to detect directional change before the DC confirmation point is reached (Section 5.1).

Contribution 4: Introduction of new DC indicator: BM

We have introduced BM, a new indicator in DC (Definition 5.4, Section 5.2.3). We have shown that it is useful for nowcasting new trends.

Contribution 5: Nowcasting new trends

We have proposed an algorithm (NCA, Section 5.6.1) to nowcast new trends using aTMV and BM. We have demonstrated how this algorithm allows us to nowcast directional effectively.

6.3. Future work

We have shown in Chapter 4 that by using DC indicators to track the market, we could not only reinforce what we learn from TC, we could also gain information beyond TS. In this thesis, we have only used two indicators that represent different kinds of volatility under DC. Is there any chance that we can combine DC indicators and SD? If yes, could we introduce a formulate that contain DC and TS indicators altogether? This is the first direction that we would like to explore in the future.

In Chapter 5, we have only proposed a simple nowcasting algorithm, NCA. Parameters for NCA were picked arbitrarily. They show that the performance of NCA does not rely on fine-tuning of its parameters. That does not mean that NCA could not be improved with fine-tuned parameters.

NCA uses fixed P_{TMV} and P_{BM} values. There is no reason why P_{BM} should not be varied depending on Max[13]. In DC, when aTMV is high, the trend is more likely to end soon (see Section 5.2.1). Therefore, when Max is high, we may nowcast a change of direction with a smaller BM value (see Section 5.2.2). On the other hand, when the aTMV of Max is low, we demand a bigger BM to increase the probability that we enter into new

[13] Readers are reminded that Max is the maximum aTMV in the current trend (Section 5.2.2).

trends.

Following this principle, let us introduce a new algorithm which we shall call Nowcast Progressive Algorithm (NPA). NFA takes three parameters, P_{TMV}, P_{BM1} and P_{BM2}. Here P_{BM1} and P_{BM2} are the parameters that we choose used under different ranges of P_{TMV} we choose, with P_{BM1} smaller P_{BM2}. When the aTMV of Max reaches the predefined threshold P_{TMV}, we apply the smaller BM value P_{BM1}: a new trend is nowcasted if the BM of the current price is greater than P_{BM1}. When the aTMV of Max is less than P_{TMV}, we apply the bigger BM value P_{BM2}.

Nowcast Progressive Algorithm (P_{TMV}, P_{BM1} and P_{BM2}):

Given a current transaction *ct*:

- If (1) aTMV(*Max*) $\geq P_{TMV}$ and (2) BM(*ct*) $\geq P_{BM1}$, then we nowcast that *ct* is in a new trend;
- If (3) aTMV(*Max*)$<P_{TMV}$ and (4) BM(*ct*) $\geq P_{BM2}$, then we nowcast that *ct* is in a new trend
- Otherwise, we do not nowcast that *ct* is in a new trend.

For example, we track each new transaction as it emerges. We record the Max of the current trend. If its aTMV of Max is greater than the P_{TMV} value, then we track the BM value of the transactions. As soon as BM(*ct*) exceeds

P_{BM1}, we nowcast that a new trend has started. If the aTMV of Max is smaller than P_{TMV}, then P_{BM2} is applied.

It is worth noting that both P_{BM1} and P_{BM2} cannot be bigger than one, this is because once BM is bigger than one, we have already entered into a new trend. In other words:

$$0 < P_{BM1} < P_{BM2} < 1$$

There is no reason why one should stop at NPA. With deeper analysis, one could extend NPA by making P_{BM} a function of P_{TMV}: as the aTMV of Max increases, P_{BM} is progressively reduced. This will be left for future research.

■

7. Appendix

Table A.1. Distribution of TMV_{EXT} in Training period under threshold 0.0016

Bin	TMV_{EXT} Frequency	TMV_{EXT} Probability	Probabilities of reaching this TMV_{EXT}
1	-	-	100.00%
1.09575712	1101	11.53%	88.47%
1.19151424	897	9.39%	79.08%
1.28727136	729	7.63%	71.45%
1.38302848	689	7.21%	64.24%
1.4787856	595	6.23%	58.01%
1.57454272	570	5.97%	52.04%
1.67029984	524	5.49%	46.56%
1.76605696	437	4.57%	41.98%
1.86181408	370	3.87%	38.11%
1.9575712	386	4.04%	34.07%
2.05332832	321	3.36%	30.71%
2.14908544	279	2.92%	27.78%
2.24484256	276	2.89%	24.90%
2.34059968	227	2.38%	22.52%
2.4363568	203	2.13%	20.39%
2.53211392	170	1.78%	18.61%
2.62787104	159	1.66%	16.95%
2.72362816	164	1.72%	15.23%
2.81938528	140	1.47%	13.77%
2.9151424	140	1.47%	12.30%
3.01089952	99	1.04%	11.26%
3.10665664	102	1.07%	10.20%
3.20241376	102	1.07%	9.13%
3.29817088	85	0.89%	8.24%
3.393928	69	0.72%	7.52%
3.48968512	63	0.66%	6.86%
3.58544224	56	0.59%	6.27%
3.68119936	73	0.76%	5.51%
3.77695648	55	0.58%	4.93%
3.8727136	60	0.63%	4.30%
3.96847072	31	0.32%	3.98%
4.06422784	29	0.30%	3.67%
4.15998496	39	0.41%	3.27%

4.25574208	24	0.25%	3.02%
4.3514992	30	0.31%	2.70%
4.44725632	26	0.27%	2.43%
4.54301344	22	0.23%	2.20%
4.63877056	18	0.19%	2.01%
4.73452768	10	0.10%	1.91%
4.8302848	17	0.18%	1.73%
4.92604192	11	0.12%	1.61%
5.02179904	15	0.16%	1.46%
5.11755616	13	0.14%	1.32%
5.21331328	11	0.12%	1.20%
5.3090704	14	0.15%	1.06%
5.40482752	8	0.08%	0.97%
5.50058464	7	0.07%	0.90%
5.59634176	8	0.08%	0.82%
5.69209888	5	0.05%	0.76%
5.787856	9	0.09%	0.67%
5.88361312	6	0.06%	0.61%
5.97937024	7	0.07%	0.53%
6.07512736	4	0.04%	0.49%
6.17088448	4	0.04%	0.45%
6.2666416	6	0.06%	0.39%
6.36239872	2	0.02%	0.37%
6.45815584	6	0.06%	0.30%
6.55391296	5	0.05%	0.25%
6.64967008	2	0.02%	0.23%
6.7454272	2	0.02%	0.21%
6.84118432	4	0.04%	0.17%
6.93694144	0	0.00%	0.17%
7.03269856	0	0.00%	0.17%
7.12845568	2	0.02%	0.15%
7.2242128	1	0.01%	0.14%
7.31996992	0	0.00%	0.14%
7.41572704	1	0.01%	0.13%
7.51148416	0	0.00%	0.13%
7.60724128	0	0.00%	0.13%
7.7029984	1	0.01%	0.12%
7.79875552	1	0.01%	0.10%
7.89451264	1	0.01%	0.09%
7.99026976	2	0.02%	0.07%
8.08602688	0	0.00%	0.07%
8.181784	1	0.01%	0.06%
8.27754112	1	0.01%	0.05%

8.37329824	1	0.01%	0.04%
8.46905536	0	0.00%	0.04%
8.56481248	0	0.00%	0.04%
8.6605696	1	0.01%	0.03%
8.75632672	1	0.01%	0.02%
8.85208384	1	0.01%	0.01%
8.94784096	0	0.00%	0.01%
9.04359808	0	0.00%	0.01%
9.1393552	0	0.00%	0.01%
9.23511232	0	0.00%	0.01%
9.33086944	0	0.00%	0.01%
9.42662656	0	0.00%	0.01%
9.52238368	0	0.00%	0.01%
9.6181408	0	0.00%	0.01%
9.71389792	0	0.00%	0.01%
9.80965504	0	0.00%	0.01%
9.90541216	0	0.00%	0.01%
10.00116928	0	0.00%	0.01%
10.0969264	0	0.00%	0.01%
10.19268352	0	0.00%	0.01%
10.28844064	0	0.00%	0.01%
10.38419776	0	0.00%	0.01%
10.47995488	0	0.00%	0.01%
10.575712	1	0.01%	0.01%

Table A.2. Distribution of TMV_{EXT} in the Nowcasting period under threshold 0.0016

Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1			100.00%
1.09157793	557	11.99%	88.01%
1.18315586	491	10.57%	77.43%
1.27473379	351	7.56%	69.88%
1.36631172	319	6.87%	63.01%
1.45788965	287	6.18%	56.83%
1.54946758	226	4.87%	51.96%
1.64104551	213	4.59%	47.37%
1.73262344	191	4.11%	43.26%
1.82420137	170	3.66%	39.60%
1.9157793	152	3.27%	36.33%
2.00735723	131	2.82%	33.51%
2.09893516	137	2.95%	30.56%
2.19051309	162	3.49%	27.07%
2.28209102	113	2.43%	24.63%
2.37366895	134	2.89%	21.75%
2.46524688	94	2.02%	19.72%
2.55682481	108	2.33%	17.40%
2.64840274	84	1.81%	15.59%
2.73998067	67	1.44%	14.15%
2.8315586	65	1.40%	12.75%
2.92313653	74	1.59%	11.15%
3.01471446	56	1.21%	9.95%
3.10629239	46	0.99%	8.96%
3.19787032	42	0.90%	8.05%
3.28944825	27	0.58%	7.47%
3.38102618	36	0.78%	6.70%
3.47260411	29	0.62%	6.07%
3.56418204	21	0.45%	5.62%
3.65575997	29	0.62%	5.00%
3.7473379	22	0.47%	4.52%
3.83891583	22	0.47%	4.05%
3.93049376	13	0.28%	3.77%
4.02207169	16	0.34%	3.42%
4.11364962	15	0.32%	3.10%
4.20522755	13	0.28%	2.82%
4.29680548	11	0.24%	2.58%

4.38838341	10	0.22%	2.37%
4.47996134	12	0.26%	2.11%
4.57153927	12	0.26%	1.85%
4.6631172	8	0.17%	1.68%
4.75469513	8	0.17%	1.51%
4.84627306	3	0.06%	1.44%
4.93785099	8	0.17%	1.27%
5.02942892	2	0.04%	1.23%
5.12100685	4	0.09%	1.14%
5.21258478	8	0.17%	0.97%
5.30416271	4	0.09%	0.88%
5.39574064	7	0.15%	0.73%
5.48731857	3	0.06%	0.67%
5.5788965	4	0.09%	0.58%
5.67047443	0	0.00%	0.58%
5.76205236	2	0.04%	0.54%
5.85363029	3	0.06%	0.47%
5.94520822	4	0.09%	0.39%
6.03678615	1	0.02%	0.37%
6.12836408	3	0.06%	0.30%
6.21994201	1	0.02%	0.28%
6.31151994	1	0.02%	0.26%
6.40309787	1	0.02%	0.24%
6.4946758	1	0.02%	0.22%
6.58625373	1	0.02%	0.19%
6.67783166	0	0.00%	0.19%
6.76940959	0	0.00%	0.19%
6.86098752	2	0.04%	0.15%
6.95256545	0	0.00%	0.15%
7.04414338	1	0.02%	0.13%
7.13572131	0	0.00%	0.13%
7.22729924	0	0.00%	0.13%
7.31887717	1	0.02%	0.11%
7.4104551	0	0.00%	0.11%
7.50203303	1	0.02%	0.09%
7.59361096	1	0.02%	0.06%
7.68518889	0	0.00%	0.06%
7.77676682	0	0.00%	0.06%
7.86834475	0	0.00%	0.06%
7.95992268	0	0.00%	0.06%
8.05150061	0	0.00%	0.06%
8.14307854	0	0.00%	0.06%
8.23465647	0	0.00%	0.06%

8.3262344	0	0.00%	0.06%
8.41781233	0	0.00%	0.06%
8.50939026	0	0.00%	0.06%
8.60096819	0	0.00%	0.06%
8.69254612	0	0.00%	0.06%
8.78412405	1	0.02%	0.04%
8.87570198	0	0.00%	0.04%
8.96727991	0	0.00%	0.04%
9.05885784	0	0.00%	0.04%
9.15043577	1	0.02%	0.02%
9.2420137	0	0.00%	0.02%
9.33359163	0	0.00%	0.02%
9.42516956	0	0.00%	0.02%
9.51674749	0	0.00%	0.02%
9.60832542	0	0.00%	0.02%
9.69990335	0	0.00%	0.02%
9.79148128	0	0.00%	0.02%
9.88305921	0	0.00%	0.02%
9.97463714	0	0.00%	0.02%
10.06621507	0	0.00%	0.02%
10.157793	1	0.02%	

Table A.3. Distribution of TMV_{EXT} in the Training period under threshold 0.0032			
Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1			100.00%
1.07790733	154	7.50%	92.50%
1.15581466	144	7.02%	85.48%
1.23372199	108	5.26%	80.21%
1.31162932	129	6.29%	73.93%
1.38953665	126	6.14%	67.79%
1.46744398	109	5.31%	62.48%
1.54535131	99	4.82%	57.65%
1.62325864	98	4.78%	52.88%
1.70116597	79	3.85%	49.03%
1.7790733	65	3.17%	45.86%
1.85698063	82	4.00%	41.86%
1.93488796	51	2.49%	39.38%
2.01279529	57	2.78%	36.60%
2.09070262	56	2.73%	33.87%
2.16860995	47	2.29%	31.58%
2.24651728	49	2.39%	29.19%
2.32442461	42	2.05%	27.14%
2.40233194	36	1.75%	25.39%
2.48023927	46	2.24%	23.15%
2.5581466	36	1.75%	21.39%
2.63605393	35	1.71%	19.69%
2.71396126	22	1.07%	18.62%
2.79186859	29	1.41%	17.20%
2.86977592	18	0.88%	16.33%
2.94768325	25	1.22%	15.11%
3.02559058	20	0.97%	14.13%
3.10349791	28	1.36%	12.77%
3.18140524	20	0.97%	11.79%
3.25931257	15	0.73%	11.06%
3.3372199	18	0.88%	10.19%
3.41512723	19	0.93%	9.26%
3.49303456	11	0.54%	8.72%
3.57094189	14	0.68%	8.04%
3.64884922	12	0.58%	7.46%
3.72675655	14	0.68%	6.77%
3.80466388	7	0.34%	6.43%

3.88257121	9	0.44%	5.99%
3.96047854	9	0.44%	5.56%
4.03838587	7	0.34%	5.21%
4.1162932	11	0.54%	4.68%
4.19420053	10	0.49%	4.19%
4.27210786	6	0.29%	3.90%
4.35001519	6	0.29%	3.61%
4.42792252	6	0.29%	3.31%
4.50582985	4	0.19%	3.12%
4.58373718	4	0.19%	2.92%
4.66164451	4	0.19%	2.73%
4.73955184	5	0.24%	2.49%
4.81745917	3	0.15%	2.34%
4.8953665	2	0.10%	2.24%
4.97327383	2	0.10%	2.14%
5.05118116	1	0.05%	2.10%
5.12908849	2	0.10%	2.00%
5.20699582	4	0.19%	1.80%
5.28490315	2	0.10%	1.71%
5.36281048	5	0.24%	1.46%
5.44071781	5	0.24%	1.22%
5.51862514	4	0.19%	1.02%
5.59653247	0	0.00%	1.02%
5.6744398	3	0.15%	0.88%
5.75234713	1	0.05%	0.83%
5.83025446	3	0.15%	0.68%
5.90816179	2	0.10%	0.58%
5.98606912	1	0.05%	0.54%
6.06397645	0	0.00%	0.54%
6.14188378	0	0.00%	0.54%
6.21979111	2	0.10%	0.44%
6.29769844	1	0.05%	0.39%
6.37560577	0	0.00%	0.39%
6.4535131	1	0.05%	0.34%
6.53142043	0	0.00%	0.34%
6.60932776	0	0.00%	0.34%
6.68723509	0	0.00%	0.34%
6.76514242	0	0.00%	0.34%
6.84304975	0	0.00%	0.34%
6.92095708	0	0.00%	0.34%
6.99886441	1	0.05%	0.29%
7.07677174	0	0.00%	0.29%
7.15467907	1	0.05%	0.24%

7.2325864	0	0.00%	0.24%
7.31049373	0	0.00%	0.24%
7.38840106	0	0.00%	0.24%
7.46630839	0	0.00%	0.24%
7.54421572	1	0.05%	0.19%
7.62212305	0	0.00%	0.19%
7.70003038	0	0.00%	0.19%
7.77793771	1	0.05%	0.15%
7.85584504	0	0.00%	0.15%
7.93375237	0	0.00%	0.15%
8.0116597	0	0.00%	0.15%
8.08956703	0	0.00%	0.15%
8.16747436	0	0.00%	0.15%
8.24538169	2	0.10%	0.05%
8.32328902	0	0.00%	0.05%
8.40119635	0	0.00%	0.05%
8.47910368	0	0.00%	0.05%
8.55701101	0	0.00%	0.05%
8.63491834	0	0.00%	0.05%
8.71282567	0	0.00%	0.05%
8.790733	1	0.05%	

Table A.4. Distribution of TMV_{EXT} in the Nowcasting period under threshold 0.0032			
Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1			100.00%
1.05194809	39	3.29%	96.71%
1.10389618	72	6.08%	90.63%
1.15584427	61	5.15%	85.49%
1.20779236	129	10.89%	74.60%
1.25974045	75	6.33%	68.27%
1.31168854	56	4.73%	63.54%
1.36363663	54	4.56%	58.99%
1.41558472	67	5.65%	53.33%
1.46753281	55	4.64%	48.69%
1.5194809	28	2.36%	46.33%
1.57142899	28	2.36%	43.97%
1.62337708	22	1.86%	42.11%
1.67532517	21	1.77%	40.34%
1.72727326	29	2.45%	37.89%
1.77922135	23	1.94%	35.95%
1.83116944	27	2.28%	33.67%
1.88311753	19	1.60%	32.07%
1.93506562	26	2.19%	29.87%
1.98701371	17	1.43%	28.44%
2.0389618	20	1.69%	26.75%
2.09090989	12	1.01%	25.74%
2.14285798	22	1.86%	23.88%
2.19480607	13	1.10%	22.78%
2.24675416	12	1.01%	21.77%
2.29870225	11	0.93%	20.84%
2.35065034	7	0.59%	20.25%
2.40259843	10	0.84%	19.41%
2.45454652	9	0.76%	18.65%
2.50649461	11	0.93%	17.72%
2.5584427	11	0.93%	16.79%
2.61039079	14	1.18%	15.61%
2.66233888	12	1.01%	14.60%
2.71428697	10	0.84%	13.76%
2.76623506	10	0.84%	12.91%
2.81818315	7	0.59%	12.32%
2.87013124	10	0.84%	11.48%

2.92207933	9	0.76%	10.72%
2.97402742	6	0.51%	10.21%
3.02597551	3	0.25%	9.96%
3.0779236	5	0.42%	9.54%
3.12987169	5	0.42%	9.11%
3.18181978	2	0.17%	8.95%
3.23376787	3	0.25%	8.69%
3.28571596	7	0.59%	8.10%
3.33766405	6	0.51%	7.59%
3.38961214	4	0.34%	7.26%
3.44156023	6	0.51%	6.75%
3.49350832	6	0.51%	6.24%
3.54545641	5	0.42%	5.82%
3.5974045	5	0.42%	5.40%
3.64935259	3	0.25%	5.15%
3.70130068	2	0.17%	4.98%
3.75324877	3	0.25%	4.73%
3.80519686	1	0.08%	4.64%
3.85714495	3	0.25%	4.39%
3.90909304	4	0.34%	4.05%
3.96104113	1	0.08%	3.97%
4.01298922	0	0.00%	3.97%
4.06493731	3	0.25%	3.71%
4.1168854	4	0.34%	3.38%
4.16883349	2	0.17%	3.21%
4.22078158	1	0.08%	3.12%
4.27272967	2	0.17%	2.95%
4.32467776	2	0.17%	2.78%
4.37662585	1	0.08%	2.70%
4.42857394	2	0.17%	2.53%
4.48052203	0	0.00%	2.53%
4.53247012	2	0.17%	2.36%
4.58441821	0	0.00%	2.36%
4.6363663	4	0.34%	2.03%
4.68831439	1	0.08%	1.94%
4.74026248	0	0.00%	1.94%
4.79221057	3	0.25%	1.69%
4.84415866	2	0.17%	1.52%
4.89610675	0	0.00%	1.52%
4.94805484	6	0.51%	1.01%
5.00000293	2	0.17%	0.84%
5.05195102	0	0.00%	0.84%
5.10389911	2	0.17%	0.68%

5.1558472	0	0.00%	0.68%
5.20779529	2	0.17%	0.51%
5.25974338	0	0.00%	0.51%
5.31169147	1	0.08%	0.42%
5.36363956	1	0.08%	0.34%
5.41558765	1	0.08%	0.25%
5.46753574	0	0.00%	0.25%
5.51948383	0	0.00%	0.25%
5.57143192	0	0.00%	0.25%
5.62338001	0	0.00%	0.25%
5.6753281	0	0.00%	0.25%
5.72727619	1	0.08%	0.17%
5.77922428	0	0.00%	0.17%
5.83117237	0	0.00%	0.17%
5.88312046	0	0.00%	0.17%
5.93506855	0	0.00%	0.17%
5.98701664	0	0.00%	0.17%
6.03896473	0	0.00%	0.17%
6.09091282	0	0.00%	0.17%
6.14286091	1	0.08%	0.08%
6.194809	1	0.08%	

Table A.5. Distribution of Max Below Max in the Training period under Threshold 0.0016

Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	16168	2.87%	100.00%
0.02	14776	2.62%	97.13%
0.03	14521	2.58%	94.51%
0.04	15027	2.67%	91.93%
0.05	140909	25.01%	89.26%
0.06	36137	6.41%	64.26%
0.07	21582	3.83%	57.85%
0.08	16566	2.94%	54.02%
0.09	36879	6.54%	51.08%
0.1	61216	10.86%	44.53%
0.11	17758	3.15%	33.67%
0.12	8824	1.57%	30.52%
0.13	10646	1.89%	28.95%
0.14	23065	4.09%	27.06%
0.15	19592	3.48%	22.97%
0.16	7719	1.37%	19.49%
0.17	5019	0.89%	18.12%
0.18	7669	1.36%	17.23%
0.19	9359	1.66%	15.87%
0.2	6557	1.16%	14.21%
0.21	4215	0.75%	13.05%
0.22	4204	0.75%	12.30%
0.23	4783	0.85%	11.55%
0.24	4552	0.81%	10.70%
0.25	3167	0.56%	9.89%
0.26	3003	0.53%	9.33%
0.27	2886	0.51%	8.80%
0.28	2956	0.52%	8.29%
0.29	2527	0.45%	7.76%
0.3	2285	0.41%	7.31%
0.31	2407	0.43%	6.91%
0.32	1887	0.33%	6.48%
0.33	1956	0.35%	6.15%
0.34	1756	0.31%	5.80%
0.35	1585	0.28%	5.49%
0.36	1555	0.28%	5.21%

0.37	1548	0.27%	4.93%
0.38	1393	0.25%	4.66%
0.39	1208	0.21%	4.41%
0.4	1262	0.22%	4.19%
0.41	1100	0.20%	3.97%
0.42	1186	0.21%	3.78%
0.43	1004	0.18%	3.57%
0.44	1036	0.18%	3.39%
0.45	863	0.15%	3.20%
0.46	908	0.16%	3.05%
0.47	807	0.14%	2.89%
0.48	772	0.14%	2.75%
0.49	706	0.13%	2.61%
0.5	634	0.11%	2.48%
0.51	730	0.13%	2.37%
0.52	629	0.11%	2.24%
0.53	695	0.12%	2.13%
0.54	448	0.08%	2.01%
0.55	559	0.10%	1.93%
0.56	495	0.09%	1.83%
0.57	542	0.10%	1.74%
0.58	429	0.08%	1.64%
0.59	409	0.07%	1.57%
0.6	458	0.08%	1.49%
0.61	448	0.08%	1.41%
0.62	430	0.08%	1.33%
0.63	329	0.06%	1.26%
0.64	352	0.06%	1.20%
0.65	389	0.07%	1.14%
0.66	356	0.06%	1.07%
0.67	339	0.06%	1.00%
0.68	273	0.05%	0.95%
0.69	262	0.05%	0.90%
0.7	261	0.05%	0.85%
0.71	248	0.04%	0.80%
0.72	260	0.05%	0.76%
0.73	233	0.04%	0.71%
0.74	274	0.05%	0.67%
0.75	224	0.04%	0.62%
0.76	209	0.04%	0.58%
0.77	184	0.03%	0.55%
0.78	171	0.03%	0.51%

0.79	209	0.04%	0.48%
0.8	196	0.03%	0.45%
0.81	153	0.03%	0.41%
0.82	171	0.03%	0.38%
0.83	174	0.03%	0.35%
0.84	160	0.03%	0.32%
0.85	125	0.02%	0.29%
0.86	125	0.02%	0.27%
0.87	162	0.03%	0.25%
0.88	132	0.02%	0.22%
0.89	126	0.02%	0.20%
0.9	74	0.01%	0.18%
0.91	120	0.02%	0.16%
0.92	139	0.02%	0.14%
0.93	89	0.02%	0.12%
0.94	74	0.01%	0.10%
0.95	86	0.02%	0.09%
0.96	82	0.01%	0.07%
0.97	105	0.02%	0.06%
0.98	78	0.01%	0.04%
0.99	71	0.01%	0.03%
1	72	0.01%	0.01%

Table A.6. Distribution of Max Below Max in Training period under Threshold 0.0032			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	12461	5.32%	100.00%
0.02	11570	4.94%	94.68%
0.03	68389	29.17%	89.75%
0.04	13744	5.86%	60.58%
0.05	38282	16.33%	54.72%
0.06	10034	4.28%	38.39%
0.07	13846	5.91%	34.11%
0.08	11238	4.79%	28.20%
0.09	5455	2.33%	23.41%
0.1	6983	2.98%	21.08%
0.11	3760	1.60%	18.10%
0.12	4348	1.85%	16.50%
0.13	2937	1.25%	14.65%
0.14	2798	1.19%	13.39%
0.15	2390	1.02%	12.20%
0.16	2187	0.93%	11.18%
0.17	1851	0.79%	10.25%
0.18	1691	0.72%	9.46%
0.19	1553	0.66%	8.74%
0.2	1391	0.59%	8.07%
0.21	1267	0.54%	7.48%
0.22	1164	0.50%	6.94%
0.23	1010	0.43%	6.44%
0.24	960	0.41%	6.01%
0.25	809	0.35%	5.60%
0.26	814	0.35%	5.26%
0.27	709	0.30%	4.91%
0.28	650	0.28%	4.61%
0.29	636	0.27%	4.33%
0.3	611	0.26%	4.06%
0.31	592	0.25%	3.80%
0.32	465	0.20%	3.55%
0.33	512	0.22%	3.35%
0.34	442	0.19%	3.13%
0.35	423	0.18%	2.94%
0.36	373	0.16%	2.76%

0.37	351	0.15%	2.60%
0.38	353	0.15%	2.45%
0.39	245	0.10%	2.30%
0.4	308	0.13%	2.20%
0.41	290	0.12%	2.07%
0.42	282	0.12%	1.94%
0.43	206	0.09%	1.82%
0.44	254	0.11%	1.73%
0.45	164	0.07%	1.63%
0.46	203	0.09%	1.56%
0.47	176	0.08%	1.47%
0.48	169	0.07%	1.39%
0.49	196	0.08%	1.32%
0.5	142	0.06%	1.24%
0.51	132	0.06%	1.18%
0.52	125	0.05%	1.12%
0.53	112	0.05%	1.07%
0.54	136	0.06%	1.02%
0.55	111	0.05%	0.96%
0.56	92	0.04%	0.91%
0.57	117	0.05%	0.88%
0.58	91	0.04%	0.83%
0.59	89	0.04%	0.79%
0.6	87	0.04%	0.75%
0.61	84	0.04%	0.71%
0.62	84	0.04%	0.68%
0.63	78	0.03%	0.64%
0.64	70	0.03%	0.61%
0.65	82	0.03%	0.58%
0.66	59	0.03%	0.54%
0.67	62	0.03%	0.52%
0.68	53	0.02%	0.49%
0.69	60	0.03%	0.47%
0.7	48	0.02%	0.44%
0.71	48	0.02%	0.42%
0.72	50	0.02%	0.40%
0.73	49	0.02%	0.38%
0.74	53	0.02%	0.36%
0.75	50	0.02%	0.34%
0.76	45	0.02%	0.32%
0.77	55	0.02%	0.30%
0.78	38	0.02%	0.27%

0.79	30	0.01%	0.26%
0.8	37	0.02%	0.24%
0.81	28	0.01%	0.23%
0.82	31	0.01%	0.22%
0.83	38	0.02%	0.20%
0.84	35	0.01%	0.19%
0.85	45	0.02%	0.17%
0.86	29	0.01%	0.15%
0.87	42	0.02%	0.14%
0.88	30	0.01%	0.12%
0.89	25	0.01%	0.11%
0.9	22	0.01%	0.10%
0.91	19	0.01%	0.09%
0.92	23	0.01%	0.08%
0.93	24	0.01%	0.07%
0.94	17	0.01%	0.06%
0.95	22	0.01%	0.05%
0.96	17	0.01%	0.04%
0.97	28	0.01%	0.04%
0.98	22	0.01%	0.03%
0.99	13	0.01%	0.02%
1	24	0.01%	0.01%

Table A.7. Distribution of Max Below Max in Nowcasting period under Threshold 0.0016

Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	16251	7.38%	100.00%
0.02	17496	7.94%	92.62%
0.03	13634	6.19%	84.68%
0.04	10815	4.91%	78.49%
0.05	17552	7.97%	73.58%
0.06	27085	12.29%	65.62%
0.07	14920	6.77%	53.32%
0.08	12031	5.46%	46.55%
0.09	9580	4.35%	41.09%
0.1	10507	4.77%	36.74%
0.11	7801	3.54%	31.97%
0.12	8003	3.63%	28.43%
0.13	4888	2.22%	24.80%
0.14	5038	2.29%	22.58%
0.15	3940	1.79%	20.29%
0.16	3278	1.49%	18.50%
0.17	3198	1.45%	17.01%
0.18	2644	1.20%	15.56%
0.19	2427	1.10%	14.36%
0.2	1999	0.91%	13.26%
0.21	1794	0.81%	12.35%
0.22	1569	0.71%	11.54%
0.23	1780	0.81%	10.83%
0.24	1465	0.67%	10.02%
0.25	1271	0.58%	9.35%
0.26	1164	0.53%	8.78%
0.27	935	0.42%	8.25%
0.28	1136	0.52%	7.82%
0.29	981	0.45%	7.31%
0.3	879	0.40%	6.86%
0.31	813	0.37%	6.46%
0.32	650	0.30%	6.09%
0.33	642	0.29%	5.80%
0.34	680	0.31%	5.51%
0.35	660	0.30%	5.20%
0.36	617	0.28%	4.90%

0.37	520	0.24%	4.62%
0.38	470	0.21%	4.38%
0.39	456	0.21%	4.17%
0.4	476	0.22%	3.96%
0.41	436	0.20%	3.75%
0.42	416	0.19%	3.55%
0.43	307	0.14%	3.36%
0.44	326	0.15%	3.22%
0.45	367	0.17%	3.07%
0.46	319	0.14%	2.91%
0.47	336	0.15%	2.76%
0.48	306	0.14%	2.61%
0.49	247	0.11%	2.47%
0.5	228	0.10%	2.36%
0.51	283	0.13%	2.25%
0.52	262	0.12%	2.13%
0.53	204	0.09%	2.01%
0.54	205	0.09%	1.91%
0.55	180	0.08%	1.82%
0.56	223	0.10%	1.74%
0.57	192	0.09%	1.64%
0.58	188	0.09%	1.55%
0.59	170	0.08%	1.47%
0.6	191	0.09%	1.39%
0.61	158	0.07%	1.30%
0.62	168	0.08%	1.23%
0.63	124	0.06%	1.15%
0.64	141	0.06%	1.10%
0.65	135	0.06%	1.03%
0.66	116	0.05%	0.97%
0.67	121	0.05%	0.92%
0.68	107	0.05%	0.87%
0.69	105	0.05%	0.82%
0.7	112	0.05%	0.77%
0.71	106	0.05%	0.72%
0.72	82	0.04%	0.67%
0.73	82	0.04%	0.63%
0.74	79	0.04%	0.60%
0.75	79	0.04%	0.56%
0.76	59	0.03%	0.52%
0.77	94	0.04%	0.50%
0.78	61	0.03%	0.45%

0.79	53	0.02%	0.43%
0.8	33	0.01%	0.40%
0.81	56	0.03%	0.39%
0.82	50	0.02%	0.36%
0.83	43	0.02%	0.34%
0.84	52	0.02%	0.32%
0.85	49	0.02%	0.30%
0.86	58	0.03%	0.27%
0.87	57	0.03%	0.25%
0.88	47	0.02%	0.22%
0.89	39	0.02%	0.20%
0.9	45	0.02%	0.18%
0.91	44	0.02%	0.16%
0.92	38	0.02%	0.14%
0.93	36	0.02%	0.13%
0.94	46	0.02%	0.11%
0.95	31	0.01%	0.09%
0.96	40	0.02%	0.07%
0.97	35	0.02%	0.06%
0.98	24	0.01%	0.04%
0.99	36	0.02%	0.03%
1	28	0.01%	0.01%

Table A.8. Distribution of Max Below Max in Nowcasting period under Threshold 0.0032			
Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0.01	13260	14.52%	100.00%
0.02	9658	10.57%	85.48%
0.03	17676	19.35%	74.91%
0.04	10491	11.48%	55.56%
0.05	7635	8.36%	44.08%
0.06	6369	6.97%	35.72%
0.07	3885	4.25%	28.75%
0.08	2969	3.25%	24.49%
0.09	2435	2.67%	21.24%
0.1	1845	2.02%	18.58%
0.11	1485	1.63%	16.56%
0.12	1433	1.57%	14.93%
0.13	1083	1.19%	13.36%
0.14	1025	1.12%	12.18%
0.15	884	0.97%	11.05%
0.16	757	0.83%	10.09%
0.17	605	0.66%	9.26%
0.18	658	0.72%	8.60%
0.19	491	0.54%	7.88%
0.2	499	0.55%	7.34%
0.21	480	0.53%	6.79%
0.22	365	0.40%	6.27%
0.23	367	0.40%	5.87%
0.24	381	0.42%	5.46%
0.25	294	0.32%	5.05%
0.26	302	0.33%	4.73%
0.27	232	0.25%	4.40%
0.28	240	0.26%	4.14%
0.29	236	0.26%	3.88%
0.3	218	0.24%	3.62%
0.31	192	0.21%	3.38%
0.32	186	0.20%	3.17%
0.33	148	0.16%	2.97%
0.34	147	0.16%	2.81%
0.35	131	0.14%	2.64%
0.36	130	0.14%	2.50%

0.37	110	0.12%	2.36%
0.38	100	0.11%	2.24%
0.39	107	0.12%	2.13%
0.4	78	0.09%	2.01%
0.41	80	0.09%	1.93%
0.42	71	0.08%	1.84%
0.43	93	0.10%	1.76%
0.44	81	0.09%	1.66%
0.45	84	0.09%	1.57%
0.46	66	0.07%	1.48%
0.47	73	0.08%	1.41%
0.48	70	0.08%	1.33%
0.49	57	0.06%	1.25%
0.5	68	0.07%	1.19%
0.51	65	0.07%	1.11%
0.52	45	0.05%	1.04%
0.53	41	0.04%	0.99%
0.54	38	0.04%	0.95%
0.55	42	0.05%	0.91%
0.56	36	0.04%	0.86%
0.57	43	0.05%	0.82%
0.58	33	0.04%	0.77%
0.59	22	0.02%	0.74%
0.6	38	0.04%	0.71%
0.61	31	0.03%	0.67%
0.62	28	0.03%	0.64%
0.63	22	0.02%	0.61%
0.64	26	0.03%	0.58%
0.65	31	0.03%	0.56%
0.66	19	0.02%	0.52%
0.67	29	0.03%	0.50%
0.68	26	0.03%	0.47%
0.69	28	0.03%	0.44%
0.7	28	0.03%	0.41%
0.71	23	0.03%	0.38%
0.72	16	0.02%	0.35%
0.73	16	0.02%	0.34%
0.74	16	0.02%	0.32%
0.75	21	0.02%	0.30%
0.76	23	0.03%	0.28%
0.77	14	0.02%	0.25%
0.78	13	0.01%	0.24%

0.79	15	0.02%	0.22%
0.8	12	0.01%	0.21%
0.81	18	0.02%	0.19%
0.82	9	0.01%	0.17%
0.83	18	0.02%	0.16%
0.84	8	0.01%	0.14%
0.85	11	0.01%	0.14%
0.86	6	0.01%	0.12%
0.87	10	0.01%	0.12%
0.88	8	0.01%	0.11%
0.89	10	0.01%	0.10%
0.9	7	0.01%	0.09%
0.91	14	0.02%	0.08%
0.92	8	0.01%	0.06%
0.93	9	0.01%	0.05%
0.94	6	0.01%	0.04%
0.95	7	0.01%	0.04%
0.96	4	0.00%	0.03%
0.97	6	0.01%	0.03%
0.98	8	0.01%	0.02%
0.99	5	0.01%	0.01%
1	5	0.01%	0.01%

Table A.9. Distribution of TMVEXT in the Traning period$_T$ under threshold 0.00002			
Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1	0		
3.08433502	44311	50.16%	100.00%
5.16867004	24804	28.08%	49.84%
7.25300506	8336	9.44%	21.76%
9.33734008	4517	5.11%	12.32%
11.4216751	2656	3.01%	7.21%
13.50601012	1552	1.76%	4.20%
15.59034514	787	0.89%	2.44%
17.67468016	464	0.53%	1.55%
19.75901518	330	0.37%	1.03%
21.8433502	186	0.21%	0.65%
23.92768522	110	0.12%	0.44%
26.01202024	75	0.08%	0.32%
28.09635526	55	0.06%	0.23%
30.18069028	44	0.05%	0.17%
32.2650253	32	0.04%	0.12%
34.34936032	17	0.02%	0.08%
36.43369534	6	0.01%	0.06%
38.51803036	7	0.01%	0.06%
40.60236538	9	0.01%	0.05%
42.6867004	5	0.01%	0.04%
44.77103542	4	0.00%	0.03%
46.85537044	7	0.01%	0.03%
48.93970546	1	0.00%	0.02%
51.02404048	1	0.00%	0.02%
53.1083755	1	0.00%	0.02%
55.19271052	1	0.00%	0.02%
57.27704554	2	0.00%	0.02%
59.36138056	1	0.00%	0.01%
61.44571558	0	0.00%	0.01%
63.5300506	0	0.00%	0.01%
65.61438562	2	0.00%	0.01%
67.69872064	2	0.00%	0.01%
69.78305566	0	0.00%	0.01%
71.86739068	0	0.00%	0.01%
73.9517257	1	0.00%	0.01%
76.03606072	0	0.00%	0.01%

78.12039574	0	0.00%	0.01%
80.20473076	1	0.00%	0.01%
82.28906578	0	0.00%	0.01%
84.3734008	0	0.00%	0.01%
86.45773582	0	0.00%	0.01%
88.54207084	1	0.00%	0.01%
90.62640586	0	0.00%	0.01%
92.71074088	0	0.00%	0.01%
94.7950759	0	0.00%	0.01%
96.87941092	1	0.00%	0.01%
98.96374594	0	0.00%	0.00%
101.048081	0	0.00%	0.00%
103.132416	1	0.00%	0.00%
105.216751	0	0.00%	0.00%
107.301086	0	0.00%	0.00%
109.385421	0	0.00%	0.00%
111.4697561	0	0.00%	0.00%
113.5540911	0	0.00%	0.00%
115.6384261	0	0.00%	0.00%
117.7227611	1	0.00%	0.00%
119.8070961	0	0.00%	0.00%
121.8914312	0	0.00%	0.00%
123.9757662	0	0.00%	0.00%
126.0601012	0	0.00%	0.00%
128.1444362	0	0.00%	0.00%
130.2287712	0	0.00%	0.00%
132.3131063	0	0.00%	0.00%
134.3974413	0	0.00%	0.00%
136.4817763	0	0.00%	0.00%
138.5661113	0	0.00%	0.00%
140.6504463	0	0.00%	0.00%
142.7347814	0	0.00%	0.00%
144.8191164	0	0.00%	0.00%
146.9034514	0	0.00%	0.00%
148.9877864	1	0.00%	0.00%
151.0721214	0	0.00%	0.00%
153.1564565	0	0.00%	0.00%
155.2407915	0	0.00%	0.00%
157.3251265	0	0.00%	0.00%
159.4094615	0	0.00%	0.00%
161.4937965	0	0.00%	0.00%
163.5781316	0	0.00%	0.00%
165.6624666	0	0.00%	0.00%

167.7468016	0	0.00%	0.00%
169.8311366	0	0.00%	0.00%
171.9154716	0	0.00%	0.00%
173.9998067	0	0.00%	0.00%
176.0841417	0	0.00%	0.00%
178.1684767	0	0.00%	0.00%
180.2528117	0	0.00%	0.00%
182.3371467	0	0.00%	0.00%
184.4214818	0	0.00%	0.00%
186.5058168	0	0.00%	0.00%
188.5901518	0	0.00%	0.00%
190.6744868	0	0.00%	0.00%
192.7588218	0	0.00%	0.00%
194.8431569	0	0.00%	0.00%
196.9274919	0	0.00%	0.00%
199.0118269	0	0.00%	0.00%
201.0961619	0	0.00%	0.00%
203.1804969	0	0.00%	0.00%
205.264832	0	0.00%	0.00%
207.349167	0	0.00%	0.00%
209.433502	1	0.00%	0.00%

Table A.10. Distribution of Max Below Max in Traning period$_T$ under Threshold 0.00002

Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0	0	0	
0.01	0	0.00%	100.00%
0.02	0	0.00%	100.00%
0.03	0	0.00%	100.00%
0.04	0	0.00%	100.00%
0.05	0	0.00%	100.00%
0.06	0	0.00%	100.00%
0.07	0	0.00%	100.00%
0.08	0	0.00%	100.00%
0.09	0	0.00%	100.00%
0.1	0	0.00%	100.00%
0.11	0	0.00%	100.00%
0.12	0	0.00%	100.00%
0.13	0	0.00%	100.00%
0.14	0	0.00%	100.00%
0.15	0	0.00%	100.00%
0.16	0	0.00%	100.00%
0.17	0	0.00%	100.00%
0.18	0	0.00%	100.00%
0.19	0	0.00%	100.00%
0.2	0	0.00%	100.00%
0.21	0	0.00%	100.00%
0.22	0	0.00%	100.00%
0.23	0	0.00%	100.00%
0.24	0	0.00%	100.00%
0.25	0	0.00%	100.00%
0.26	0	0.00%	100.00%
0.27	0	0.00%	100.00%
0.28	0	0.00%	100.00%
0.29	0	0.00%	100.00%
0.3	0	0.00%	100.00%
0.31	0	0.00%	100.00%
0.32	0	0.00%	100.00%
0.33	0	0.00%	100.00%
0.34	0	0.00%	100.00%
0.35	0	0.00%	100.00%

0.36	0	0.00%	100.00%
0.37	0	0.00%	100.00%
0.38	0	0.00%	100.00%
0.39	0	0.00%	100.00%
0.4	0	0.00%	100.00%
0.41	0	0.00%	100.00%
0.42	0	0.00%	100.00%
0.43	0	0.00%	100.00%
0.44	0	0.00%	100.00%
0.45	0	0.00%	100.00%
0.46	0	0.00%	100.00%
0.47	0	0.00%	100.00%
0.48	0	0.00%	100.00%
0.49	0	0.00%	100.00%
0.5	0	0.00%	100.00%
0.51	0	0.00%	100.00%
0.52	0	0.00%	100.00%
0.53	0	0.00%	100.00%
0.54	0	0.00%	100.00%
0.55	0	0.00%	100.00%
0.56	0	0.00%	100.00%
0.57	0	0.00%	100.00%
0.58	0	0.00%	100.00%
0.59	0	0.00%	100.00%
0.6	0	0.00%	100.00%
0.61	0	0.00%	100.00%
0.62	0	0.00%	100.00%
0.63	183	0.58%	100.00%
0.64	18133	57.43%	99.42%
0.65	13256	41.99%	41.99%
0.66	0	0.00%	0.00%
0.67	0	0.00%	0.00%
0.68	0	0.00%	0.00%
0.69	0	0.00%	0.00%
0.7	0	0.00%	0.00%
0.71	0	0.00%	0.00%
0.72	0	0.00%	0.00%
0.73	0	0.00%	0.00%
0.74	0	0.00%	0.00%
0.75	0	0.00%	0.00%
0.76	0	0.00%	0.00%
0.77	0	0.00%	0.00%

0.78	0	0.00%	0.00%
0.79	0	0.00%	0.00%
0.8	0	0.00%	0.00%
0.81	0	0.00%	0.00%
0.82	0	0.00%	0.00%
0.83	0	0.00%	0.00%
0.84	0	0.00%	0.00%
0.85	0	0.00%	0.00%
0.86	0	0.00%	0.00%
0.87	0	0.00%	0.00%
0.88	0	0.00%	0.00%
0.89	0	0.00%	0.00%
0.9	0	0.00%	0.00%
0.91	0	0.00%	0.00%
0.92	0	0.00%	0.00%
0.93	0	0.00%	0.00%
0.94	0	0.00%	0.00%
0.95	0	0.00%	0.00%
0.96	0	0.00%	0.00%
0.97	0	0.00%	0.00%
0.98	0	0.00%	0.00%
0.99	0	0.00%	0.00%
1	0	0.00%	0.00%

Table A.11. Distribution of Max Below Max in Traning period USD/JPY under Threshold 0.00016

Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0	0	0	
0.01	18239	9.66%	100.00%
0.02	28020	14.84%	90.34%
0.03	26549	14.06%	75.50%
0.04	26029	13.79%	61.44%
0.05	19077	10.10%	47.65%
0.06	13493	7.15%	37.55%
0.07	9464	5.01%	30.40%
0.08	7027	3.72%	25.39%
0.09	5318	2.82%	21.67%
0.1	4179	2.21%	18.85%

0.11	3372	1.79%	16.64%
0.12	2949	1.56%	14.85%
0.13	2126	1.13%	13.29%
0.14	2149	1.14%	12.17%
0.15	1653	0.88%	11.03%
0.16	1688	0.89%	10.15%
0.17	1229	0.65%	9.26%
0.18	1332	0.71%	8.61%
0.19	990	0.52%	7.90%
0.2	1092	0.58%	7.38%
0.21	847	0.45%	6.80%
0.22	867	0.46%	6.35%
0.23	773	0.41%	5.89%
0.24	743	0.39%	5.48%
0.25	632	0.33%	5.09%
0.26	608	0.32%	4.75%
0.27	531	0.28%	4.43%
0.28	536	0.28%	4.15%
0.29	441	0.23%	3.87%
0.3	432	0.23%	3.63%
0.31	381	0.20%	3.40%
0.32	345	0.18%	3.20%
0.33	332	0.18%	3.02%
0.34	290	0.15%	2.84%
0.35	263	0.14%	2.69%
0.36	243	0.13%	2.55%
0.37	255	0.14%	2.42%
0.38	219	0.12%	2.29%
0.39	204	0.11%	2.17%
0.4	203	0.11%	2.06%
0.41	217	0.11%	1.96%
0.42	209	0.11%	1.84%
0.43	192	0.10%	1.73%
0.44	148	0.08%	1.63%
0.45	156	0.08%	1.55%
0.46	154	0.08%	1.47%
0.47	146	0.08%	1.39%
0.48	124	0.07%	1.31%
0.49	146	0.08%	1.24%
0.5	97	0.05%	1.17%
0.51	97	0.05%	1.11%
0.52	115	0.06%	1.06%

0.53	95	0.05%	1.00%
0.54	100	0.05%	0.95%
0.55	96	0.05%	0.90%
0.56	100	0.05%	0.85%
0.57	78	0.04%	0.79%
0.58	76	0.04%	0.75%
0.59	78	0.04%	0.71%
0.6	54	0.03%	0.67%
0.61	78	0.04%	0.64%
0.62	54	0.03%	0.60%
0.63	57	0.03%	0.57%
0.64	52	0.03%	0.54%
0.65	62	0.03%	0.52%
0.66	35	0.02%	0.48%
0.67	56	0.03%	0.46%
0.68	42	0.02%	0.43%
0.69	64	0.03%	0.41%
0.7	39	0.02%	0.38%
0.71	36	0.02%	0.36%
0.72	42	0.02%	0.34%
0.73	41	0.02%	0.32%
0.74	34	0.02%	0.29%
0.75	28	0.01%	0.28%
0.76	24	0.01%	0.26%
0.77	35	0.02%	0.25%
0.78	16	0.01%	0.23%
0.79	34	0.02%	0.22%
0.8	25	0.01%	0.20%
0.81	27	0.01%	0.19%
0.82	34	0.02%	0.18%
0.83	27	0.01%	0.16%
0.84	22	0.01%	0.14%
0.85	25	0.01%	0.13%
0.86	23	0.01%	0.12%
0.87	19	0.01%	0.11%
0.88	15	0.01%	0.10%
0.89	10	0.01%	0.09%
0.9	16	0.01%	0.08%
0.91	19	0.01%	0.08%
0.92	16	0.01%	0.07%
0.93	16	0.01%	0.06%
0.94	11	0.01%	0.05%

0.95	10	0.01%	0.04%
0.96	14	0.01%	0.04%
0.97	11	0.01%	0.03%
0.98	17	0.01%	0.02%
0.99	18	0.01%	0.01%
1	10	0.01%	0.01%

Table A.12. Distribution of TMV_{EXT} in the Traning period USD/JPY under threshold 0.00016

Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1	0		
1.08110049	271	8.08%	100.00%
1.16220098	251	7.48%	91.92%
1.24330147	191	5.69%	84.44%
1.32440196	201	5.99%	78.74%
1.40550245	193	5.75%	72.75%
1.48660294	178	5.31%	66.99%
1.56770343	169	5.04%	61.69%
1.64880392	135	4.03%	56.65%
1.72990441	129	3.85%	52.62%
1.8110049	132	3.94%	48.78%
1.89210539	119	3.55%	44.84%
1.97320588	109	3.25%	41.29%
2.05430637	107	3.19%	38.04%
2.13540686	101	3.01%	34.85%
2.21650735	74	2.21%	31.84%
2.29760784	70	2.09%	29.64%
2.37870833	62	1.85%	27.55%
2.45980882	79	2.36%	25.70%
2.54090931	72	2.15%	23.35%
2.6220098	46	1.37%	21.20%
2.70311029	55	1.64%	19.83%
2.78421078	47	1.40%	18.19%
2.86531127	44	1.31%	16.79%
2.94641176	51	1.52%	15.47%
3.02751225	35	1.04%	13.95%
3.10861274	42	1.25%	12.91%
3.18971323	29	0.86%	11.66%
3.27081372	30	0.89%	10.79%

3.35191421	27	0.81%	9.90%
3.4330147	23	0.69%	9.09%
3.51411519	23	0.69%	8.41%
3.59521568	28	0.83%	7.72%
3.67631617	21	0.63%	6.89%
3.75741666	20	0.60%	6.26%
3.83851715	21	0.63%	5.66%
3.91961764	18	0.54%	5.04%
4.00071813	14	0.42%	4.50%
4.08181862	16	0.48%	4.08%
4.16291911	9	0.27%	3.61%
4.2440196	15	0.45%	3.34%
4.32512009	9	0.27%	2.89%
4.40622058	7	0.21%	2.62%
4.48732107	7	0.21%	2.42%
4.56842156	4	0.12%	2.21%
4.64952205	3	0.09%	2.09%
4.73062254	7	0.21%	2.00%
4.81172303	6	0.18%	1.79%
4.89282352	4	0.12%	1.61%
4.97392401	3	0.09%	1.49%
5.0550245	2	0.06%	1.40%
5.13612499	1	0.03%	1.34%
5.21722548	5	0.15%	1.31%
5.29832597	5	0.15%	1.16%
5.37942646	2	0.06%	1.01%
5.46052695	1	0.03%	0.95%
5.54162744	4	0.12%	0.92%
5.62272793	3	0.09%	0.81%
5.70382842	2	0.06%	0.72%
5.78492891	3	0.09%	0.66%
5.8660294	1	0.03%	0.57%
5.94712989	2	0.06%	0.54%
6.02823038	3	0.09%	0.48%
6.10933087	1	0.03%	0.39%
6.19043136	0	0.00%	0.36%
6.27153185	0	0.00%	0.36%
6.35263234	0	0.00%	0.36%
6.43373283	2	0.06%	0.36%
6.51483332	0	0.00%	0.30%
6.59593381	1	0.03%	0.30%
6.6770343	2	0.06%	0.27%
6.75813479	0	0.00%	0.21%

6.83923528	0	0.00%	0.21%
6.92033577	0	0.00%	0.21%
7.00143626	0	0.00%	0.21%
7.08253675	0	0.00%	0.21%
7.16363724	2	0.06%	0.21%
7.24473773	0	0.00%	0.15%
7.32583822	0	0.00%	0.15%
7.40693871	0	0.00%	0.15%
7.4880392	1	0.03%	0.15%
7.56913969	1	0.03%	0.12%
7.65024018	0	0.00%	0.09%
7.73134067	0	0.00%	0.09%
7.81244116	0	0.00%	0.09%
7.89354165	0	0.00%	0.09%
7.97464214	1	0.03%	0.09%
8.05574263	0	0.00%	0.06%
8.13684312	0	0.00%	0.06%
8.21794361	1	0.03%	0.06%
8.2990441	0	0.00%	0.03%
8.38014459	0	0.00%	0.03%
8.46124508	0	0.00%	0.03%
8.54234557	0	0.00%	0.03%
8.62344606	0	0.00%	0.03%
8.70454655	0	0.00%	0.03%
8.78564704	0	0.00%	0.03%
8.86674753	0	0.00%	0.03%
8.94784802	0	0.00%	0.03%
9.02894851	0	0.00%	0.03%
9.110049	1	0.03%	0.03%

Table A.13. Distribution of Max Below Max in Traning period GBP/USD under Threshold 0.0032

Bin	MBM Frequency	MBM Probabilities	Probabilities of reaching this MBM value
0	0	0	
0.01	18239	9.66%	100.00%
0.02	28020	14.84%	90.34%
0.03	26549	14.06%	75.50%
0.04	26029	13.79%	61.44%
0.05	19077	10.10%	47.65%

0.06	13493	7.15%	37.55%
0.07	9464	5.01%	30.40%
0.08	7027	3.72%	25.39%
0.09	5318	2.82%	21.67%
0.1	4179	2.21%	18.85%
0.11	3372	1.79%	16.64%
0.12	2949	1.56%	14.85%
0.13	2126	1.13%	13.29%
0.14	2149	1.14%	12.17%
0.15	1653	0.88%	11.03%
0.16	1688	0.89%	10.15%
0.17	1229	0.65%	9.26%
0.18	1332	0.71%	8.61%
0.19	990	0.52%	7.90%
0.2	1092	0.58%	7.38%
0.21	847	0.45%	6.80%
0.22	867	0.46%	6.35%
0.23	773	0.41%	5.89%
0.24	743	0.39%	5.48%
0.25	632	0.33%	5.09%
0.26	608	0.32%	4.75%
0.27	531	0.28%	4.43%
0.28	536	0.28%	4.15%
0.29	441	0.23%	3.87%
0.3	432	0.23%	3.63%
0.31	381	0.20%	3.40%
0.32	345	0.18%	3.20%
0.33	332	0.18%	3.02%
0.34	290	0.15%	2.84%
0.35	263	0.14%	2.69%
0.36	243	0.13%	2.55%
0.37	255	0.14%	2.42%
0.38	219	0.12%	2.29%
0.39	204	0.11%	2.17%
0.4	203	0.11%	2.06%
0.41	217	0.11%	1.96%
0.42	209	0.11%	1.84%
0.43	192	0.10%	1.73%
0.44	148	0.08%	1.63%
0.45	156	0.08%	1.55%
0.46	154	0.08%	1.47%
0.47	146	0.08%	1.39%

0.48	124	0.07%	1.31%
0.49	146	0.08%	1.24%
0.5	97	0.05%	1.17%
0.51	97	0.05%	1.11%
0.52	115	0.06%	1.06%
0.53	95	0.05%	1.00%
0.54	100	0.05%	0.95%
0.55	96	0.05%	0.90%
0.56	100	0.05%	0.85%
0.57	78	0.04%	0.79%
0.58	76	0.04%	0.75%
0.59	78	0.04%	0.71%
0.6	54	0.03%	0.67%
0.61	78	0.04%	0.64%
0.62	54	0.03%	0.60%
0.63	57	0.03%	0.57%
0.64	52	0.03%	0.54%
0.65	62	0.03%	0.52%
0.66	35	0.02%	0.48%
0.67	56	0.03%	0.46%
0.68	42	0.02%	0.43%
0.69	64	0.03%	0.41%
0.7	39	0.02%	0.38%
0.71	36	0.02%	0.36%
0.72	42	0.02%	0.34%
0.73	41	0.02%	0.32%
0.74	34	0.02%	0.29%
0.75	28	0.01%	0.28%
0.76	24	0.01%	0.26%
0.77	35	0.02%	0.25%
0.78	16	0.01%	0.23%
0.79	34	0.02%	0.22%
0.8	25	0.01%	0.20%
0.81	27	0.01%	0.19%
0.82	34	0.02%	0.18%
0.83	27	0.01%	0.16%
0.84	22	0.01%	0.14%
0.85	25	0.01%	0.13%
0.86	23	0.01%	0.12%
0.87	19	0.01%	0.11%
0.88	15	0.01%	0.10%
0.89	10	0.01%	0.09%

0.9	16	0.01%	0.08%
0.91	19	0.01%	0.08%
0.92	16	0.01%	0.07%
0.93	16	0.01%	0.06%
0.94	11	0.01%	0.05%
0.95	10	0.01%	0.04%
0.96	14	0.01%	0.04%
0.97	11	0.01%	0.03%
0.98	17	0.01%	0.02%
0.99	18	0.01%	0.01%
1	10	0.01%	0.01%

Table A.14. Distribution of TMV_{EXT} in the Traning period GBP/USD under threshold 0.0032

Bin	TMV Frequency	TMV Probabilities	Probabilities of reaching this TMV
1	0		
1.08110049	271	8.08%	100.00%
1.16220098	251	7.48%	91.92%
1.24330147	191	5.69%	84.44%
1.32440196	201	5.99%	78.74%
1.40550245	193	5.75%	72.75%
1.48660294	178	5.31%	66.99%
1.56770343	169	5.04%	61.69%
1.64880392	135	4.03%	56.65%
1.72990441	129	3.85%	52.62%
1.8110049	132	3.94%	48.78%
1.89210539	119	3.55%	44.84%
1.97320588	109	3.25%	41.29%
2.05430637	107	3.19%	38.04%
2.13540686	101	3.01%	34.85%
2.21650735	74	2.21%	31.84%
2.29760784	70	2.09%	29.64%
2.37870833	62	1.85%	27.55%
2.45980882	79	2.36%	25.70%
2.54090931	72	2.15%	23.35%
2.6220098	46	1.37%	21.20%
2.70311029	55	1.64%	19.83%
2.78421078	47	1.40%	18.19%
2.86531127	44	1.31%	16.79%

2.94641176	51	1.52%	15.47%
3.02751225	35	1.04%	13.95%
3.10861274	42	1.25%	12.91%
3.18971323	29	0.86%	11.66%
3.27081372	30	0.89%	10.79%
3.35191421	27	0.81%	9.90%
3.4330147	23	0.69%	9.09%
3.51411519	23	0.69%	8.41%
3.59521568	28	0.83%	7.72%
3.67631617	21	0.63%	6.89%
3.75741666	20	0.60%	6.26%
3.83851715	21	0.63%	5.66%
3.91961764	18	0.54%	5.04%
4.00071813	14	0.42%	4.50%
4.08181862	16	0.48%	4.08%
4.16291911	9	0.27%	3.61%
4.2440196	15	0.45%	3.34%
4.32512009	9	0.27%	2.89%
4.40622058	7	0.21%	2.62%
4.48732107	7	0.21%	2.42%
4.56842156	4	0.12%	2.21%
4.64952205	3	0.09%	2.09%
4.73062254	7	0.21%	2.00%
4.81172303	6	0.18%	1.79%
4.89282352	4	0.12%	1.61%
4.97392401	3	0.09%	1.49%
5.0550245	2	0.06%	1.40%
5.13612499	1	0.03%	1.34%
5.21722548	5	0.15%	1.31%
5.29832597	5	0.15%	1.16%
5.37942646	2	0.06%	1.01%
5.46052695	1	0.03%	0.95%
5.54162744	4	0.12%	0.92%
5.62272793	3	0.09%	0.81%
5.70382842	2	0.06%	0.72%
5.78492891	3	0.09%	0.66%
5.8660294	1	0.03%	0.57%
5.94712989	2	0.06%	0.54%
6.02823038	3	0.09%	0.48%
6.10933087	1	0.03%	0.39%
6.19043136	0	0.00%	0.36%
6.27153185	0	0.00%	0.36%
6.35263234	0	0.00%	0.36%

6.43373283	2	0.06%	0.36%
6.51483332	0	0.00%	0.30%
6.59593381	1	0.03%	0.30%
6.6770343	2	0.06%	0.27%
6.75813479	0	0.00%	0.21%
6.83923528	0	0.00%	0.21%
6.92033577	0	0.00%	0.21%
7.00143626	0	0.00%	0.21%
7.08253675	0	0.00%	0.21%
7.16363724	2	0.06%	0.21%
7.24473773	0	0.00%	0.15%
7.32583822	0	0.00%	0.15%
7.40693871	0	0.00%	0.15%
7.4880392	1	0.03%	0.15%
7.56913969	1	0.03%	0.12%
7.65024018	0	0.00%	0.09%
7.73134067	0	0.00%	0.09%
7.81244116	0	0.00%	0.09%
7.89354165	0	0.00%	0.09%
7.97464214	1	0.03%	0.09%
8.05574263	0	0.00%	0.06%
8.13684312	0	0.00%	0.06%
8.21794361	1	0.03%	0.06%
8.2990441	0	0.00%	0.03%
8.38014459	0	0.00%	0.03%
8.46124508	0	0.00%	0.03%
8.54234557	0	0.00%	0.03%
8.62344606	0	0.00%	0.03%
8.70454655	0	0.00%	0.03%
8.78564704	0	0.00%	0.03%
8.86674753	0	0.00%	0.03%
8.94784802	0	0.00%	0.03%
9.02894851	0	0.00%	0.03%
9.110049	1	0.03%	0.03%

8. Reference

[1] Abdalla Kablan and Wing Lon Ng Intraday high-frequency FX trading with adaptive neuro-fuzzy inference systems, January 2011

[2] Adegboye, A., Kampouridis, M, "Machine Learning Classification and Regression Models for Predicting Directional Changes Trend Reversal in FX Markets",(2021) Expert Systems with Applications, Vol. 173, Elsevier, pp. 114645

[3] Aloud, M. E.P.K.Tsang, R.Olsen & A. Dupuis, A Directional-Change Events Approach for Studying Financial Time Series, Economics, No. 2012-36, 7 September 2012, e-journal, http://www.economics-ejournal.org/economics/journalarticles/2012-36

[4] Ang, A. and Bekaert, G. (2004). "How regimes affect asset allocation". Financial Analysts Journal 60.2, pp. 86–99.

[5] Ao, H. & Tsang, E.P.K., Trading algorithms built with Direction Changes, IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr) 2019 Conference, Shenzhen, China, 4th-5th May 2019

[6] Bisig, T., Depuis, A., Impagliazzo, V., & Olsen, R. (2012). The scale of market quakes. Quantative Finance, 12 (4), 501-508.

[7] Bakhach, A., Tsang, E.P.K. & Jalalian, H., Forecasting Directional

Changes in FX Markets, IEEE Symposium on Computational Intelligence for Financial Engineering & Economics (IEEE CIFEr'16), Athens, Greece, 6-9 December 2016

[8] Bakhach, A., Tsang, E.P.K. & Raju Chinthalapati, V.L., TSFDC: A trading strategy based on forecasting directional change, Intelligent Systems in Accounting, Finance and Management, Vol.25, Issue 3, May 2018

[9] Bakhach, A V.L.R. Chinthalapati, E.P.K. Tsang & A.R. El Sayed, Intelligent Dynamic Backlash Agent: a trading strategy based on the directional change framework, Algorithms, Special Issue on Algorithms in Computational Finance, MDPI Open Access Publishing (ISSN 19994893), 11(11), 2018

[10] Bakhach, A., Developing trading strategies under the Directional Changes framework, with application in the FX market, PhD Thesis, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, 2018

[11] Bollerslev, T. (1986): Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327

[12] Black, F. and Scholes, M. (1973): The pricing of options and corporate liabilities. Journal of Political Economy

[13] Black, F. (1976): The pricing of commodity contracts. Journal of Financial Economics 3, 167–791

[14] Black, F. and J. C. Cox (1976). "Valuing corporate securities: Some effects of bond indenture provisions." The Journal of Finance

[15] Chen, J. & E.P.K. Tsang, Classification of Normal and Abnormal Regimes in Financial Markets, Algorithms, Special Issue on Algorithms in Computational Finance, 11(12), 202, MDPI Open Access Publishing (ISSN 19994893), published online, December 2018

[16] Chen, J. & Tsang, E.P.K., Detecting Regime Change in Computational Finance, Data Science, Machine Learning and Algorithmic Trading, CRC Press, September 2020

[17] Campbell, J., Lo, A. and McKinlay, C. (1996): The Econometrics of Financial Markets. Princeton University Press, Princeton

[18] Culp, C. L., Y. Nozawa and P. Veronesi (2014). The Empirical Merton Model. working paper, Chicago University

[19] Derman, E., Iraj Kani (1994). "Riding on a Smile." RISK, 7(2) Feb.1994" (PDF). Risk. Retrieved 1 June 2007.

[20] David Einhorn (June–July 2008), Private Profits and Socialized Risk (PDF), GARP Risk Review, archived (PDF) from the original on April 26, 2016

[21] Dupuis, A., Olsen, R. (2012). High-Frequency Finance: Using Scaling Laws to Build Trading Models. In J. James et al (Ed.),

Handbook of Exchange Rates, NJ, USA: Wiley.

[22] Delbaen, F. and Schachermayer, W. (1994): A general version of the fundamental theorem of asset pricing. Mathematische Annalen

[23] Duan, J.-C. (1997): Augmented GARCH(p,q) process and its diffusion limit. Journal of Econometric

[24] Engle, R.F. (1982): Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation

[25] Eugene F. Fama and Kenneth R. French (2004) The Capital Asset Pricing Model: Theory and Evidence Journal of Economic Perspectives—Volume 18, Number 3—Summer 2004—Pages 25–46

[26] Eric Renault (2009) Handbook of Financial Time Series "Moment–Based Estimation of Stochastic Volatility Models" Page 297-Page 311

[27] Engle, R.F. (1995): ARCH: Selected Readings. Oxford University Press

[28] Glattfelder, J., Dupuis, A., and Olsen, R. (2011). Patterns in high-frequency FX data: discovery of 12 empirical scaling laws. Quantitative Finance, 11(4): 599– 614. URL: http://www.tandfonline.com/doi/abs/10.1080/14697688.2010.48

1632

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

[30] Giannone, Reichlin, and Small (2008) Nowcasting: The real-time informational content of macroeconomic data, Journal of Monetary Economics 55 (2008) 665– 676.

[31] Guillaume, D., Dacorogna, M., Davé, R., Müller, U., Olsen, R., and Pictet, O. (1997). From the bird’s eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. Finance Stoch., 1(2): 95–129

[32] Golub, Anton and Glattfelder, James and Olsen, Richard B., The Alpha Engine: Designing an Automated Trading Algorithm, High-Performance Computing in Finance, Chapman & Hall/CRC Series in Mathematical Finance, April 2017

[33] Geske, R. (1979). "The valuation of compound options." Journal of financial economics 7(1): 63-81

[34] Geske, R. and H. E. Johnson (1984). "The American put option valued analytically." The Journal of Finance 39(5): 1511-1524

[35] Hamilton, J. D. (1989). “A new approach to the economic analysis of nonstationary time series and the business cycle”.

Econometrica: Journal of the Econometric Society, pp. 357–384.

[36] Holton, Glyn A. (2014). Value-at-Risk: Theory and Practice second edition, e-book

[37] Jorion, Philippe (2006). Value at Risk: The New Benchmark for Managing Financial Risk (3rd ed.). McGraw-Hill. ISBN 978-0-07-146495-6

[38] Jones, R.H. (1985): Time series analysis with unequally spaced data. In: Hannan E.J., Krishnaiah P.R. and Rao M.M. (Eds.): Time Series in the Time Domain, Handbook of Statistics 5, North Holland, Amsterdam

[39] Jensen, Michael C. (1968). "The Performance of Mutual Funds in the Period 1945–1964." Journal of Finance

[40] Jones, E. P., S. P. Mason and E. Rosenfeld (1984). "Contingent claims analysis of corporate capital structures: An empirical investigation." The journal of finance 39(3): 611-625

[41] Kampouridis, M., Otero, F.E.B., (2017) "Evolving Trading Strategies Using Directional Changes", Expert Systems with Applications, Elsevier, Vol. 73, pp. 145-160

[42] Lintner, John. (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." Review of Economics and Statistics

[43] Longstaff, F. A. and E. S. Schwartz (1995). "A simple approach to

valuing risky fixed and floating rate debt." The Journal of Finance 50(3): 789-819

[44] Leland, H. E. (1994). "Corporate debt value, bond covenants, and optimal capital structure." The journal of finance 49(4): 1213-1252

[45] Leland, H. E. and K. B. Toft (1996). "Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads." The Journal of Finance 51(3): 987-1019

[46] Markovich, N. (2007), Nonparametric analysis of univariate heavy-tailed data, Wiley

[47] Martinez-Jaramillo, S., Artificial financial markets: an agent-based approach to reproduce stylized facts and to study the Red Queen Effect, PhD Thesis, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, 2007

[48] Masry, S. (2013). Event-Based Microscopic Analysis of the FX Market. PhD thesis, Centre for Computational Finance and Economic Agents (CCFEQ), University of Essex.

[49] Markowitz, Harry. (1952). “Portfolio Selection.” Journal of Finance

[50] Markowitz, Harry. (1959). Portfolio Selection: Efficient Diversification of Investments. Cowles Foundation Monograph No. 16. New York: John Wiley & Sons, Inc

[51] Mandelbrot, B. (1963): The variation of certain speculative prices. Journal of Business 36 , 394–419

[52] Merton, R. (1973): The theory of rational option pricing. Bell J. Economics and Management Science

[53] Neil Shephard and Torben G. Andersen (2009), Handbook of Financial Time Series, Stochastic Volatility: Origins and Overview

[54] Nelson, D. (1990): ARCH models as diffusion approximations. J. Econometrics

[55] Nielsen, L. T., J. Saà-Requejo and P. Santa-Clara (2001). Default risk and interest rate risk: The term structure of default spreads

[56] O'Hara, Maureen., (1995) Market Microstructure Theory. Blackwell, Oxford.

[57] Pole, A., West, M., & Harrison, J., 1994, Applied Bayesian Forecasting and Time Series Analysis, London. Chapman and Hall.

[58] Piger, J. (2009). "Econometrics: Models of regime changes". Encyclopedia of Complexity and Systems Science. New York, NY: Springer, pp. 2744–2757.

[59] Rachev, S. and Mittnik, S. (2000): Stable Paretian Models in Finance. Wiley, Chichester

[60] Ruey S, Tsay (2005), Analysis of Financial Time Series, Second

Edition, by Wiley & Sons, Inc., Hoboken, New Jersey

[61] Roll. R. 1984: A Simple Implicit Measure of the Effective Bid Ask Spread in an Efficient Market, Journal of Finance 39 (4).

[62] Rosenberg, B. (1972): The behaviour of random variables with nonstationary variance and the distribution of security prices. WP11 GSBA, University of California, Berkeley

[63] Sarkissian, Jack (2016). "Express Measurement of Market Volatility Using Ergodicity Concept". SSRN 2812353

[64] Paniangtong, P., The Evaluation of the Trend-Following Directional Change with the Trailing Stop and Major-Trend-Adjusted Strategies on Algorithmic Trading in the Foreign Exchange Markets, MSc Thesis, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, September 2015

[65] Petris, G. et al. (2009). Dynamic Linear Models with R. Springer Science & Business Media

[66] Peter J. Brockwell (2009) Handbook of Financial Time Series, An Overview of Asset–Price Models ISBN 978-3-540-71296-1 Page 403-419

[67] Ronn, E. I. and A. K. Verma (1986). "Pricing risk-adjusted deposit insurance: An option-based model." The Journal of Finance 41(4): 871-895

[68] Samuelson, P.A. (1965): Rational theory of warrant pricing. Ind. Management Review 6

[69] Sharpe, William F. (1964). “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” Journal of Finance

[70] Saa-Requejo, J. and P. Santa-Clara (1997). "Bond pricing with default risk."

[71] Tao, R., Using Directional Change for Information Extraction in Financial Market Data, PhD Thesis, Centre for Computational Finance and Economic Agents (CCFEA), University of Essex, 2018

[72] Torben G. Andersen, Richard A. Davis, Jens-Peter Kreiß and Thomas Mikosch (2007) Handbook of Financial Time Series ISBN 978-3-540-71296-1

[73] Tsang, E.P.K., Tao, R., Serguieva, A. & Ma, S., Profiling High-Frequency Equity Price Movements in Directional Changes, Quantitative Finance, Vol.17, Issue 2, 2017, 217-225

[74] Tsang, E.P.K. & J. Chen, Regime change detection using directional change indicators in the foreign exchange market to chart Brexit, IEEE Transactions in Emerging Technology in Computational Intelligence (TETCI), Vol.2, Issue 3, June 2018, pages 185-193 (DOI: 10.1109/TETCI.2017.2775235 / Electronic

ISSN: 2471-285X)

[75] Tsang, E.P.K., Directional changes: a new way to look at price dynamics, In Mandal J., Dutta P., Mukhopadhyay S. (eds) Computational Intelligence, Communications, and Business Analytics (CICBA), Communications in Computer and Information Science, Vol. 775, Springer, 2017, 45-55

[76] Taylor, S. J. (1986): Modelling Financial Time Series. Wiley, Chichester.

[77] Timo Teräsvirta (2009) An Introduction to Univariate GARCH Models

[78] Ye, A., Chinthalapati, V.L.R., Serguieva, A. & Tsang, E., Developing sustainable trading strategies using directional changes with high-frequency data, IEEE International Conference on Big Data, Boston, 11-14 December 2017