Applications of Eigensystem Analysis to Derivatives and Portfolio Management

XIAO LIANG

Applications of Eigensystem Analysis to Derivatives and Portfolio Management

Xiao Liang

Durham University Business School

Durham University

A thesis submitted for

Doctor of Philosophy

APR 2021

Abstract

Eigensystem structure plays the key role in principal component analysis (PCA). However, the application of it in high-frequency datasets is noticeably thin, especially for derivatives pricing. In my thesis, I will present the predictive power of eigenvalue/eigenvector analysis in several financial markets. Performance of prediction based on eigenvalue/eigenvector structure shows the result that this methodology is reliable compared with traditional methodology.

To verify the performance of eigensystem analysis in derivatives pricing, I select one of the most important financial markets: the foreign exchange(FX) option market as datasets. The traditional pricing models for FX options are highly reliant on historical data, which leads to the dilemma that for those contracts with less liquidity investors find it difficult to provide reliable guidance on price. I will present a brand-new model based on eigensystem analysis to provide accurate guidance for option pricing, especially in cases where the underlying asset is considered to be an illiquid currency pair. The importance of eigenvalues and eigenvectors structure in asset pricing will be explored in this thesis. The empirical study covers FX option contracts across deltas and maturities. The performance of eigensystem model are compared with other widely used models, results indicate that traditional models are outperformed in all selected underlying assets, maturities and deltas.

In addition, I perform analysis of machine learning performance based on the

FX market's empirical asset pricing problem. I demonstrate the advantage of machine learning in promoting the predictive power of eigensystem based on multiple predictors from the OTC market. Black-Scholes implied volatility is used as predictors for the eigenvalue error between market and our innovative eigensystem. I identify the regression tree algorithm's predictive gain with empirical study across contracts. The effect of currency pairs is numerical and sorted to generate an overview for global FX market structure.

I also implement eigenstructure analysis based on the S&P500 market. I discover the convergence of first principal component explanatory power. In order to generate the statistical summary for trend of principal components, I raise a set of measurements and thresholds to describe eigenvalue and eigenvector structure in market portfolios.

Declaration

I, Xiao Liang, hereby declare that this is entirely my own work unless referenced to the contrary in the text. No part of this thesis has previously been submitted else where for any other degree or qualification in this or any other university.

To my parents, for always loving and supporting me.

Acknowledgements

I would like to extend my deep gratitude to all those who have offered me help and support in the process of my studies at Durham University.

First of all, my profound gratitude goes to my supervisor Professor Julian Williams, for his support during my doctoral studies. He has always given me great instruction and encouragements through the process of confirming research topics, improving my skills in all aspects such as finance, mathematics and programming. He was always positive when I faced difficulties. This thesis would not have been possible without the help from Professor Julian Williams, I would never forget the days we spent together in Durham University Business School and the Institute of Hazard Risk and Resilience.

Many thanks to my second supervisor Professor Dennis Philip for his encouragement, guidance and help. His knowledge improve my thesis significantly and expand my research. I always remember the day we met when I was given the chance to come to Durham as a Ph.D student.

I would like also to thanks Julian Cook and Fenics Software. Inc., BGC Partners. They provided funding for my research and the chance to work together with their talented engineers in New York. Julian Cook always provide valuable advise from the aspect of real market. I appreciate the patience and suggestions.

I would like to thank my colleagues Dr Jing Nie, Dr Yang Zhang, Eleftheria Vavadaki, Xiangyu Wu, Yaodong Liu and Dr XiaoXiao Ma, for their great

support and encouragement. Especially Dr Handing Sun, my best friend in Durham, his attitude towards research and incredible talent always encouraged me to become a better person.

Thanks to Johann Sebastian Bach, Marcus Tullius Cicero, Liang Shih-Chiu and Inoue Takehiko, you let me travel without moving feet.

Last but not least, I would like to thank my parents for their love and support. You are always the most important persons in my life. You show me the way of being a nice man.

Contents

List of Figures

Nomenclature

Abbreviations

ACP	Autoregressive conditional Poisson
AD	Affine diffusion model
AJD	Affine jump diffusion model
ASTSV	Affine stochastic term structure with stochastic volatility
ATM	At-the-money
BS	Black-Scholes
BF	Butterfly
BIS	Bank for International Settlements
BV	Bipower variation
c.d.f	Cumulative density function
CF	Characteristic function
CIR	Cox-Ingersoll-Ross process
FX	Foreign exchange
GARCH	The generalized autoregressive conditional heteroskedasticity
GMM	Generalized method of moments
GMT	Greenwich mean time
HN	Heston-Nandi
IMF	International Monetary Fund
LSV	Local stochastic volatility
ME	Mean error
MGF	Moment generating function
MLE	Maximum likelihood estimation
NDF	Non-deliverable forward
ODE	Ordinery differencial equation
OTC	Over-the-counter
OTM	Out-of-the-money
ON	Over night
p.d.f	Probability density function
RMSE	Root mean squared error

RR	Risk reversal
RV	Realized volatility
SW	Spot week
TP	Tripower quarticity
1W/M/Y	One week/month/year tenor

Non-greeks Conventions

$W(t)$	Wiener process at time t
$Z(t)$	pure Poisson jump process at time t
$Y_i(t)$	state vector for economy i at time t
ξ_i	discount rate
$\boldsymbol{\alpha_i}$	expected rates of changes
B	State vector in Affine Jump Diffusion Model
K	Option Strike Price
T	Option Maturity
$C(.)$	Call Option price
$V_{ij}(t)$	Spot variance
$S_{ij}(t)$	Spot rate
Δ^{BS}	Black-Scholes Delta
q_i	interest rate of currency i
λ	arrival rate of jump
η	risk premium on stochastic volatility component
μ	jump size in Poisson jump process
σ_h	volatility coefficient in jump diffusion model θ long run mean
κ	mean reversion rate
φ	Parameter vector of interest movement
φ^{MKT}	Parameter vector of jump-diffusion model under data-generating measurement
φ	Parameter vector of jump-diffusion model under risk neutral measurement
σ^{BS}	Implied volatility
r_i	interest rate for currency i in Black-Scholes model
X_{ij}	Spot rate of currency pair ij
Y_{ij}	Log of spot rate of currency pair ij
Ψ_{ij}	Characteristic Function of currency pair ij
$R_{ij}(t)$	Log increment of spot rate of ij at time t
μ_{ij}	Mean of currency pair ij log return
$\rho_{ij,ik}$	Correlation coefficient between currency pair ij and ik
$\boldsymbol{C}$	Variance-Covariance matrix
$\boldsymbol{R}$	Implied correlation matrix
$V(t,T,\Delta)$	matrix of eigenvector as a function of $T-t$ and BS Delta Δ
$\Lambda(t,T,\Delta)$	diagonal matrix of eigenvalues as a function of $T-t$ and BS Delta Δ

Greeks Conventions

τ	Tenor
$\Psi(\cdot)$	Characteristic function
σ	Volatility
Δ	Black-Scholes implied delta
μ	Drift
θ	Parameter vector
ϖ	Call/put option indicator
η	Increment

Currencies

AUD	Australian dollar
BRL	Brazilian real
CAD	Canadian dollar
CHF	Swiss franc
CLP	Chilean peso
CNY	Chinese Yuan Renminbi
COP	Colombian peso
CZK	Czech koruna
DKK	Danish krone
EUR	European euro
GBP	Pound sterling
HKD	Hong Kong dollar
HUF	Hungarian forint
IDR	Indonesian rupiah
ILS	Israeli new shekel
INR	Indian rupee
JPY	Japanese yen
KRW	South Korean won
MXN	Mexican peso
MYR	Malaysian ringgit
NOK	Norwegian krone
NZD	New Zealand dollar
PHP	Philippine peso
PLN	Polish zloty
SEK	Swedish krona
SGD	Singapore dollar
THB	Thai baht
TRY	Turkish lira
TWD	New Taiwan dollar

USD	United States dollar
ZAR	South African rand
XAG	Gold
XAU	Silver

Chapter 1

Introduction

1.1 Introduction

Decomposing a positive definite matrix via eigen-decomposition is one of the most common quantitative methods in finance. Typically, but not always, this matrix is the covariance of returns from asset prices. However,

This collection of research essays looks at a set of applications within the area of derivatives pricing and portfolio management. My first application addresses the complex issue of pricing foreign exchange (FX) derivatives and in particular foreign exchange options. Typically, FX options are looked at in terms of small triangles and these are used to fill gaps using parametric models. I extend the use of functional correlation analysis (where the correlation matrix varies continuously as a matrix function) to look at large scale sets of cross currency options. Using this approach I can price infrequently traded options as a function of the collection of options on a broad basket of other currency pairs. Uncertainty can be reduced to a single correlation guided and bounded by the eigen decomposition.

My second application is to portfolio management. Here I extend existing work to look at the stability of the eigenvector of the largest eigenvalue of the realized covariance

matrix. In contrast to the FX options case, where the eigenvalue and vector pairs are deterministic functionals, here the matrix is stochastic. I use simulation evidence to illustrate the stability of the portfolio structure and answer a series of missing questions from the principal component literature.

1.1.1 Thesis topic

Correlation and covariance matrix decomposition via eigensystems generates a set of eigenvalues and eigenvectors, used in a number of statistical methods. Eigenvalue analysis plays a key role in a wide range of fields from mathematics to telecommunication engineering. For finance and economic research, Principal Component Analysis(PCA) and Factor Analysis(FA) are known to reduce the data dimension by summarizing and selecting a smaller number of factors while retaining the major variation of the data set.

Eigenvalue analysis has been tested and proved be a reliable method for low-frequency data. But with the improvement of computer science and financial technology, it is necessary to verify the eigenvalue and eigenvector structure across high frequency dataset. My thesis will look at two novel applications and develop the mathematical preliminaries and statistical artefacts to solve a number of difficult asset pricing problems.

First I look at derivatives and in particular the difficult pricing problem of constructing volatility/co-volatility matrices for large cross-sections of currency options. The unique difficulty of this problem is that the space of the eigensystem is now itself a Hilbert space that is a function of the option delta. This is the currency option version of the classic option smile. As the currency market is one of the largest financial markets, it is important to price the options precisely.

For currencies with a high level of liquidity, there is sufficient data for investors to follow and analyse in this fast-changing market. But when there are a large number of currency pairs, some options may not have sufficient transactions to construct an implied

volatility smile. I can use eigen decomposition of actively traded pairs to narrow down the bounds on the un-traded pairs implied volatility smile. I will then define a metric to determine the correct point estimate for the correct eigenvalue. In addition to that, machine learning will be used to forecast and predict the missing prices.

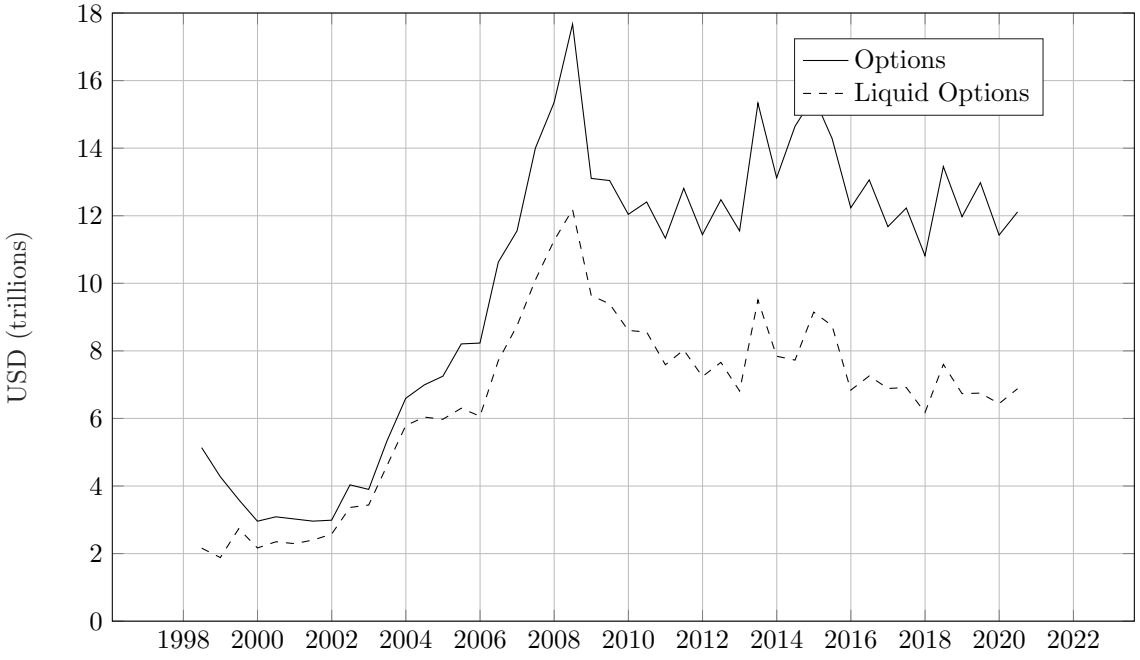
Second, I will look at eigen decomposition of the realized variance covariance matrix. S&P 500 stocks will be taken as an empirical dataset. I will go through equities covariance matrix to explore the eigenvalue/eigenvector structure. With the high-frequency data in the equity market, I will show the convergence critical point for the principal components. A set of thresholds and measurements are presented to illustrate the eigensystem structure and provide guidance for portfolio management.

Unlike the FX derivative OTC market, the S&P 500 has sufficient tick data to make it possible for us to check the effect of different data steps and provide an estimation window to eigenvalue structure. An overview is presented to generate the principal components structure, I will generate the constant characteristic across time series and other factors (estimation period and steps between data).

High frequency data, especially tick-by-tick data either from the FX options market or equity market, is known to be noisy and hard to call directly. This thesis tries to eliminate the influence of noisy data in two ways: 1) the data are carefully cleaned based on a specific algorithm 2) Monte-Carlo method is implemented for the simulation, which aim to create an ideal market pricing method to identify the accurate level of eigen decomposition.

The remainder of the chapter is organized as follows. I firstly present some essential background on the FX options and equity markets. Then I look back at the development of financial technology such as machine learning algorithms and the Monte-Carlo methods.

Figure 1.1: FX nominal Amount



Note: This figure presents the notional amount outstanding(in USD trillions) of FX option market, including gold, reported by BIS from 1998 to 2020. BIS categorizes option contracts into liquid currency pairs and illiquid pairs based on the underlying asset.

1.1.2 Background on the foreign exchange options market

The foreign exchange options market is known as one of the most active and important over-the-counter(OTC) derivatives market. According to data published by the Bank for International Settlements(BIS), the overall FX option market nominal amount was $ 5.13 trillion.In the first half year of 2020, the market volume had increased to $ 12.1 trillion.

The FX option is a financial tool to hedge the FX rate risk and it shares similar characteristics with options on other types of underlying assets, such as equity options. But it also forms its unique trading properties to fit its underlying asset. The first feature of FX related derivatives trading is that it involves a currency pair in the trading, which

means there is a potential need to measure and take two currencies into consideration. Investors need to pay attention not only to the policy and trends within a certain country, as two or even more sets of factors need to be measured. For example the adjusted option pricing model for FX would require two risk free rate from corresponding countries. After the trading, which currency is used to pay the premium is another point need to be predetermined.

A second characteristic of the FX option as distinct from the equity option is the quotation style. This is also caused mainly by the fact that currency is traded as underlying asset in contracts. For the equity option, the volatility surface is quoted by plotting option premium versus strike price, normally with both of them measured in USD. However in the FX option market, options are quoted by implied volatility surface derived from the Black-Scholes model, meanwhile the moneyness of options is presented by the delta, the sensitivity of the option price to underlying asset price. Further information on this will be given in the following chapters.

Another important feature is that FX options trading is significantly affected by the underlying spot liquidity. As global dominating spot, in 2020 86.8% of FX options were trading on USD. Further more, the other leg currency in addition to USD is mainly traded between six spot: EUR, JPY, GBP, CHF CAD and SEK. Fig. 1.1 presents the FX option market volume time series. It can be concluded that all six spots(EUR, JPY, GBP, CHF, CAD and SEK) with high-level liquidity took more than half of the whole market. In 2000, the amount of illiquid FX options was $ 0.73 trillion, taking 25.3%. This market rose to $ 5.23 trillion in 2020, with the proportion of illiquid option rising to 44%.

In addition, Table 1.1 reports the the nominal amount outstanding of OTC FX options market by currency for the last twenty years(2000-2020). From the table we can see that USD related trading was constantly higher than 80%, taking the major part in the whole market. EUR and JPY are considered to be two global spots providing sufficient liquidity

Table 1.1: OTC FX option notional amounts outstanding by currency (in trillion USD)

	Total	USD	EUR	JPY	GBP	CHF	CAD	SEK	Other
2000	2.89	2.5	0.97	0.92	0.34	0.18	0.11	0.03	0.73
2002	3.69	3.16	1.77	1.21	0.35	0.21	0.12	0.12	0.46
2004	7	6.22	2.75	2.43	0.68	0.31	0.22	0.12	1.27
2006	11.03	8.82	4.26	4.15	0.93	0.64	0.33	0.12	2.81
2008	12.8	10.01	5.28	4.85	0.68	0.98	0.21	0.12	3.46
2010	11.07	8.87	3.79	3.54	0.8	1.01	0.27	0.12	3.75
2012	11.26	8.71	3.85	3.6	0.58	0.65	0.3	0.09	4.74
2014	15.33	13.6	4.38	3.95	0.88	0.97	0.4	0.1	6.37
2016	11.53	10.38	3.33	2.83	0.8	0.39	0.33	0.21	4.79
2018	11.84	10.34	3.67	2.49	1.13	0.37	0.33	0.1	5.24
2020	11.82	10.26	3.71	2.62	0.73	0.46	0.55	0.08	5.23

Notes: This table presents the notional amounts outstanding of OTC FX option market by currency for the twenty years from 2000 to 2020, reported by BIS. Notional amounts are quoted in trillion USD.

for the derivative market not only option contracts. GBP,CHF CAD and SEK, as the currencies of dominant regional economies, are also traded at high frequency for risk management and investment.

But with the advancement of developing countries, market participants have been investing more into options trading on these illiquid spots. Investors used to trade these spots with USD and then transfer among themselves, now they require direct trades between two illiquid countries in order to reduce the transaction cost and risk. However, due to the limitation of low liquidity, there is little historical data for investors to refer. This has led to the problem in option markets where there is a large spread in the distribution of quotation for options on illiquid currency pairs.

In practice in the options market, pricing models such as Black-Scholes normally can not provide an ideal fair price for contracts by themselves. Investors improve the model performance in two ways: firstly, additional data called 'local factors' are used for adjustment; Secondly, historical data can be used for calibration. Calibration is a widely used process in correcting pricing error, which normally requires a continuous historical quote. The reason why calibration is important is that in most cases the pricing

model can provide a relative rational movement and trend for the price changes, but in a sharply floating option market, the level of price is hard to directly calculate from model. Participants will collect the aim historical data from market, normally the window period is between seven days and two weeks, then the calculation result from the model will be corrected by moving the level to the historical level.

By calibration and re-calibration, investors can generate acceptable predictions for a market quote in most cases. But as mentioned above, for illiquid currency pairs, there is insufficient data to calibrate with. So for these contracts, the key question is how to generate a relatively reliable model without historical data supplementation. A rapidly increasing number of industry institutions are eager to obtain a model for pricing these unexplored options markets.

The first section of this thesis presents a model based on the analysis of the market's eigen value structure to derive these 'missing values' for the options market. We have a deep insight into the implied volatility, correlation and covariance matrix, promote a model which does not rely on the historical market quotes. Then I implement the model for a wide range of data to verify model performance. The practical section covers different contract types, from contracts that are relatively stable and for which it is easier to forecast at-the-money option, to deep out-of-money contracts. The eigen analysis based model shows a constant outstanding accuracy level for all contracts.

1.1.3 Market Conventions

Because currency is viewed as a commodity in the FX market, there are two kinds of quotation styles generated naturally: domestic currency per foreign currency or foreign currency per domestic currency. Further, the definition of 'domestic currency' in FX trading can be ambiguous because both counterparts will think of themselves as domestic investors. In order to evade these conflicts, there are a number of market conventions

to quote the currency pairs, regardless of the nationality of investors. The US dollar, as the most important leg currency in the FX market, is always defined as foreign currency, except for trading with EUR, GBP, AUD, NZD.

In addition to quotation, settlement for FX option is another point that needs to be mentioned. For most currency pairs, option contracts follow the 'T+2' settlement rule. This settlement rule means that the payment of a certain contract is made two days later than the date when the transaction is agreed. These two days take the holidays and non-trading days into consideration for both domestic and foreign countries. There are some exceptions 'T+1' settlement for certain currency pairs, eg. EURTRY, EURRUB, CADTRY, CADRUB and TRYRUB. These specific settlements introduce more risk into the trading for both counterparts.

1.1.4 Option quotation style

The FX options market quote contracts in a way that is significantly different to the equity option market. Table 1.2 reports quotes cross maturities and strategies on 19 October 2018 for option trading on EURUSD. It can be concluded from the table that the FX options market participants quote contract by Black-Scholes implied volatility instead of price. For each trade the actual price for both counterparts requires further calculation.

Because the FX market quotes Black-Scholes implied volatility, the contracts' strike price is replaced by one of the Greeks, delta. The European volatility surface for OTC FX option market is formed by five key deltas: 10/25 delta for put and call option and at-the-money delta. The steps to get the volatility surface are as follows:

1) Collect Black-Scholes implied volatility for at-the-money contracts.

2) Collect quotes for four key strategies: 10/25 delta risk reversal and 10/25 delta butterfly.

Table 1.2: Volatility surface of EURUSD on October 19th, 2018

	ATM	25D RR	25D BF	10DRR	10DBF
1D	4.340/5.630	-1.135/-0.235	-0.170/0.470	-1.910/-0.365	-0.110/0.920
1W	6.925/7.525	-1.075/-0.655	-0.005/0.295	-1.845/-1.125	0.155/0.635
1M	7.000/7.200	-0.980/-0.840	0.095/0.195	-1.680/-1.440	0.310/0.470
6M	7.455/7.655	-1.315/-1.175	0.240/0.340	-2.310/-2.070	0.775/0.866
1Y	7.605/7.805	-1.340/-1.200	0.300/0.400	-2.385/-2.145	1.020/1.180
18M	7.825/8.035	-1.230/-1.080	0.300/0.405	-2.180/-1.930	1.015/1.185
2Y	7.940/8.165	-1.175/-1.020	0.305/0.415	-2.090/-1.820	1.065/1.245

Note: This table presents volatility surface data of EURUSD on 19 October, 2018. It includes the volatility bid (left) and ask (right) quotes with tenor from one day up to two years. Reported by Bloomberg.

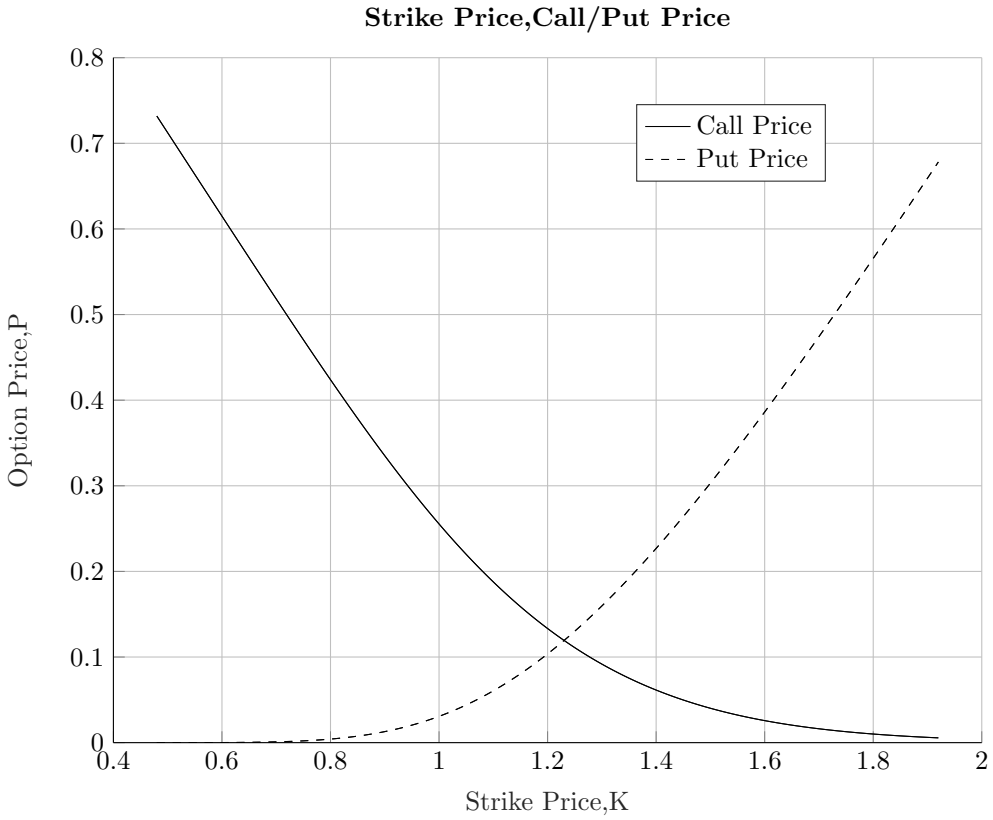
3) Derive the quotes for 10/25 put and call option with known implied volatility.

Fig. 1.2 to Fig. 1.5 show the process for transferring standard market quote (price versus strike price) to FX derivative quote style. Fig. 1.2 presents a normal volatility smirk; Fig. 1.3 generates a 'volatility smile' with strike price K on x-axis. Then Fig. 1.4 reveals the fact that for a certain strike price, I would have an 'opposite' delta for call option and put option. Fig. 1.5 generates a volatility surface in the format of the FX options market, the implied volatility is quoted versus call delta.

The details for risk reversal, butterfly and at-the-money will be introduced later. Market participants taking any positions(bid or ask) in this market can compare the price with volatility surface. The reason why the FX options market chose Black-Scholes implied volatility as quotes for all contracts is closely related to the fact that the underlying asset is a currency pair. If the actual price was taken, counterparts would find it confusing to make any comparison across this global market. Black-Scholes implied volatility act as a device to generate a normalized price for all participants. Risk reversals and butterflies provide information on the shape of the volatility surface by describing skewness and kurtosis level.

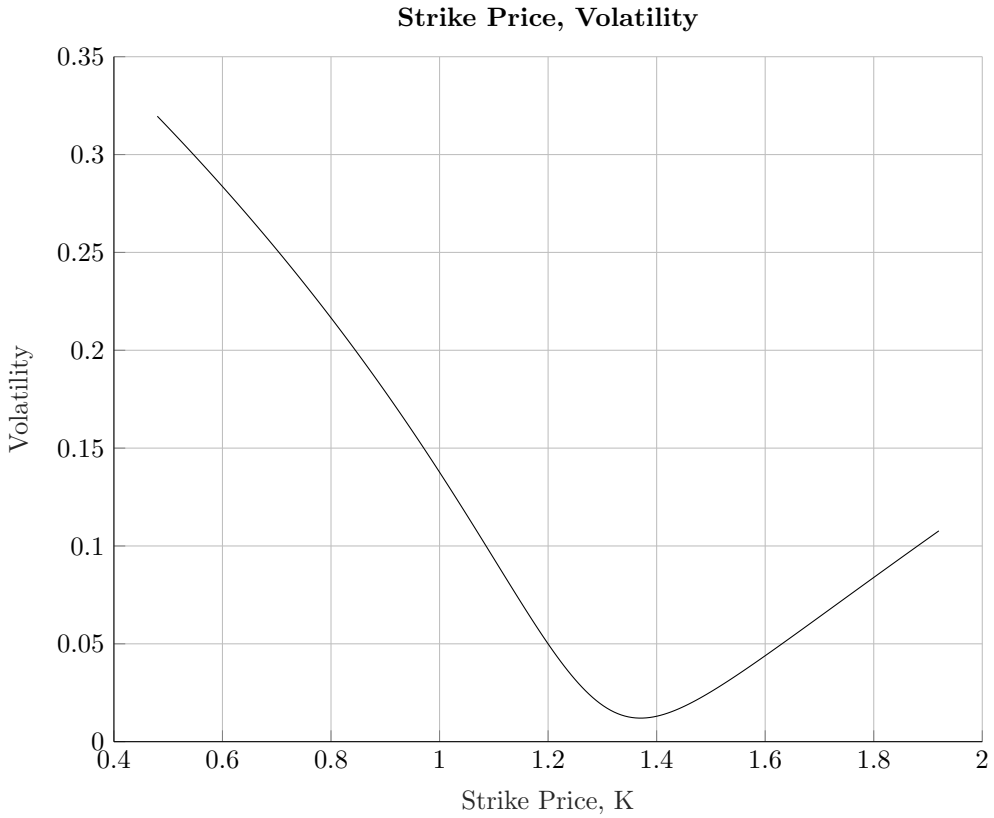
In practical study, the volatility surface is completed based on these key deltas' quotes,

Figure 1.2: Option price versus strike price



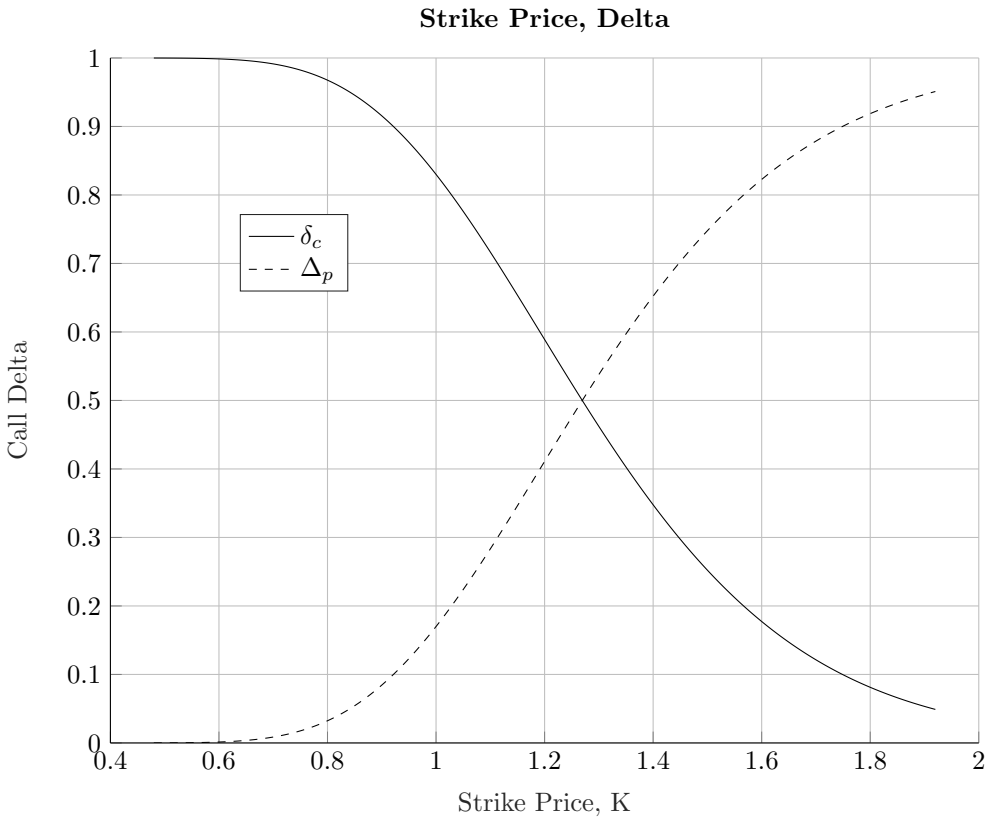
Note: This table presents option price versus strike price. Put price is generated via put-call parity.

Figure 1.3: Implied volatility versus strike price



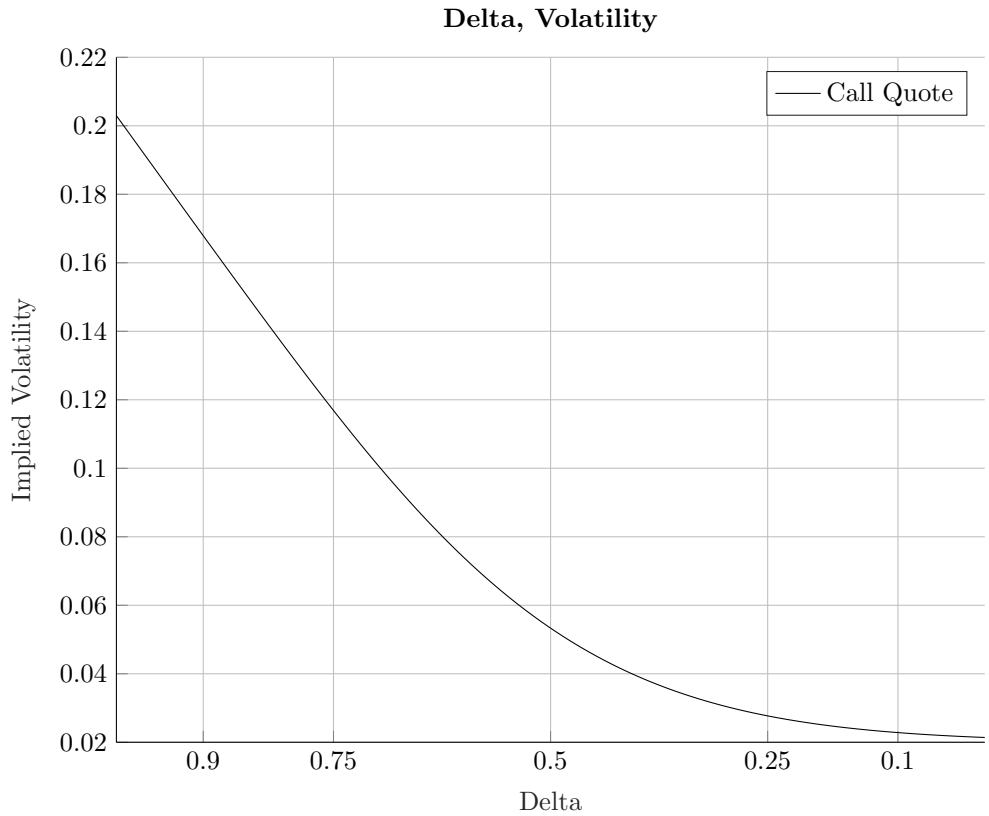
Note: This table presents the change of implied volatility versus strike price. Implied volatility is derived by making Black-Scholes model's calculation result equal to market quote.

Figure 1.4: Call delta versus strike price



Note: This table presents changes of call delta versus strike price. In FX market, options are widely used for hedging, it is important to transfer the strike price to delta for hedge.

Figure 1.5: Implied volatility versus Delta



Note: This table presents implied volatility versus call delta. This is the volatility surface used in FX option market instead of 'option price versus strike price'.

then an approach such as SABR or Vanna-Volga will be implemented to complete the surface. The quotes for contracts located in the gap between 10-25-atm deltas also provide important liquidity for the market, and methods to forecast the prices for these contracts are well established with a reliable accuracy level. One of the puzzles in the FX options market is how to price the contracts with delta lower than 10 delta. Firstly, the historical data for these deltas are rare, and secondly nonsensical negative implied volatility can be generated using traditional methodology.

In order to better understand FX options, It is necessary to discuss the delta used in option trading. Delta is the first order derivative of the option price with respect to the changes in the underlying asset. So, depending on the quotation format, the delta will be slightly different. Generally speaking, the only point that needs to be noted here is the difference caused by counterpart valuation adjustment.

For domestic investors in FX option trading, when the premium is exchanged in domestic currency, I will consider the premium itself as risk free and that no more risk is involved to the investor's portfolio. In this circumstance no adjustment needs to be made. But for foreign investors who will take foreign currency as the premium, the premium is not risk free and will contribute an extra risk. So the delta will be adjusted for this extra risk.

In practice, whether one adjusts the premium or not leads to different types of delta. In cases where the premium is risk free, the delta is called percentage delta. Meanwhile if the premium is risk free itself, I will implement the pips delta directly. The premium format is fixed for the most part; for example US dollar is accepted as premium currency for most contracts trading on developing currencies versus USD.

Although market participants rely on volatility surface to provide guidance for the comparison of and guidance on investments, normally in FX options institutions seldom directly transform quotes for a certain delta. Instead, they will describe the Black-Scholes

implied volatility for a specific strategy(risk reversal or butterfly). In the following I will denote the implied volatility as $\sigma(\Delta)$ where $\Delta \in 10, 25, 50$ present the implied volatility for contracts and their corresponding delta. The strategy widely used will be denoted as $\tilde{AT}$ for at-the-money, $\tilde{BF}$ for butterfly and $\tilde{RR}$ for risk reversal.

1.1.4.1 At-the-money

At-the-money, denoted as $\tilde{AT}$, provides the largest liquidity to the option market. In most cases, at-the-money refers to the volatility($\tilde{AT}$) and its corresponding strike $K(\tilde{AT})$, which can be used to construct a straddle where the volatility is eliminated:

$$\Delta_c(K(\tilde{AT}), \tilde{AT}) + \Delta_p(K(\tilde{AT}), \tilde{AT}) = 0 \tag{1.1}$$

where $\Delta_c(\cdot)$ denotes the delta for call option and $\Delta_p(\cdot)$ denotes the delta for put option. FX option market participants normally buy at-the money by this straddle, the purest way to buy the volatility at the middle of strikes.

In addition to the at-the-money given above, there is another at-the-money definition. In this special case, at-the-money could be interpolated as forward price:

$$K(\tilde{AT}) \equiv F(t, T) \tag{1.2}$$

This convention is used only for currency where USD is against a currency from the Latin American emerging market.

1.1.4.2 Butterfly

Butterfly is a confusing concept in industry and academia. The term 'butterfly' in market quotation refers to a single approximated volatility premium, which if added to the at-the-money volatility can be used to price a market strangle that longs out-of-the-money put

and call option at the same delta level. Generally speaking, OTC FX option quotation for market strangles are butterfly for 10 and 25 delta respectively. But in practice, volatility surface is usually heavily smirked, so the 2-vol butterfly can diverge substantially from the 1-vol butterfly strangle.

Most academic literature reviews the butterfly in 2-vol butterfly, which is considered to be a description of the volatility surface that contains more information. The 2-vol butterfly is defined as the premium of the average of 25/10-delta call and 25/10-delta put over the at-the-money volatility:

$$\tilde{BF_{25/10-2vol}} = (\sigma_c(25/10) + \sigma_p(25/10))/2 - \tilde{AT}. \tag{1.3}$$

In contrast with the academic definition and quotation style, the market participants in the FX options market turn to use the 1-vol butterfly as approximation. Because strike price could be back derived from 25/10 delta, the 1-vol butterfly is expressed as:

$$\Delta_p(K_p(\tilde{BF}_{25-1vol}, \tilde{AT} + \tilde{BF}_{25-1vol}) = -0.25 \tag{1.4}$$

$$\Delta_c(K_c(\tilde{BF}_{25-1vol}, \tilde{AT} + \tilde{BF}_{25-1vol}) = 0.25 \tag{1.5}$$

It is clear from the equation that 1-vol butterfly is back-derived from option price and this is the actual definition which FX options trading institutions refer to 'butterfly'. Compared with 2-vol butterfly, although the 2-vol butterfly follows the definition, I still need to transfer the quotes between two methods. For a volatility surface which is not greatly skewed, there is no significant difference between the two kinds of butterflies, but as mentioned above, the FX options surface is described as a special surface with a high level of skewness. This can lead to significant differences between academia and industry's values for a set of chosen contracts.

1.1.4.3 Risk reversal

Risk reversal is used to describe the anticipated skewness of volatility surface by tenor. It is relatively straight forward to calculate as the difference in volatility of call and put options with the same delta. The 10-delta risk reversal($\tilde{RR}_{10}$) is computed as follows:

$$\tilde{RR}_{10} = \sigma_c(10) - \sigma_p(10) \tag{1.6}$$

25-delta risk reversal follows the same rule.

1.1.5 Special Volatility Surface

Volatility surface is a concept that arises in option market. Formally, volatility surface is a plot of the implied volatility versus strike price for a set of option contracts for a given maturity. In practical study, options whose strike price differs from underlying asset price command higher price, which lead to a 'smile' type plot, so volatility surface is also known as 'volatility smile'.

The particular format of volatility surface show slight difference with the underlying asset. For stock option, the volatility surface is in the normal format with implied volatility versus strike price. FX options, due to the quotation style, have a special volatility surface in practical.

We can conclude from the quotation style for strategies(atm,butterfly and risk reversal) that OTC FX options market participants focus on the Black-Scholes implied volatility but not actual price, and the description for strike price is also replaced by delta. This quotation method might confuse new investors at the very beginning compared with the straightforward price quotation style, but it makes the information exchange and deal more efficient in this special market.

Implied volatility quotation allows market participants to ignore the spot rate prob-

lem in negotiating. The price for all currencies is 'normalized' and put into a unique measurement. All participants are discussing delta and volatility, but not price. This will make it easier for both domestic and foreign investors to communicate and make deals. The actual price would need to be calculated respectively for both sides only after the agreement and when they transfer the premium.

In addition to the convenience in terms of communication, another advantage of this quotation style comes from the background of OTC FX options market participants. Many of the investors in this market are financial institutions rather than single investors, and the aim of these institutions is constructing delta-hedged, gamma-hedged or vega-hedged portfolio but not to simply buy or sell an option. The risk and delta are directly quoted in this quotation style, so the portfolio management is straight forward for these investors.

1.1.6 Implied volatility and Realized volatility

As discussed in the previous section, implied volatility is the normalized price to some extent,which is closely related to option trading for all kinds of underlying assets. So implied volatility is a forward-looking information filtration; the market participants' forecasting and opinion on an underlying asset's future trend is presented in the implied volatility. Compared with realized volatility estimated from historical data, the implied volatility is more sensitive to market news, and the fluctuation is more significant.

Firstly, the average level of implied volatility is significantly higher than realized volatility in a long enough window period, because implied volatility contains not only the historical risk but also the unknown future risk. Secondly, implied volatility reaction to market changes is consistently more sharp than realized volatility. Realized volatility, as the back-looking estimation of risk, although adjusted by giving more weight to recent data in some method, still can not present market participants' behavior in as timely a

way as implied volatility.

The point that needs to be noted is that although implied and realized volatility come from different estimated/calculated procedures, have different reactions and levels to the market, they still share the same trend in time series. It is worth viewing one of them as the predictor/variable for the prediction of the other. Several studies focus on the predictive power of implied/realized volatility on realized/volatility, and in most cases the predictive power is identified to be significant.

Due to the fluctuation of implied volatility, it is extremely difficult to predict it based on the traditional pricing methodology, such as the Heston-Nandi model. So in practical investments, the price of the option is highly reliant on the historical data. Institutions will use historical data to do calibration/re-calibration.

While the trend and movement of model calculated is considered to be rational and relatively reliable for most FX options contracts, the main problem is how to adjust these calculated results to the market level. The calibration procedure is straightforward: firstly, market participants obtain calculated results from their chosen methodology. These results are expected to be far away from the market quote. Then historical data from the most recent window period(normally set to be two to four weeks) are collected, and the calculated results will be adjusted to move the level to market quotes. 'Re-calibration' involves re-running the calibration, and is the most widely used methodology to get a relatively accurate price(implied volatility) for options contracts.

In situations where there are insufficient data to calibrate with, it is hard for investors to price the options contracts. This is exactly what is happening for FX options based on illiquid currencies. Many of these are emerging currencies, and the only foreign currency in options contracts is the US dollar. For direct trading between two illiquid currencies one would go into a wilderness where there is no historical data for investors to calibrate with. In the past investors would give up the direct trading and chose the US dollar as a

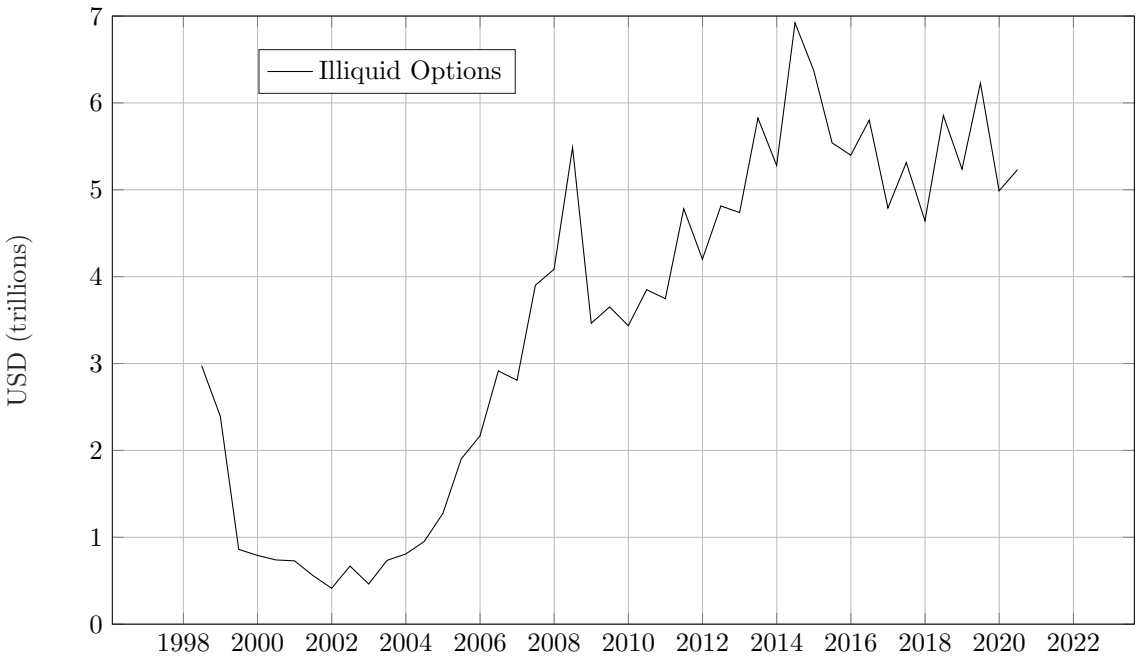
leg currency for constructing a portfolio. But with the development of emerging countries, these 'illiquid' currencies options are growing to form a market which can not be neglected. Market participants need a solution based on a non-calibration environment.

1.1.7 Volatility surface for illiquid currency pair

Building a volatility surface from market quotes is straightforward and easily done with sufficient data provided. As mentioned in the previous sections, there are five important quotes(two butterflies, two risk reversals, and at-the-money contract) over a set of deltas. With the combination of these quotes, the volatility surface is easily recovered. Some data vendors such as Bloomberg and Thomson Reuters provide a bid/ask price for every single tick of these key quotes. For the delta outside 10-call and 10-put, a structural model such as one/two factor Heston or other local stochastic volatility models could provide rational guidance. The deltas inside this range are also interpolated by methodology such as Vanna-Volga or SABR, and the volatility surface is smoothed based on five key quotes.

The key problem for volatility surface is how to generate accurate predictions over key quotes. The traditional method of calibration, as discussed above, is useless for illiquid currency pairs, therefore I turn to seek the information from the FX options market but not micro factors. As one of the largest and most liquid financial markets, the contracts provide liquidity to make the system stable, which indicates that the eigen value structure for a set of selected currency pairs can generate forecast-able movement. 'Illiquid' does not mean corresponding leg currencies are rarely traded in the FX options market - on the contrary, the amount of these currency pairs has increased to a level that can not be ignored. Fig. 1.6 presents the trend for illiquid currency nominal amount from 1998. It can be concluded from the figure that although the 2001 and 2008 financial crises momentarily stopped the increment, the market scale increased sharply to $7 trillions after 2010. Option trading based on one of the illiquid currencies and a global currency

Figure 1.6: Nominal Amount of Illiquid FX option



Note: The graph presents the notional amount (in USD trillions) of FX option trading on illiquid underlying asset, reported by BIS from 1998 to 2020.

(such as US dollar, Euro and Japanese Yen) has a high level of liquidity. In addition to the combination of emerging-global pairs, the contracts for these currencies mainly focus on the emerging currency and its regional trading partners, for example Argentina and Chile. Option contracts for cross-region emerging currency pairs are extremely rare in all financial institutions.

Based on the background discussed, the volatility surface is hard to forecast, especially the deep out-of-money delta such as 10-call and 10-put. This leads to a vicious circle that with no historical data to refer and calibrate with, market participants give up the direct trading between two emerging currencies, finally resulting in a forbidden zone for FX investment. Our study provides a methodology to skip the traditional re-calibration

procedure and generate price guidance

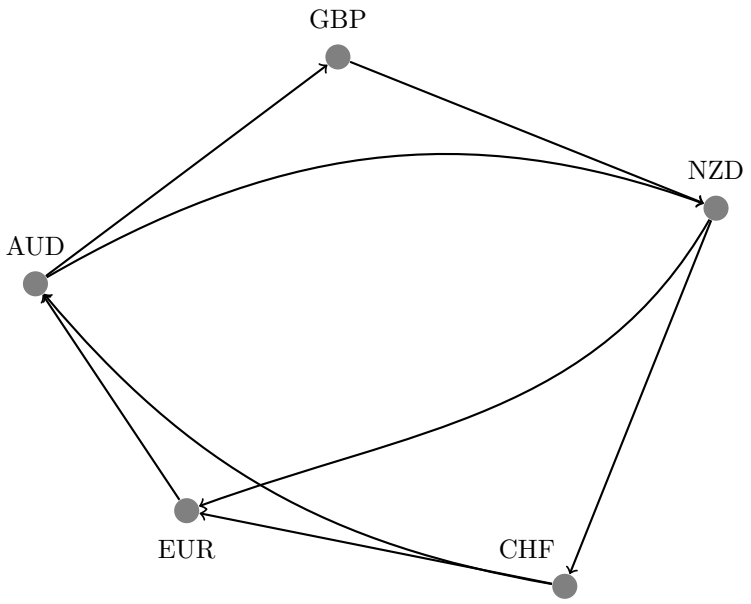
Traditional methodology in option pricing focuses on micro factors, e.g. interest rate, forward price and underlying volatility. This thesis aims to take the whole FX option into consideration, instead of only paying attention to the currency and its corresponding country.I build a leg currency set containing 6-10 related currency pairs and organize a co-variance matrix, with eigen decomposition made towards it.

In my study, I construct a process called 'leg currency journey', in which all related option information is organized into a single matrix. The illiquid currency pair's implied volatility is viewed as a 'missing value' in the system, and the eigen value structure is used to back-derive the missing value. The key point here is how to control the freedom level of the incomplete matrix, arrangement of leg currency journey and covariance matrix limit the currency amount involved. Fig. 1.7 presents an example of leg journey between several liquid currencies (GBP, EUR,AUD, NZD, AUD and CHF). Different choices of 'journey path' are shown in the figure. For any currency pair, there is a set of contracts with different maturities and delta. The leg journey provides the background to construct a covariance matrix for the purpose of implementing eigensystem analysis.

In contrast to option trading with strategies and market straddle, our methodology organizes the matrix by delta. For all deltas (10/25 put/call delta and at-the-money contract), I have a set of leg currency journeys' implied volatility filtration to complete the system. The advantage is that I could directly generate the volatility surface and make comparison through deltas. No further transformation is required; the mislead between 1-vol butterfly and 2-vol butterfly can be perfectly avoided in our methodology.

Further more, I apply technology such as simulation and machine learning in our study to improve the forecast ability for FX options traded on illiquid currency pairs. I firstly give a brief introduction to the implementation of machine learning in financial realm.

Figure 1.7: Leg Journey for selected Currencies based on USD



Note: This figure presents an example of possible leg journey among a set of selected currency pairs(with USD).

1.1.8 Machine learning adjustment

The fundamental goal of pricing is to understand and measure the risk premia. Quoted by implied volatility, FX options contracts are made to manage the risk. Even for options trades on liquid currency pairs where expected price is well observed, it is still required to explain the behavior and economic factors behind. But implied volatility is widely known to be difficult to measure: firstly the market quotes are noisy, not all of the quotes are rational as some are made by algorithm to check the spread, and it is quite common to detect irrational bid/ask prices in tick history. Secondly, market efficiency makes implied volatility return to be dominated by unpredictable market news, which make the prices unforeseeable to some extent.

My study uses machine learning to capture factors that might have been neglected in previous literature. Before we go into detail about the implementation of machine learning in the FX option market, it is necessary to introduce this popular concept.The

definition of machine learning is context-specific. In academic study of finance, the term of 'machine learning' is usually used to describe a set of high dimensional models used for predictions which enjoy the advantage of searching out the most effective factors in a vast amount of variables.

Compared with traditional economical models, the high-dimensional nature of machine learning promote the ability to find the key factors from unknown variables. The nature of leg journey methodology decides that there are rare fixed factors, and all variables need to be specified to present the connection between chosen illiquid currency pairs and the FX option market. Underlying asset, strategy, maturity and even the number of factors are different across over 1400 contracts, so it would be exhausting to value the parameters and unfeasible to implement in practice.

The nature of machine learning makes it possible to find neglected connections behind the intricate FX market. In my study, I firstly generate an implied volatility prediction based on leg journey and corresponding implied covariance matrix, then I organized a set of related FX option prices as factors for machine learning. The advantages of using machine learning are significant in a number of respects:

1) Although eigenvalue analysis has led to an accurate forecast for missing foreign exchange options, there are still slight differences between calculation result and market value, especially for those currency pairs with a relatively 'short leg journey'.

2) In our case, machine learning directly takes known implied volatility into the model as adjustments;this information increases the information filtration for the currency leg journey. More extensive historical data is used in our model.

3) There are huge amounts of derivatives traded based on foreign exchange, but it is still relatively straight forward to recognize key factors/variables in empirical study. Global currencies with high levels of liquidity are naturally considered to be the factors in our study.

To conclude, I will present the impact of factors based on the foreign exchange market to explode the relationship between currency pairs. In previous literature, researchers mainly focus on impact caused by regional effect or from trading partners. I would show a broader point of view for this unpredictable market.

1.1.9 Equity eigenstructure analysis

Another practice of eigenstructure analysis is in the equity market. I organize the S&P500 equity historical data from 1996 to 2020 as dataset to generate the eigenvalue/eigenvector distribution. In previous study, scale of portfolio is normally limited to around 100, the distribution of eigenvalue is not significant. I expand the scale by constructing covariance matrix with one of the largest markets: S&P500 equities. the rule of eigenvalue are explored in my practice: the eigenvalue will converge when the scale of portfolio is large enough.

Eigenvalue analysis provides guidance on the portfolio constructing. I will show the simulation result to compare the eigenvalues between market and ideal circumstance, bounds are presented for market portfolio. In practice, there are two noted findings: 1) the largest eigenvalue are converged when the scale of portfolio is larger than 200 2) measurement and threshold for 'convergence' is defined in my study.

I aim to discover and discuss the implementation of eigenvalue structure analysis in equity market. I pay more attention to eigenvalue distribution in equity study more compared with FX option. The simulation results and bounds will also presented. Time series study cover the historical data from 1996 to 2020 and focus mainly on the trend of eigenvalues after 2015.

1.1.10 Conclusion

In my thesis, I focus on the eigenstructure analysis and implications in different circumstances. Firstly I raise a model to provide pricing guidance for FX option contracts. Compared with traditional models, an eigenvalue-based model reduces the dependency on historical data. For illiquid option contracts with few historical data to calibrate with, my model provides a reliable price guidance. Eigen models try to integrate information from option contracts all over the world and derive the 'missing' price for those illiquid contracts. The accuracy level is verified to outperform the traditional historical data-based model. In practice, I will use historical FX option data as the data set. The information of FX option data is recorded in Appendix A Table 5.2.

In order to improve the performance of the eigen model, I implement machine learning to reduce the error between calculation result and market quote. The algorithm I select is the regression tree, and results indicate that machine learning improves the prediction performance in all respects. In addition to prediction, I numerical importance level of currency pairs' by machine learning.

I focus not only on the FX options market, but also pay attention to the equity market. In the equity market, I focus on the convergence of the first principal component. I selected S&P500 as the empirical dataset for our study(details are recorded in Appendix A Table 5.1). In my empirical study, I expand the scale of covariance to 500 to check the trend of the explanation power of the first principal component. With the increment of portfolio scale, convergence of the first component is significant. I set methodology to set the threshold to distinguish the stable process from turbulent process of eigenvalue.

To conclude, I focus on the eigenstructure of market equities. The idea arose from different aspects of financial markets for which eigenvalue analysis can provide guidance for both pricing and knowledge of portfolio management. The rest of the chapters are organized as follows: in Chapter2 I introduce the eigenvalue analysis, and in Chapter3 I

provide the machine learning solution for the FX market. Then in Chapter4, I introduce the implementation of eigenvalue analysis in equity market. The last chapter forms the conclusion for the thesis.

Chapter 2

Eigenvalue Model

2.1 Introduction

Foreign exchange option (FXO) markets are notoriously difficult to calibrate. Highly liquid markets such as the dollar versus the currencies of the other G7 countries have specific rules of thumb to generate implied volatility smiles. However, a large number of currency pairs (including many versus the dollar) do not have any quotes for FXOs, and this is the subject of interest for this chapter. That many currency pairs do not have FXO quotes is something of a puzzle. Unlike equities which have high degrees of freedom relative to their underlying factor structure currency dynamics- under both the physical and risk neutral measures- must be consistent to eliminate triangular arbitrage opportunities.

Hence, with a number of traded quotes delivering the market price of contingent securities, if the average correlation between exchange rates is relatively high (say above 50%) then the portfolio of exchange rates has very few valid exchange rate volatilities to preserve positive semi definiteness of the conditional variance covariance matrix describing the individual and joint second moments of exchange rates. Interestingly, there is a discrepancy between the practitioner pricing of FXOs, which are apparently priced under

a Geometric Brownian Motion ($\mathcal{GBM}$), but are actually priced under a Brownian Semi-Martingale ($\mathcal{BSM}$), which has a far richer set of dynamics. I will illustrate how the $\mathcal{BSM}$ instantiates itself in the pricing curves for FXOs and then extend the current standard set of models to a richer framework of pricing models.

In this chapter I outline a global consumption economy with Cox et al. [1985] type nominal interest rates and equivalent zero coupon bond securities. I then derive the equilibrium exchange rate dynamics in the presence of time homogenous jumps and stochastic volatility. I will show that once risk neutral dynamics are derived the bounds on FXO pricing for un-quoted currency pairs are quite tight, particularly when considering a large number of alternate leg currencies; that is extending beyond triangular arbitrage into a case where more leg pairs are possibly paths from one numéraires to another.

My main contribution in this chapter is to develop a multi-currency economy whereby the spot price dynamics adjust as a Poisson point process relative to the prevailing economic conditions. For an N-numéraires (floating currencies, gold and silver) set up, to my knowledge this is the first model of its type. I then solve the model for a set of contingent claims that complete the market and translate this equilibrium economy into a transformed Black-Scholes setting in keeping with the standard quoting approach of FXO markets.

I then implement an eigenfunction approach which presumes that the delta portfolio returns (the return on a portfolio of spot and options contracts with a specific target delta) are correlated by an at least positive semi-definite matrix, hence by building a set of options prices, I can determine the upper and lower bounds on currency correlations when a suitable leg currency pair is absent.

My eigenfunction approach generates a European implied correlation surface and solves a mini-max analysis of the smallest eigenvalue of the implied correlation matrix. As the number of leg currencies increases the size of the pricing bounds on missing cur-

rency pairs decreases and this rate of decrease is proportional to the average correlation across the vector of numéraires.

In this chapter, I will firstly present basic assumptions I made towards underlying assets: a Cox-Ingersoll interest rate for both domestic and foreign interest rates and jump-diffusion model for spot process. Then the conception of FX leg journey will be introduced, I will present how to derive implied correlation matrices from leg journey and FX implied volatility.

Based on the characteristic of positive definite matrix, investors can derive any missing element in corresponding implied correlation matrix. In the eigen model, the illiquid option contracts' implied volatility are set as 'missing element' in the implied correlation matrices. In the empirical study, I access over 1700 FX vanilla contracts to apply the predictive power of the model, calculation results are compared with the market quote. Empirical section presents the performance of eigen model from two aspects: market quote as benchmark to estimate the predictive power of the model and the comparison with widely-used parametric model.

The rest of chapter is organized as following: Section 2.2 gives a general literature review of the prior work in the foreign exchange option market. Section 2.4 presents the jump-diffusion option pricing model and a brief introduction to the eigenvalue analysis for implied correlation matrix. Section 2.6 gives data information in this study.Section 2.8 gives summary and conclusion to the method.

2.2 Prior work on FXOs and FX options

Since Garman and Kohlhagen [1983] adapted the the Black and Scholes [1973] pricing model of contingent securities for foreign exchange options there has been a strong push to generate option pricing surfaces to value hedging and speculation on exchange rate volatility. In particular, there has been a separation of the option implied structure

of the FX market and analysis and forecasting of real and nominal exchange rates via some form of regression analysis, see for instance Hansen and Hodrick [1980] for an early example, using forward nominal exchange rates.

All starting points for an FXO pricing model must begin with a particular parity assumption. Typically, practitioner models utilize covered interest parities as opposed to real exchange rate models favoured in the macroeconomic forecasting literature. Garman and Kohlhagen [1983] provide the base model for FXOs and this is typically the model that is used by practitioners for quoting FXOs. This is a Black-Scholes type pricing model, with the driven driven by a non-stochastic difference in riskless deposit rates between two numéraires, for instance two currencies or a currency and a commodity such as a precious metal (typically gold and silver, presumed to exhibit non-stochastic forward curves in line with US deposit rates). For empirical implementation the theoretical riskless timed deposit accounts are replaced by quoted time deposits, hence in the pure Black-Scholes universe arbitrage is presumed to follow a Rubinstein [1976] style approach, with rebalancing always relative to a fixed point to point interest rate, as opposed to a continuously dynamic rebalancing in the pure Black-Scholes PDE framework. For an exact FXO PDE approach in a Black-Scholes universe, see [Chapter 9 of Shreve, 2004] for a full description. As it turns out in both cases the one factor asset pricing model yields the same solution, however, certain different assumptions are needed and this has relevance to my implementation later on. Some ad-hoc adjustments, in model, have been proposed an example of which is the Bossens et al. [2010] approach that uses the sensitivity of the Black-Scholes valuation to changes in spot and volatility to adjust prices to compensate, hence describing the implied volatility curve typically observed in most option markets, FXO or otherwise.

First, [Chapter 9 of Shreve, 2004] posits a case where three representative assets are available for domestic and overseas investments: two money market and one risky

'stock'. Each asset is within the Black-Scholes framework ($\mathcal{GBM}$, frictionless rebalancing) and can be dynamically swapped by changing numéraire. Each asset is valued under its own risk neutral measure, versus a money market account, with the domestic money market account being fixed to unity in its own numéraire. In each case the domestic and overseas money market accounts are both pure martingale processes. Hence, one can either stick to a domestic money market account, exchange to the overseas account or invest in a risky asset. [Chapter 9 of Shreve, 2004] illustrates that the resulting pricing model for call options on exchange rates, when the stock process is a $\mathcal{GBM}$ is equivalent to the Garman and Kohlhagen [1983] pricing under the Rubinstein [1976] timed account conditions. When specifying my general $\mathcal{BSM}$ model I will need to switch between these two pricing frameworks to derive a closed form price, both under a no-arbitrage and equivalent risk neutral pricing framework.[1]

Most standard models of economies do not presume fixed interest rates, but that the money-market account is subject to random shocks; see for instance, the classic equilibrium model of Vasicek [1977] or in a mean reverting framework Cox et al. [1985]. Indeed, for my re-implementation of a jump diffusion for multiple currencies I will presume that the underlying economy has a Cox et al. [1985] style consumption framework with mean reverting technology shocks. This forms the base risk neutral measure driving the domestic and overseas money markets. Timed deposit accounts are therefore the integrated expectations of the forward spot rate path under a one-dimensional mean reverting stochastic process. This might seem restrictive, but for the purposes of pricing FXOs, I am interested in the difference in the term structure and not modelling both individually, hence this is technically a two-factor model of the difference in the term structures and not two one-factor models of individual exchange rates (although I can take initial values from these models).

[1] That is by writing down a no-arbitrage, delta neutral, price, there must be an equivalent risk neutral measure, or by writing down equivalent risk neutral dynamics, there must be an arbitrage free price.

Modern option pricing utilizes a rich class of $\mathcal{BSM}$ models, most of which include some form of point process to complement the continuous diffusion dynamics. These models can be applied to both free float and pegged currencies. See Yu [2007] for an example of a point process being used to model a peg break shock. Hull and White [1987] and Heston [1993a] propose early examples of stochastic volatility type models for stock and bond option pricing, whilst Bates [2000] incorporated simple diffusive jumps into an index option pricing model to explain the change in the pricing structure of the US equity options market after the 1987 stock market crash. More recent work by Da Fonseca et al. [2014] has extended the continuous component of stochastic volatility to a multivariate case using Wishart volatility models. However, these models do not contain jumps and closed form solutions are somewhat difficult, given the computational nature of a Wishart characteristic function.

One-dimensional models of jump diffusions are common; see Duffie et al. [2000] for the definitive review. While Bates [2000] offers arguably the simplest model of a jump diffusion, Pan [2002] provides a set of dynamics for index options that are most applicable to the FXO case. It is along these lines (Cox et al. [1985] type dynamics of the numéraire and jumps in the mean equation only) that my benchmark model is created. In my case jump compensation is designed to mimic the friction of interventions (such as pegs) forcing the exchange rate away from its natural point. Hence a jump, is a jump-to-adjust and this is a useful interpretation for pure pegs such as the US dollar versus Saudi Arabia and China.

Estimating jump diffusions from spot data is inherently difficult. Singleton [2001], Duffie and Glynn [2004] Ait-Sahalia et al. [2008] and Aït-Sahalia et al. [2015] provide a variety of frameworks for parameterizing the physical process from spot data. However, the majority of option models are specified in close form and then calibrated to observed quotes. My approach is a halfway house. I presume that market quotes are not available

for certain pairs and it is then necessary to combine calibration from cross option pairs and estimation from spot data to complete the set of available options.

More recent work by Du et al. [2017] has postulated an equilibrium global economy with multiple numéraires whereby a deviation from the covered interest parity (and hence from the domestic and overseas risk neutral measures) are persistent, resulting from a variety of frictions. Their model encapsulates this premium and calibration suggests that pegs and spot market frictions are the most prominent source of systematic deviations from the nominal parity conditions.

In the two numéraires case the parity conditions are relatively straightforward to model. However, in a case where more than two numéraires exist, leg currencies must satisfy both the mean and variance conditions to be fully arbitrage free; that is, one cannot synthetically create a position in two leg currency transactions that replicates the cross (Dollar-Yen, Yen-Pound versus Dollar-Pound, for instance), that yields both an expected forward rate and a volatility inconsistent with the direct transaction. For instance, consider a dynamically rebalanced portfolio of 'stock' and options in two exchange rates with identical 'inner' pair, (e.g. Dollar-Yen and Yen-Pound). If the position replicates a position directly in the 'outer' pair (e.g. Dollar-Pound) then the two portfolios should track exactly, in terms of mean and variance. One can then resolve from the outer pair a factor (referred to as an implied correlation) that ensures that the leg exchange rate dynamics (e.g. Dollar-Yen and Yen-Pound correlations) are consistent with the outer' pair. This consistency should be maintained for both delta neutral positions and positions with partial hedging exposure.

Other than specification and times series forecasting, (see Campa and Chang [1998]) there is little specific academic research across a wide range of currencies in this area/ Ballotta et al. [2017] is one such study, and this case looks at a very narrow interpretation of the at-the-money straddle correlation versus the implied correlation (under the physical

measure) estimated from the relative spot price dynamics. Interestingly, there are no current research papers that look beyond triangular arbitrage arguments and I will show that there is considerable value to analysing broader numbers of currency pairs in this type of framework.

2.3 Contribution of the Chapter

In this chapter, I aim to provide a brand new methodology for the pricing of FX options with illiquid currency underlying asset. The model is based on the eigen decomposition of implied correlation matrix. As a widely used conception in FX option market, implied correlation comes from the implied volatility of pre-determined currency triangle.

In previous study of FX option pricing, traditional parametric models need sufficient historical data to calibrate with. In practical, they will use very last data to make adjustment for parameters, this process is known as 'recalibration'. For liquid and widely traded currency pairs, recalibration greatly improve the performance of the model, but for those illiquid contracts there is insufficient historical data to do the calibration.

Lack of data for illiquid FX option contracts reduces the predictive power of pricing models, which increases the difficulty of relative derivatives' trade. The Eigen model raised in this chapter overcomes this disadvantage and provides a reliable pricing method for illiquid contracts. In this chapter, I test the performance of the eigen model with more than 1800 contracts over different currency pairs, deltas and maturities. The results indicate that the eigen model's predictive power is consistent across underlying assets.

The performance of the eigen model shows slight difference between deltas and maturities. At-the-money contracts show the highest level of predictive power and the accurate level decrease with the increment of delta to the volatility surface wings. For maturity, the eigen model generates the best performance for short maturities from 1M to 6M.

This model provides a guidance of these illiquid FX option contracts for investors.

With the development of these corresponding countries, there is increasing requirement for the direct derivatives' trades between these currencies. As mentioned above, parametric models' performance are limited without sufficient historical data, but the eigen model could work without historical data. The data required in my model are just a well-structured leg journey and the implied volatility of related contracts.

In the rest sections of this chapter, I would introduce the basic assumptions towards FX market, the introduction of leg journey and practical performance of eigen model in the market.

2.4 A General Model of FX rates and FX Correlations

I consider a global economy with well-defined $i \in \{1, \ldots, N\}$ numéraires for well-defined component economies. For each component economy I presume a Cox-Ingersoll-Ross type consumptions economy, see Cox et al. [1985] in continuous time with a single representative consumer with a logarithmic utility function $\exp(-\xi_i t) \ln C_i(t)$, where ξ_i is a discount rate. Each economy has a vector of capital stock (technology and commodity stock valuations) positions $Y_i(t)$ with a corresponding vector of expected rates of changes in capital stock $\boldsymbol{\alpha}_i$ and covariance matrix $\boldsymbol{S}_i\boldsymbol{S}_i'$. I presume that the M_i length vector of capital stock is a vector of independent Ornstein-Uhlenbeck processes as follows:

$$\mathrm{d}Y_i(t) = (\xi_i Y_i(t) + \zeta_i)\mathrm{d}t + \mathrm{diag}[\boldsymbol{\nu}_i]\sqrt{Y_i(t)}\mathrm{d}\tilde{\boldsymbol{w}}^i(t)$$

where the first two moments of $\mathrm{d}Y_i(t)$ are given by $\tilde{\boldsymbol{a}}s_i$ and $\boldsymbol{G}_i\boldsymbol{G}_i'$ respectively. Setting $\tilde{a}_i = \tilde{a}_i Y_i(t)$, $\boldsymbol{G}_i\boldsymbol{G}_i' = \boldsymbol{\Omega}_i Y_i(t)$ and $\boldsymbol{G}_i\boldsymbol{S}_i' = \boldsymbol{\Sigma}_i Y_i(t)$, where $\tilde{\boldsymbol{a}}_i$, $\boldsymbol{\Omega}_i$ and $\boldsymbol{\Sigma}_i$ are arbitrary constants, the classic result of the CIR model is that when consumption is chosen to maximize

$\mathbb{E}_t[\int_t^T U(C(s), Y(s), s)\mathrm{d}s]$, the equilibrium interest rate under logarithmic utility is given in closed form by:

$$q_i(Y_i(t)) = \frac{\mathbf{1}'\mathbf{\Omega}_i\tilde{\mathbf{a}}_i - 1}{\mathbf{1}'\mathbf{\Omega}_i\mathbf{1}} \tag{2.1}$$

where $\mathbf{1}$ is a unit vector of length M_i, essentially solving the classic consumption minimization problem in continuous time. From this construct the nominal interest rate dynamics for the i^{th} numéraires is a one-dimensional square root diffusion model of the following form:

$$\mathrm{d}q_i(t) = \left(\frac{\mathbf{1}'\mathbf{\Omega}_i\tilde{\mathbf{a}}_i - 1}{\mathbf{1}'\mathbf{\Omega}_i\mathbf{1}}\right)'(\xi_i Y_i(t) + \zeta_i)dt + \left(\frac{\mathbf{1}'\mathbf{\Omega}_i\tilde{\mathbf{a}}_i - 1}{\mathbf{1}'\mathbf{\Omega}_i\mathbf{1}}\right)'\sqrt{\boldsymbol{\nu}_i\boldsymbol{\nu}_i' Y_i(t)}\mathrm{d}W^i(t) \tag{2.2}$$

$$\mathrm{d}q_i = \kappa_i(\theta_i - q_i)\mathrm{d}t + \sigma_i\sqrt{q_i}\mathrm{d}W^i(t) \tag{2.3}$$

where $\mathrm{d}W_i(t)$ is the combination of the vector Weiner process $\tilde{\boldsymbol{w}}_i(t)$. Now let $i, j \in \{1, \ldots, N\}$ be an index of our numéraires and $S_{ij}(t)$ be the exchange rate of the i^{th} 'domestic' numéraire to the j^{th} at time t. Let $q_i(t)$ and $q_j(t)$ be the nominal short rate for $\{i, j\}$ numéraire pair, with the CIR type dynamics:

$$\mathrm{d}q_i = \kappa_i(\theta_i - q_i)\mathrm{d}t + \sigma_i\sqrt{q_i}\mathrm{d}W^i(t) \tag{2.4}$$

$$\mathrm{d}q_j = \kappa_j(\theta_j - q_j)\mathrm{d}t + \sigma_j\sqrt{q_j}\mathrm{d}W^j(t) \tag{2.5}$$

by replacing the terms of (2.2) with their scalar parametric equivalents. Collecting parameters I denote $\varphi_q = (\varphi_{q_i}', \varphi_{q_j}')'$ with $\varphi_{q_i} = (\kappa_i, \theta_i, \sigma_i)'$ and $\varphi_{q_j} = (\kappa_j, \theta_j, \sigma_j)'$.

The dynamics of the spot exchange rate for the delivery of one unit of currency i for a given amount $S_{ij}(t)$ of currency j, referred to as the the ij^{th} currency pair is given by

the following jump diffusion model:

$$\begin{aligned} \mathrm{d}S_{ij}(t) =&(q_i(t) - q_j(t) + \eta V_{ij}(t) + \lambda V_{ij}(t)(\mu - \mu^*))S_{ij}(t)\mathrm{d}t + \\ &\sqrt{V_{ij}(t)}q_i(t)W^{ij}(t) + \mathrm{d}Z(t) - \mu S_{ij}(t)\lambda V_{ij}(t)\mathrm{d}t \end{aligned} \tag{2.6}$$

with spot variance given by

$$\mathrm{d}V_{ij}(t) = \kappa_h(\theta_h - V_{ij}(t))\mathrm{d}t + \sigma_h\sqrt{V_{ij}(t)}\mathrm{d}W^h(t) \tag{2.7}$$

where η is the risk premium on the stochastic volatility component, $V_{ij}(t)$, which is presumed to be a square root diffusion in keeping with Heston [1993b]. λ is the arrival rate of jumps from a pure jump process $Z(t)$, with mean jump size μ and jump compensator μ^*. Jumps are presumed to arrive with intensive λ, the arrival following Pan [2002] is as follows, let $U_{\tau^*} \sim \mathcal{N}(\mu_{ij,J}, \sigma^2_{ij,J})$ be a normally distributed random variable. The relative jump size is given by $1 + \mu = \mathbb{E}[\exp(U_{\tau^*})] = \exp(\mu_J + 0.5\sigma^2_J)$, with the change in the spot FX rate after a jump at τ^* is given by $S_{ij}(\tau^+) = S_{ij}(\tau^-)\exp(U_{\tau^*})$, where τ^- is the instant prior to the jump, τ^+ is the instant after the jump. Besides parameters in φ_q, the rest of the parameters are collected as $\vartheta^{MKT}_{ij} = (\eta, \mu, \lambda, \mu^*, \kappa_h, \theta_h, \sigma_h)'_{ij}$.

I now set out an arbitrage free economy, whereby the volatility and jump processes are risk neutralized, such that nominal bond prices denoted in each numéraires are semi-martingales in frictionless exchange. Under the martingale measure Q, I will assume that both the domestic and foreign exchange rate $q_i(t)$ and $q_j(t)$ have some joint distribution under Q and the data generating measure P and hence clear the market (this is a standard assumption).

Hence the jump and volatility terms are compensated, such that a synthetic portfolio of floating rate accounts borrowing continuously at $q_i(t)$ and financing a position in $q_j(t)$, have an expected payoff identical to $S_{ij}(t)$, with risk neutral variance identical to the

observed physical variance generated from the law of motion defined in (Eq. (2.6)) and (Eq. (2.7)). Deriving the law of motion for the equivalent risk neutral dynamics is known as risk neutralization. The standard approach for this type of model is given in Pan [2002] but several others can generate the same result, with and without jumps. See Bates [1996] for another example for equity indices.

The pure-jump process $Z_Q(t)$ distribution is identical to the data generating measure. The dynamics of $(S_{ij}(t), V_{ij}(t))$ are as follows:

$$\begin{aligned} \mathrm{d}S_{ij}(t) =&[q_i(t) - q_j(t)]S_{ij}(t)\mathrm{d}t + \sqrt{V_{ij}(t)}S_{ij}(t)\mathrm{d}W_Q^{ij} + \\ &\mathrm{d}Z_Q(t) - \mu^* S_{ij}(t)\lambda V_{ij}(t)\mathrm{d}t \end{aligned} \tag{2.8}$$

where spot variance is given by:

$$\mathrm{d}V_{ij}(t) = [\kappa_h(\bar{\theta} - V_{ij}(t)) + \eta^\theta V_{ij}(t)]\mathrm{d}t + \sigma_h\sqrt{V_{ij}(t)}(\rho_{ij}\mathrm{d}W_Q^h) + \sqrt{1-\rho_{ij}^2}\mathrm{d}W_Q^{ij}, \tag{2.9}$$

for a one-dimensional volatility process, such that $V_{ij}(t)$ is $t : \mathbb{R}_+ \to \mathbb{R}_+$. In this set up ρ_{ij} to capture the correlation between spot price $S_{ij}(t)$ and the spot continuous variance $V_{ij}(t)$ of the exchange rate. Besides parameters in φ_q, I denote the rest of the parameters as $\vartheta_{ij}^Q = (\lambda, \mu, \rho, \kappa_h, \bar{\theta}, \sigma_h, \mu^*)'_{ij}$. To simplify the notation, in the following sections I omit the dependency of S_{ij} on φ_q.

The solution of a spot rate characteristic function is quite tractable and simply aggregates the characteristic functions for each of the processes driving the specific terms of the stochastic differential equation.Starting from:

$$\Psi(c, V_{ij}(t), q_i, q_j, T-t, \vartheta_{ij}^Q) = \mathbb{E}_t^Q[e^{-\int_t^T q_s \mathrm{d}s} - e^{c\ln(S_T)}] \tag{2.10}$$

the characteristic function can be written as an exponential-affine function:

$$\Psi(c, B_{ij}(t), T-t, \vartheta_{ij}^{Q}) = \exp(\alpha + \beta B_{ij}(t)) \tag{2.11}$$

where $\beta = (\beta_x, \beta_{q_j}, \beta_{q_i}, \beta_{\theta_h})'$ and imposing q_j as the foreign currency interest rate and q_i denote domestic currency interest rate. The β is coefficient set, such that as $\beta_c = c$ the remaining CIR type components of the characteristic function for β is the sum of the individual stochastic factors of the

$$\beta_{q_j} = \frac{2c(1 - e^{-\gamma_{q_j}(T-t)})}{2\gamma_{q_j} - (\gamma_{q_j} - \kappa_{q_j})(1 - e^{-\gamma_{q_j}(T-t)})} \tag{2.12}$$

$$\beta_{q_i} = \frac{2(1-c)(1 - e^{-\gamma_{q_i}(T-t)})}{2\gamma_{q_i} - (\gamma_{q_i} - \kappa_{q_i})(1 - e^{-\gamma_{q_i}(T-t)})} \tag{2.13}$$

$$\beta_{\theta_h} = \frac{c^2(1 - e^{-\gamma_{\theta_h}(T-t)})}{2\gamma_{\theta_h} - (\gamma_{\theta_h} - \kappa_{\theta_h})(1 - e^{-\gamma_{\theta_h}(T-t)})} \tag{2.14}$$

and α is defined as:

$$\begin{aligned}
\alpha = &- \frac{\kappa_{q_d}\bar{q_d}}{\sigma_{q_d}^2}((\gamma_{q_d} - \kappa_{q_d})(T-t) + 2\ln(1 - \frac{\gamma_{q_d} - \kappa_{q_d}}{2\gamma_{q_d}}(1 - e^{-\gamma_{q_d}(T-t)})))+ \\
&- \frac{\kappa_{q_f}\bar{q_f}}{\sigma_{q_f}^2}((\gamma_{q_f} - \kappa_{q_f})(T-t) + 2\ln(1 - \frac{\gamma_{q_f} - \kappa_{q_f}}{2\gamma_{q_f}}(1 - e^{-\gamma_{q_f}(T-t)})))+ \\
&- \frac{\kappa_{\theta}\bar{\theta}}{\sigma_{\theta}^2}((\gamma_{\theta} - \kappa_{\theta})(T-t) + 2\ln(1 - \frac{\gamma_{\theta} - \kappa_{\theta}}{2\gamma_{\theta}}(1 - e^{-\gamma_{\theta}(T-t)})))+ \\
&\lambda_s(T-t)(\exp(c\mu_J) + \frac{c^2\sigma_J^2}{2}) - 1)
\end{aligned} \tag{2.15}$$

where $\gamma_{q_d} = \sqrt{\kappa_{q_d}^2 + 2\sigma_{q_d}^2 c}$, $\gamma_{q_f} = \sqrt{\kappa_{q_f}^2 + 2\sigma_{q_f}^2(1-c)}$, and $\gamma_{\theta} = \sqrt{\kappa_{\theta}^2 - c^2\sigma_{\theta}^2}$. For a given strike $K_{ij}(t)$, let $k_{ij}(t) = K_{ij}(t)/S_{ij}(t, \vartheta^Q)$ as the moneyness of currency S_{ij} at time t, the FX call option price can be shown as:

$$C(c, B_{ij}(t), T-t, k_{ij}(t), \vartheta_{ij}^{Q}) = P_{ij}^{1}(c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}) - k_{ij}(t)P_{ij}^{2}(c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}) \tag{2.16}$$

where

$$P_{ij}^{1}(c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}) = \frac{\Psi(1, B_{ij}(t), t, T, \vartheta_{ij}^{Q})}{2} - \frac{1}{\pi}\int_{0}^{\infty} \frac{\mathrm{Im}(\Psi(1-c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}))e^{c\ln(k_t)}}{c}\mathrm{d}c \tag{2.17}$$

$$P_{ij}^{2}(c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}) = \frac{\Psi(0, B_{ij}(t), t, T, \vartheta_{ij}^{Q})}{2} - \frac{1}{\pi}\int_{0}^{\infty} \frac{\mathrm{Im}(\Psi(0-c, B_{ij}(t), t, T, \vartheta_{ij}^{Q}))e^{c\ln(k_t)}}{c}\mathrm{d}c \tag{2.18}$$

where $\mathfrak{s}$ is the imaginary part of the complex number. Combining (2.12)-(2.16), we can see that there is a volatility skew relative to moneyness ratio $k_{ij}(t)$ under risk neutral measure with affine jump-diffusion model, when the velocity of mean reversion in the volatility equation is small and when the mean jump size is large.

2.4.1 The GARCH Model of Correlated FX Returns

I will use the GARCH model as a benchmark in later sections to identify the improvement made by the eigen model. Setting $R(t+\eta) = ln(S(t+\eta)/S(t))$,our GARCH model begins with the GARCH type volatility:

$$R(t+\eta) = r_d(t) - r_f(t) + \lambda_s h_s(t+\eta) + \sqrt{h_s(t+\eta)}z(t+\eta), \tag{2.19}$$

$$h_s(t+\eta) = \omega_s + \beta_s h_s(t) + \alpha_s(z(t) - \gamma_s\sqrt{h_s(t)})^2, \tag{2.20}$$

where $h_s(t)$ presents the variance process in spot and $z(t)$ ^{iid}, η. It is worth mentioning that the unconditional expectation $\mathbb{E}[h_s(t+\eta)] = (\alpha_s + \omega_s)/(1 - \beta_s - \alpha_s\gamma_s^2)$. For other parameters, β_s determines the persistence in corresponding spot variance; α_s and γ_s are the principal parameters that determine the kurtosis and skewness in implied volatility surface. λ_s determines the feedback in spot variance to the mean return.

Because FX spot markets involve investors holding derivatives based on different cur-

rencies and different measurements of the probability of the asset price movements, it is necessary to derive the risk-neutral formula for the FX market to avoid arbitrage for domestic and foreign investors. From a domestic investor's point of view, the risk-neutral FX process is given by :

$$R(t+\eta) = r_d(t) - r_f(t) - 1/2 h_s(t+\eta) + \sqrt{h_s(t+\eta)} z^*(t+\eta), \tag{2.21}$$

$$h_s(t+\eta) = \omega_s + \beta_s h_s(t) + a_s(z^*(t) - \gamma_s^* \sqrt{h_s(t)})^2, \tag{2.22}$$

with risk neutral innovations $z^* = z(t) + (\lambda_s + 1/2)\sqrt{h_s(t)}$ and parameter $\gamma_s^* = \gamma_s + \lambda_s + 1/2$.

Denoting $x(t) = lnS(t)$ as the natural logarithm of the spot FX price at time t. Let$f^*(\phi)$ denote the conditional generating function of logarithm price under domestic risk neutral measure and assume the MGF(moment generating function) takes the log-linear GARCH(1,1) form, then the MGF is in the following form:

$$\begin{aligned} f(\phi) =& \mathbb{E}^*[e^{\phi x(T)}] = exp(\phi x(t) + A(t;T,\phi) + B_d(t;T,\phi) r_d(t) - B_f(t;T,\phi) r_f(t) + \\ &+ C_d(t;T,\phi) h_d(t+\eta) - C_f(t;T,\phi) h_f(t+\eta) + C_s(t;T,\phi) h_s(t+\eta)). \end{aligned} \tag{2.23}$$

where d,f and s denote the parameters for domestic risk-free rate, foreign risk-free rate and spot price respectively. The recursive terms can be evaluated as follows:

$$\begin{aligned} A(t;T,\phi) =& A(t+\eta;T,\phi) + B_d(t+\eta;T,\phi)\mu_{0d} - B_f(t+\eta;T,\phi)\mu_{0f} + \\ &+ C_d(t+\eta;T,\phi)\omega_d - C_f(t+\eta;T,\phi)\omega_f + C_s(t+\eta;T,\phi)\omega_s + \\ &- \frac{1}{2} ln(1 - 2C_s(t+\eta;T,\phi)\alpha_s) - \frac{1}{2} ln(1 - 2C_d(t+\eta;T,\phi)\alpha_d) + \\ &+ \frac{1}{2} ln(1 - 2C_f(t+1;T,\phi)\alpha_f) \end{aligned} \tag{2.24}$$

$$B_k(t;T,\phi) = B_k(t+\eta;T,\phi)\mu_{1k} + \phi, k \in f, d, \tag{2.25}$$

$$C_k(t;T,\phi) = C_k(t+\eta;T,\phi)\beta_k + B_k(t+\eta;T,\phi)\lambda_k + \gamma_k B_k(t+\eta;T,\phi) + \\ -\frac{1}{2}\gamma_k^2 + \frac{(B_k(t+\eta;T,\phi)-\gamma_k)^2}{2(1-2C_k(t+\eta;T,\phi)\alpha_k)}, k \in f,d \tag{2.26}$$

$$C_s(t;T,\phi) = C_s(t+\eta;T,\phi)\beta_s - \frac{1}{2}\phi + \gamma_s^*\phi - \frac{1}{2}\gamma_s^{*2} + \frac{(\phi-\gamma_s^*)^2}{2(1-2C_s(t+\eta;T,\phi)\alpha_s)} \tag{2.27}$$

With the known standard boundary conditions for domestic/foreign spot deposit rate, I could derive the MGF and then the option price:

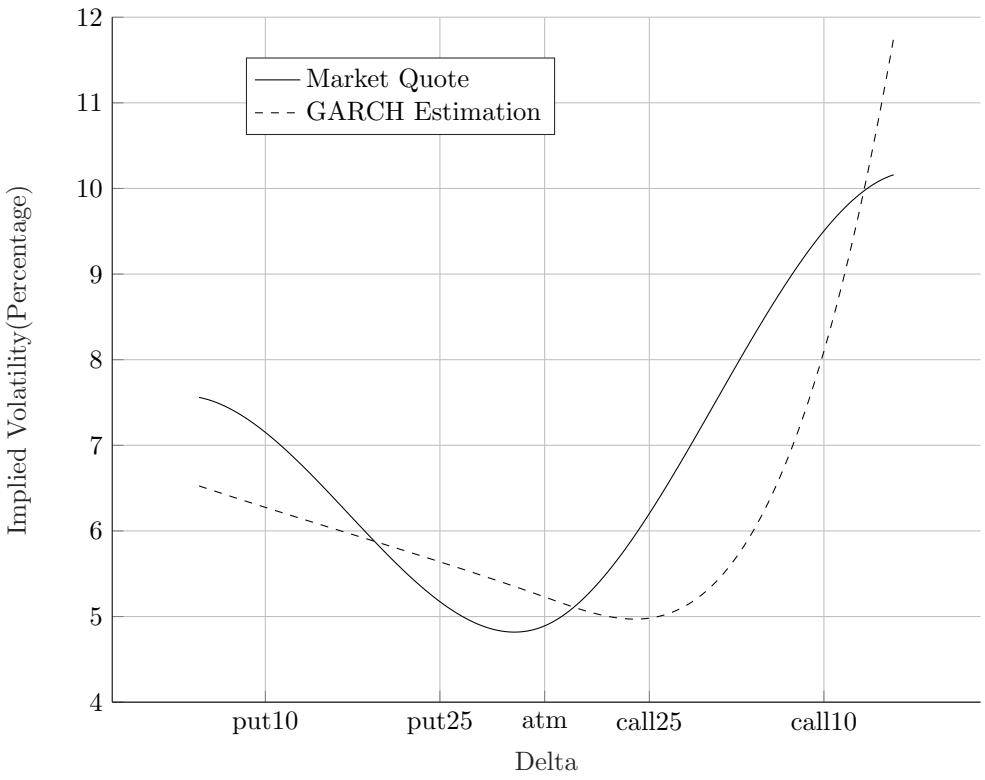
$$\mathbb{E}_t^*[max(S(T)-K),0] = f^*(1)(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathbb{R}[\frac{K^{-i\phi}f^*(i\phi+1)}{i\phi f^*(1)}]d\phi) + \\ -K(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathbb{R}[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}]d\phi). \tag{2.28}$$

Denoting $D_d(\tau)$ as the domestic discounting factor, and from $\tau = T - t$ then I can write the value of option as the present value today of expectation. The call option price for an FX option under risk neutral measure is given by the following:

$$\begin{aligned} P_c =& D_d(\tau)\mathbb{E}_t^*[max(S(T)-K),0] \\ =& D_d(\tau)f^*(1)(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathbb{R}[\frac{K^{-i\phi}f^*(i\phi+1)}{i\phi f^*(1)}]d\phi) + \\ & -K(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathbb{R}[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}]d\phi). \end{aligned} \tag{2.29}$$

Fig. 2.1 shows the comparison between market quote and GARCH estimation for volatility surface based on USDSAR 1 year maturity option. I can conclude from the figure that the error between estimation and market quote is dependent on delta. For at-the-money, the GARCH type affine model could provide a relatively reliable estimation

Figure 2.1: Market Quote versus GARCH Estimation



Note: This figure presents the USDSAR 1Y volatility surface from market quote and GARCH estimation on MAY 1,2018 respectively. GARCH parameters are estimated from spot process back to 2000, reported by TRTH.

for BS implied volatility with error less than 0.5. But for deep out-of-money contracts (put10 and call10), the error is increased to a level higher than 1. Market participants set threshold 1 as the criterion for prediction, so the GARCH type model is believed not to be reliable enough for these on-wings contracts. In our eigenvalue model based on leg journey, I will show the consistency of predictive power across volatility surface.

2.4.2 FX Leg Journeys

Let k be another numéraire with exchange rates $S_{ik}(t, \vartheta_{ik}^Q)$ and $S_{kj}(t, \vartheta_{kj}^Q)$ respectively. Triangular arbitrage imposes that for some future date $T > t$, $\mathbb{E}_t[S_{ik}(T, \vartheta_{ik}^Q)S_{kj}(T, \vartheta_{kj}^Q)] =$

$\mathbb{E}_t[S_{ij}(T, \vartheta^Q_{ij})]$. I can extend the triangular arbitrage arrangement to further legs, whereby if k_m for $m \in \{1, \ldots, M\}$ is a list of currency pairs then,

$$\mathbb{E}_t\left[\left(S_{ik_1}(T, \vartheta^Q_{ik_1})\prod_{n=1}^{M-1} S_{k_n k_{n+1}}(T, \vartheta^Q_{k_n k_{n+1}})S_{k_M j}(T, \vartheta^Q_{k_M j})\right)\left(S_{ij}(T, \vartheta^Q_{ij})\right)^{-1}\right] = 1 \quad (2.30)$$

That is, the time t expectation will be that at time T an instant journey over M leg currencies from currency i to j, should be equivalent to a direct journey from i to j of a single unit of i. Furthermore, let $R_{ij}(t, T, \vartheta^Q_{ij}) = \log S_{ij}(T, \vartheta^Q_{ij}) - \log S_{ij}(t, \vartheta^Q_{ij})$, be the natural logarithm of the difference in exchange rates from time t to T. Rewriting (2.30) I can define the following expected return on a multi-leg position instigated at at T as follows:

$$\mathbb{E}_t\left[R_{ik_1}(t, T, \vartheta^Q_{ik_1}) + \sum_{n=1}^{M-1} R_{k_n k_{n+1} j}(t, T, \vartheta^Q_{k_n k_{n+1}}) + R_{k_M j}(t, T, \vartheta^Q_{k_M j}) - R_{ij}(t, T, \vartheta^Q_{ij})\right] = 0 \quad (2.31)$$

$$\mathbb{E}_t\left[R_{ik_1}(t, T, \vartheta^Q_{ik_1}) + \sum_{n=1}^{M-1} R_{k_n k_{n+1} j}(t, T, \vartheta^Q_{k_n k_{n+1}}) + R_{k_M j}(t, T, \vartheta^Q_{k_M j})\right] = \mathbb{E}_t\left[R_{ij}(t, T, \vartheta^Q_{ij})\right] \quad (2.32)$$

Similarly, let $\text{var}_t\left[R_{ik_1}(t, T, \vartheta^Q_{ik_1})\right] = \mathbb{E}_t\left[(R_{ik_1}(t, T, \vartheta^Q_{ik_1}) - \mathbb{E}_t\left[R_{ik_1}(t, T, \vartheta^Q_{ik_1})\right])^2\right]$, be the expected variance of of a position instigated at T, the variance from directly exchanging numéraire i to numéraire j or transitioning from i to j via M intermediates should be the same, hence setting:

$$\text{cov}_t\left[R_{ik}(t, T, \vartheta^Q_{ik}), R_{kj}(t, T, \vartheta^Q_{kj})\right] = -\tilde{r}_{ik,kj}\sqrt{\text{var}_t\left[R_{ik}(t, T, \vartheta^Q_{ik})\right]\text{var}_t\left[R_{kj}(t, T, \vartheta^Q_{kj})\right]} \quad (2.33)$$

finally I can define a positive definite matrix of terminal covariation from t to T,

$$\tilde{\boldsymbol{C}}(t,T,\vartheta^{Q}) = \begin{pmatrix} \mathrm{var}_t\left[R_{ik_1}(t,T,\vartheta^{Q}_{ik_1})\right] & \dots & \mathrm{cov}_t\left[R_{ik_1}(t,T,\vartheta^{Q}_{ik_1}), R_{k_Mj}(t,T,\vartheta^{Q}_{k_Mj})\right] \\ \vdots & \ddots & \vdots \\ \mathrm{cov}_t\left[R_{ik_1}(t,T,\vartheta^{Q}_{ik_1}), R_{k_Mj}(t,T,\vartheta^{Q}_{k_Mj})\right] & \dots & \mathrm{var}_t\left[R_{k_Mj}(t,T,\vartheta^{Q}_{k_Mj})\right] \end{pmatrix} \tag{2.34}$$

for a given pair of numéraires i and j, the variance of the direct exchange rate will be such that $\mathrm{var}_t\left[R_{ij}(t,T,\vartheta^{Q}_{ij})\right] = \mathbf{1}'\tilde{\boldsymbol{C}}(t,T,\vartheta^{Q})\mathbf{1}$, where $\mathbf{1}$ is an $M+1$ length unit vector. For the tenor $T-t$, for an option portfolio to be consistent, in a mean-variance framework, hence requires the volatility under the risk neutral measure for the portfolio of direct and cross exchanges rates must be consistent.

Consider a currency list k_m for $m \in \{1,\dots,M\}$ priced by currency i under neutral measure (Q). For any $n \in m$ the dynamics of currency pairs are shown as :

$$\begin{aligned} \mathrm{d}S_{ik_n}(t,\vartheta^{Q}_{ik_n}) =& [q_i - q_{k_n}]S_{ik_n}(t,\vartheta^{Q}_{ik_n})\mathrm{d}t + \sqrt{V_{ik_n}}S_{ik_n}(t,\vartheta^{Q}_{ik_n})\mathrm{d}W^{ik_n}(t) + \\ & \mathrm{d}Z_Q(t) - \mu^* S_{ik_n}(t,\vartheta^{Q}_{ik_n})\lambda V_{ik_n}\mathrm{d}t \end{aligned} \tag{2.35}$$

With the moment generating function shown of time conditional-t transform of R_T in equation (7), the expectation of currency rate is:

$$\mu_{ik_m}(t,T,\vartheta^{Q}_{ik_m}) = R_{ik_m}(T-t)e^{0\cdot R_{ik_m}} \quad . \tag{2.36}$$

Following (2.30), the multi-leg expected return from t to T can be verified to be constant:

$$\mathbb{E}_t[\mu_{ik_1}(t,T,\vartheta^{Q}_{ik_1}) + \sum_{n=1}^{M-1}\mu_{ik_n}(t,T,\vartheta^{Q}_{ik_n}) + \mu_{k_Mj}(t,T,\vartheta^{Q}_{k_Mj})] = \mu_{ij}(t,T,\vartheta^{Q}_{ij}) \tag{2.37}$$

the variances of currency rates are shown as:

$$\mathrm{var}_t\left[R_{ik_1}(t,T,\vartheta^{Q}_{ik_1})\right] = \mathrm{Re}[\Psi''(0;V_{ik_1},q_i,q_{k_1},T-t,\vartheta^{Q}_{ik_1})]. \tag{2.38}$$

Let $\rho_{ij,jk}(t,T)$ denote the correlation between S_{ij} and S_{jk} for period t,T,$M \geq 3$, we can see that the variance will be:

$$\begin{aligned}\mathrm{var}_t\left[R_{k_1k_M}(t,T,\vartheta^Q_{k_1k_M})\right] = &\sum_{n=1}^{M-1}\mathrm{var}_t\left[R_{k_nk_{n+1}}(t,T,\vartheta^Q_{k_nk_{n+1}})\right] + \\ &2\sum_{n=2}^{M-1}\sum_{j<n}\rho_{k_nk_{n+1},k_jk_{j+1}}(t,T)\times \\ &\left(\mathrm{var}_t\left[R_{k_nk_{n+1}}(t,T,\vartheta^Q_{k_nk_{n+1}})\right]\mathrm{var}_t\left[R_{k_jk_{j+1}}(t,T,\vartheta^Q_{k_jk_{j+1}})\right]\right)^{1/2},\end{aligned} \tag{2.39}$$

with mean and volatility shown to be consistent over time range $[t,T]$, the currency exchange is verified to be consistent in the same tenor.

2.4.3 Black-Scholes Implied volatility Surface and Correlation Surface

The Black-Scholes model makes the assumption that the option price is related to underlying asset price, underlying volatility, strike price, time to maturity and interest rate. The advantage of the Black-Scholes model is that the parameters are all easy to get from market except for volatility. The Black-Scholes model assumes that volatility is constant through the life of the option, which is obviously not true for the real market. For this reason participants raise the idea of implied volatility: the volatility value makes the Black-Scholes model calculation result equal the real market option price: $C^{MKT}(t,\vartheta^Q) = C^{BS}(t)$.

Let $r_i(t,T)$ and $r_j(t,T)$ denote the returns on a riskless timed deposit account in numéraires i and j respectively over t to T. Hence $r_i(t,T) = \mathbb{E}_t[\int_t^T q_i(t)\mathrm{d}t]$. The Garman and Kohlhagen [1983] modified Black-Scholes price for foreign exchange call option price

is given as follows:

$$C_{ij}^{BS}(t,T) = S_{ij}(t,\vartheta_{ij}^{Q})e^{-r_j(T-t)}\tfrac{1}{2}(1+\mathrm{Erf}(\tfrac{\ln(S_{ij}(t,\vartheta_{ij}^{Q})/K_{ij})+(r_i-r_j+\sigma^{BS}/2)(T-t)}{\sigma^{BS}\sqrt{T}}))+ \\ -K_{ij}e^{-r_i(T-t)}\tfrac{1}{2}(1+\mathrm{Erf}(\tfrac{\ln(S_{ij}(t,\vartheta_{ij}^{Q})/K_{ij})-(r_i-r_j+\sigma^{BS}/2)(T-t)}{\sigma^{BS}\sqrt{T-t}})) \tag{2.40}$$

where $\mathrm{Erf}(\cdot)$ is the standard error function. It is useful to think of this as a pricing transform rather than having some fundamental connection to the underlying statistical properties of the asset price process. Indeed, the price quotes in the OTC-FXO tend to be delivered in terms of the Black-Scholes volatility and the Black-Scholes delta. The curiosity is that the Black-Scholes delta is a function of volatility and strike, whereas the volatility is treated as an exogenous function of strike. Set

$$\Delta^{BS}(\sigma^{BS},K_{ij},t,T,r_i,r_j) = 1+\mathrm{Erf}(\tfrac{\ln(S_{ij}(t,\vartheta_{ij}^{Q})/K_{ij})-(r_i-r_j+\sigma^{BS}/2)(T-t)}{\sigma^{BS}\sqrt{T-t}}) \tag{2.41}$$

Notice that I specify this in error function terms to illustrate that this is a functional decomposition of the call price and *not* an equilibrium model. The simplest way to think about the FXO quoting convention is that the Black-Scholes delta is a way of stating a proportional strike price and the volatility is an actual price.

As such, for a European volatility surface, market participants tend to use the implied volatility corresponding to Δ^{BS} instead of to strike price K_{ij}, which is solved for at a later point in the pricing calculation. With the market price shown by (Eq. (2.16)), implied volatility can be derived from the equation:

$$C^{MKT}(c,B_{ij}(t,T),t,T,\vartheta^{Q},K) = C_{ij}^{BS}(t,T,\Delta^{BS}) \tag{2.42}$$

The time t is considered as 0, which indicates that for both real market and Black-Scholes model I consider the process between $[0,T]$. (Eq. (2.40)) leads to the 'Black-Scholes

variance' $(\sigma^{BS}(t,T,\Delta^{BS}))^2$. Let $\text{cov}^{BS}_{ij,ik}(t,T,\Delta^{BS})$ denotes the covariance of spot pair ij and ik from Black-Scholes at period $[t,T]$ and Delta Δ^{BS}, combined (Eq. (2.33)) with the result shown by :

$$\text{cov}^{BS}_{ij,ik}(t,T,\Delta^{BS}) = r^{BS}_{ij,ik}(t)\sqrt{(\sigma^{BS}_{ij}(t,T,\Delta^{BS}))^2 \cdot (\sigma^{BS}_{ik}(t,T,\Delta^{BS}))^2}. \tag{2.43}$$

The covariance matrix $\boldsymbol{C}$ then is shown as:

$$\boldsymbol{C}(t,T,\Delta^{BS}) = \begin{pmatrix} (\sigma^{BS}_{ik_1}(t,T,\Delta^{BS}))^2 & \dots & \text{cov}^{BS}_{ik_1,k_Mj}(t,T,\Delta^{BS}) \\ \vdots & \ddots & \vdots \\ \text{cov}^{BS}_{ik_1,k_Mj}(t,T,\Delta^{BS}) & \dots & (\sigma^{BS}_{k_Mj}(t,T,\Delta^{BS}))^2 \end{pmatrix} \tag{2.44}$$

Let $\boldsymbol{R}(t,T,\Delta^{BS})$ denote the equivalent implied correlation matrix for Delta Δ^{BS} for the tenor $T-t$, this matrix is computed using the conventional normalization $\boldsymbol{R}(t,T,\Delta^{BS}) = \text{diag}[\text{diag}[\boldsymbol{C}]^{-1/2}] \cdot \boldsymbol{C} \cdot \text{diag}[\text{diag}[\boldsymbol{C}]^{-1/2}]$ with the following structure:

$$\boldsymbol{R}(t,T,\Delta^{BS}) = \begin{pmatrix} \rho^{BS}_{ik_1,ik_1}(t,T,\Delta^{BS}) & \dots & \rho^{BS}_{ik_1,k_Mj}(t,T,\Delta^{BS}) \\ \vdots & \ddots & \vdots \\ \rho^{BS}_{ik_1,k_Mj}(t,T,\Delta^{BS}) & \dots & \rho^{BS}_{k_Mj,k_Mj}(t,T,\Delta^{BS}) \end{pmatrix} \tag{2.45}$$

Unfortunately, the European volatility surface is not normally quoted directly in the FXO market, with participants preferring to use quotes from strategies that capture the structure of the implied volatility surface.. The frequently used strategy pairs are : $25\Delta^{BS}$ and $10\Delta^{BS}$ risk reversal (RR) and butterfly (BF). Let $\sigma^{BS}_{ATM}(t,T)$ denote the implied volatility of at-the-money Delta's volatility, and $\sigma^{BS}_{C}(t,T,\Delta^{BS})$ and $\sigma^{BS}_{P}(t,T,\Delta^{BS})$ denote the corresponding Delta's Black-Scholes call and put options' implied volatility respectively.

Then the volatilities of risk reversal and 2-vol butterfly can be expressed as:

$$\sigma^{BS}_{RR}(t,T,\Delta^{BS}) = \sigma^{BS}_{C}(t,T,\Delta^{BS}) - \sigma^{BS}_{P}(t,T,\Delta^{BS}) \tag{2.46}$$

$$\sigma^{BS}_{2volBF}(t,T,\Delta^{BS}) = \frac{\sigma^{BS}_{C}(t,T,\Delta^{BS}) + \sigma^{BS}_{P}(t,T,\Delta^{BS})}{2} - \sigma^{BS}_{ATM}(t,T) \tag{2.47}$$

Traders usually use 1-vol butterfly (market strangle) instead of 2-vol butterfly. The process of deriving 1-vol butterfly is firstly solving the following equation:

$$\begin{aligned} C^{BS}(K^{C}_{\Delta^{BS}}, \sigma^{BS}_{1volSTG}) + P^{BS}(K^{P}_{\Delta^{BS}}, \sigma^{BS}_{1volSTG}) = \\ C^{BS}(K^{C}_{\Delta^{BS}}, \sigma^{BS}(K^{C}_{\Delta^{BS}})) + P^{BS}(K^{P}_{\Delta^{BS}}, \sigma^{BS}(K^{P}_{\Delta^{BS}})) \end{aligned} \tag{2.48}$$

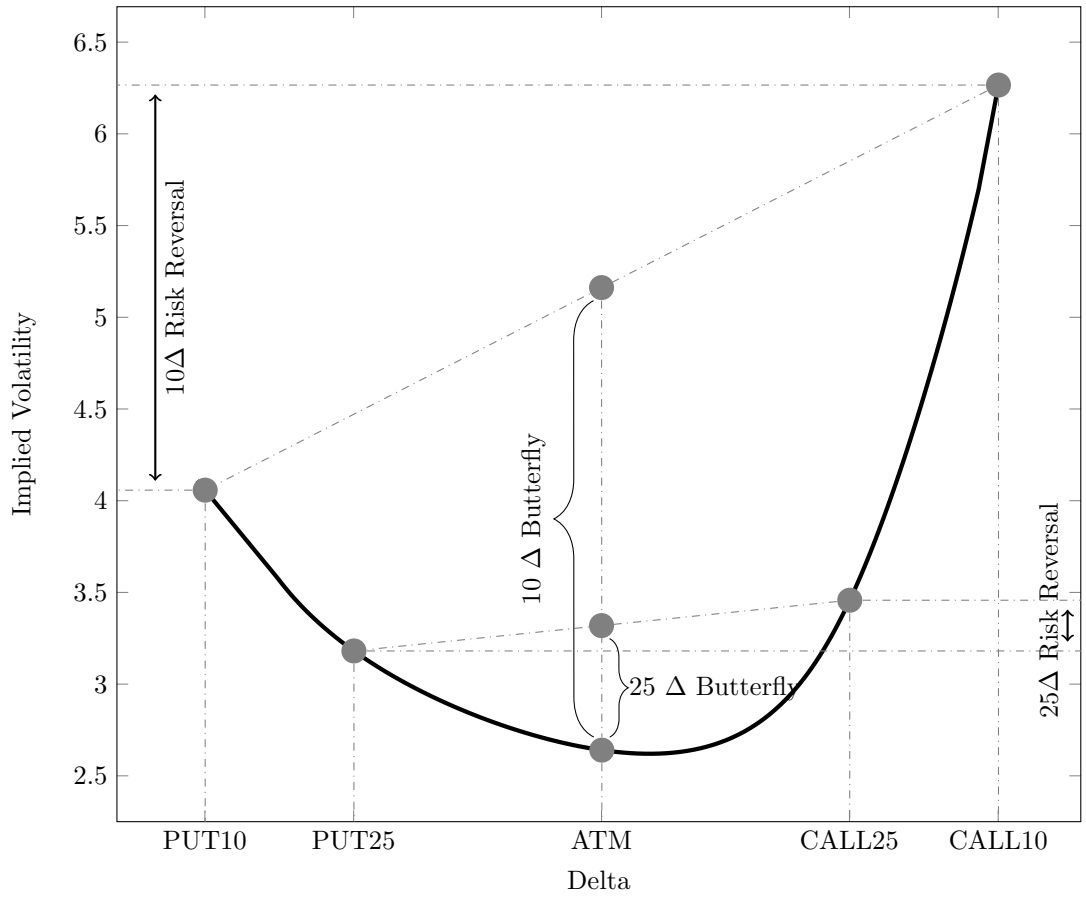
Where $K^{C}_{\Delta^{BS}}$ and $K^{P}_{\Delta^{BS}}$ denote call and put strike price at Delta Δ^{BS} respectively. Solving (2.48) to recover the $\sigma^{BS}_{1volSTG}$ that solves for the equivalent strangle position to the 2-vol butterlfy. Hence, the 1-vol butterfly for Delta Δ^{BS},denoted by $\sigma^{BS}_{1volBF}(t,T,\Delta^{BS}) = \sigma^{BS}_{1volSTG} - \sigma^{BS}_{ATM}$.

The Δ^{BS} is normally set to be $10\Delta^{BS}$ and $25\Delta^{BS}$ for quoting. A risk reversal consists of a long position in call option and a short position in put option for certain Delta. The butterfly consists of several trading options: two short positions in at-the-money option, one long position for call and put options for a certain Delta. Figure 2.2 illustrates the 1 year maturity USDSAR volatility surface (Black-Scholes implied volatility on the ordinate axis) with respect to Delta (the abscissa values) as an example, the $10\Delta^{BS}$ and $25\Delta^{BS}$ risk reversals and 2-vol butterflies are sketched against the surface for illustration.

2.4.4 The Impact of Jumps

Stochastic volatility and jumps have a substantial impact on the structure of the European implied volatility surface. Fig. 2.3 takes the example of the calibrated USDSAR volatility surface and adjusts two key parameters, ceteris paribus. The black curve is the baseline for

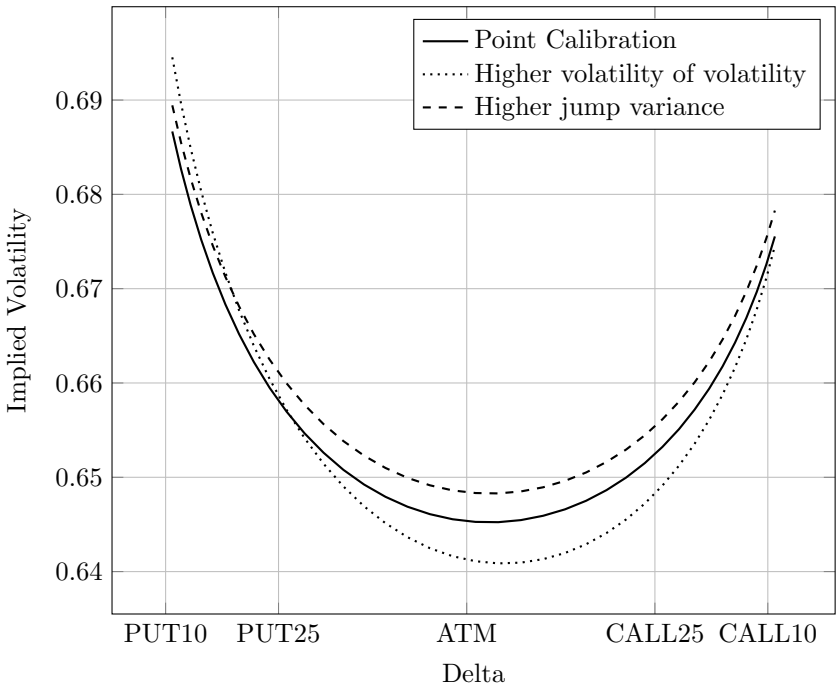
Figure 2.2: Strategies in Option Volatility Surface



Note: This figure presents widely used trading strategy(Risk Reversal and Butterfly) from implied volatility surface. Most of contracts are traded in these 5 given strategies.

the calibrated model. First, I increase the volatility of volatility by 5% (the dashed line) and we can see that this steepens the parts of the curve further from the 'at-the-money' straddle, but decreases the volatility in a neighbourhood near the the at-the-money point. Second, I increase only the jump intensity parameter by 5%; this is the dotted line. Here, we can see that the curve is vertically displaced upward, but the shift is a decreasing ratio as we move from being at the money. This would be in contrast to a 5% shift in long-run continuous volatility which would shift the whole curve by a fixed ratio vertically upwards.

Figure 2.3: 1 Year Maturity USDSAR Option Volatility Surface



Note: This figure presents The impact of jump parameters to estimation volatility surface. Solid line indicate the benchmark result from model, dot line indicate result with higher level variance of variance while keep other parameters fixed. Dash line is the result with higher level of jump variance while other parameters same as benchmark.

2.4.5 Correlation Coefficient Matrix

Constructing the implied volatility surface through method above require accurate predictive value of correlation coefficient $\rho^{BS}_{ij,ik}(t, T, \Delta^{BS})$. In the following sections, for for all the Delta and implied correlation are derived from Black-Scholes model, refer $\rho^{BS}_{ij,ik}(t, T, \Delta^{BS})$ as $\rho_{ij,ik}(t, T, \Delta)$.

As shown by (2.36) and (2.38), the exchange rate S_{ik} is stable compared wth exchange process from S_{ij} andS_{jk}. In practice, imagine a bunch of currencies k_n where $n \in \{1, \ldots, M-1\}$, all denominated by unique currency i. Correlation coefficient is not directly observed on the market because the variance is not observable. The sections above provide implied correlation as the proxy for true correlation coefficient. Using the square of implied volatility as the variance of certain currency pairs at time t and Δ, the

implied correlation in currency triangle can be shown as:

$$\rho_{ik_1,ik_2}(t,T,\Delta) = \frac{(\sigma^{BS}_{k_1k_2}(t,T,\Delta))^2 - (\sigma^{BS}_{ik_1}(t,T,\Delta))^2 - (\sigma^{BS}_{ik_2}(t,T,\Delta))^2}{2\sigma^{BS}_{ik_1}(t,T,\Delta)\sigma^{BS}_{ik_2}(t,T,\Delta)} \tag{2.49}$$

k_1 and k_2 are named as 'denominated currency' while i is the denominating currency in the currency triangle. For the list of denominated currency $k_n : n \in \{1,\ldots,M\}$, the correlation coefficient matrix is then defined as:

$$\boldsymbol{R}(t,T,\Delta) = \begin{pmatrix} \rho_{ik_1,ik_1}(t,T,\Delta) & \ldots & \rho_{ik_1,k_Mj}(t,T,\Delta) \\ \vdots & \ddots & \vdots \\ \rho_{ik_1,k_Mj}(t,T,\Delta) & \ldots & \rho_{k_Mj,k_Mj}(t,T,\Delta) \end{pmatrix}. \tag{2.50}$$

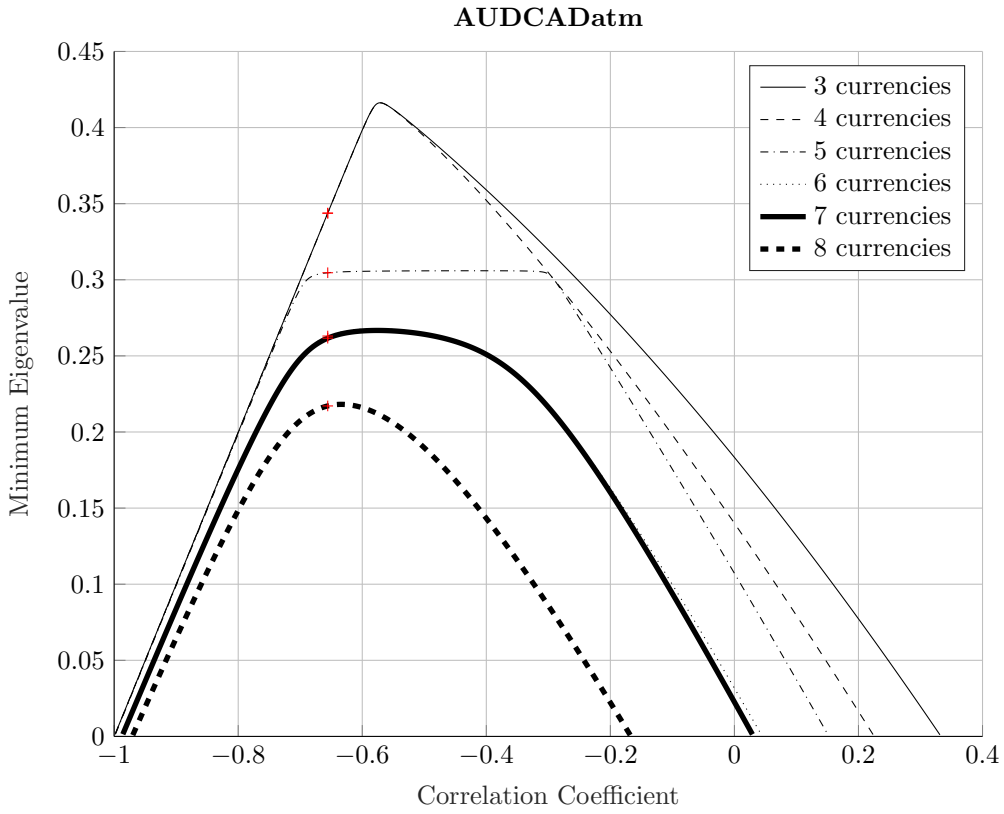
The characterization of a positive definite matrix imposes bounds on the pointwise entry of elements of the correlation matrix to ensure that the eigenvalues of the resulting matrix are always positive. By substituting the missing coefficient with a value from -1 to 1, we can get a temporary correlation $\hat{\boldsymbol{R}}(t,T,\Delta)$. Hence I now have the following structure:

$$\hat{\boldsymbol{R}}(t,T,\Delta) = V(t,T,\Delta)\cdot\Lambda(t,T,\Delta)\cdot V(t,T,\Delta)' \tag{2.51}$$

where $V(t,T,\Delta)$ is $(M\times M)$ matrix, each column is an eigenvector. $\Lambda(t,T,\Delta)$ is $(M\times M)$ diagonal matrix whose diagonal elements are corresponding eigenvalues. Functionally, the bound requires that $\min(\text{diag}[\Lambda(t,T,\Delta)]) > 0$, all substituted values that do not make the matrix meet the condition will be abandoned. The smallest and largest values that make the correlation matrix meet all positive eigenvalues condition are marked as the bounds for the missing correlation coefficient.

The missing correlation coefficient's bounds will become narrower with the increasing numbers of denominated currencies involved (denoted as M). In Fig. 2.4. I randomly select AUDCAD as underlyiny asset and generate the minimum eigenvalues with different

Figure 2.4: Minimum eigenvalue versus market data: AUDCAD atm



Note: This figure presents change of minimum eigen value versus correlation coefficient. With all other implied volatility in leg journey known, ADUCAD atm implied volatility is assumed to be the only unknown element in implied correlation matrix. Red cross is the real market quote as benchmark. Data is reported by TRTH.

numbers of leg currencies. For the plot, I have a red cross to indicate the market implied correlation and corresponding eigenvalue for different leg currency numbers. I can conclude from the figure that with the increment of leg currency amount, the performance of the model is improved in two aspects: 1) the bound provided by making minimum eigenvalue zero is narrowed 2) the difference between maximized minimum eigenvalue and market quote (indicated by red cross) is reduced. I can conclude from the trend that length of leg journey is critical to eigensystem analysis, with 7-10 currency pairs involved, the predictive power of eigen model will be greatly improved. More figures could be find in Appendix.

2.5 Preliminary Data Analysis

For the implied correlation calculation, a currency pair triangle is required with one denominating currency, and two denominated currencies. For the purpose of testing the performance of this method, the currency triangle will consist of pairs for which the option between two denominated currencies is actually trading on the market. With known real market data as a benchmark, the experiment can manually delete a correlation coefficient ρ_{ik_1,ik_2} each time, using the eigenvalue method to invert calculating correlation. For testing, it is necessary to make sure the correlation coefficient has no missing element; the maximum correlation scale currently with the constraint of no missing element is 8×8. The leg currencies included in our analysis are: GBP, USD, EUR, CHF, AUD, JPY, CAD, SEK and NOK.

The correlation coefficient calculation uses 1 month maturity option data from TRTH. The data used in this test is the end-of-day data for 2017. The volatilities for five different deltas are used for currency pair: delta 10 put, delta 25 put, at the money, delta 25 call and delta 10 call.

2.6 Empirical Results

Section 2.6 lays out our main empirical results for the predictive power of the eigen model. I summarize the results and compare the eigen model with the widely used GARCH model. I firstly introduce the GARCH type model used as the benchmark then implement the Diebold-Mariano test to investigate how consistently the eigen model outperforms the GARCH model through maturity, delta and currency pairs.

2.6.1 Data Description

To assess the accuracy of the eigen model across different deltas and maturities, I implement the eigen model based on data from the options market from Jun 2014 to Sep 2018. There are two data sources: TRTH provides the tick history and bloomberg provides the end of day data. My empirical study covers 1725 FX option contracts across deltas and maturities.

For each implied volatility surface, I quote five key points: put/call with 10/25 delta and at-the money price. All deltas are derived from butterfly and risk reversal strategies for corresponding deltas and maturities, the process is recorded in previous section. For term structure, I cover 6 widely-traded maturities: 1M, 2M, 3M, 6M, 1Y and 2Y. Although there are some even longer maturity such as 10Y for JPY, these six are the key indices in term structure.

Bid and ask price are collected separately, then mid price is taken into calculation. Cases where only bid or ask price is available are ignored in my study. All mentioned data in later section of this chapter are mid price of the trade.

To conclude, 1725 time series are used as a data set in our empirical study. 1408 of them(81.62%) are inside the bounds provided by the eigen model, meanwhile 914 of them(52.99%) have less than 1% error with eigen model point forecasts.

Table 2.1: Eigen model forecasting accurate level summary (in percentage)

	Count	Interval Forecast	Point Forecast
		Strategy	
'put10'	345	0.6724	0.4128
'put25'	345	0.8439	0.5054
'atm'	345	0.9438	0.6494
'call25'	345	0.8159	0.5589
'call10'	345	0.8066	0.5185
		Maturity	
'1M'	330	0.7885	0.5516
'1Y'	270	0.5631	0.3528
'2M'	330	0.8784	0.5815
'2Y'	265	0.8559	0.5012
'3M'	265	0.9113	0.5808
'6M'	265	0.8985	0.5908

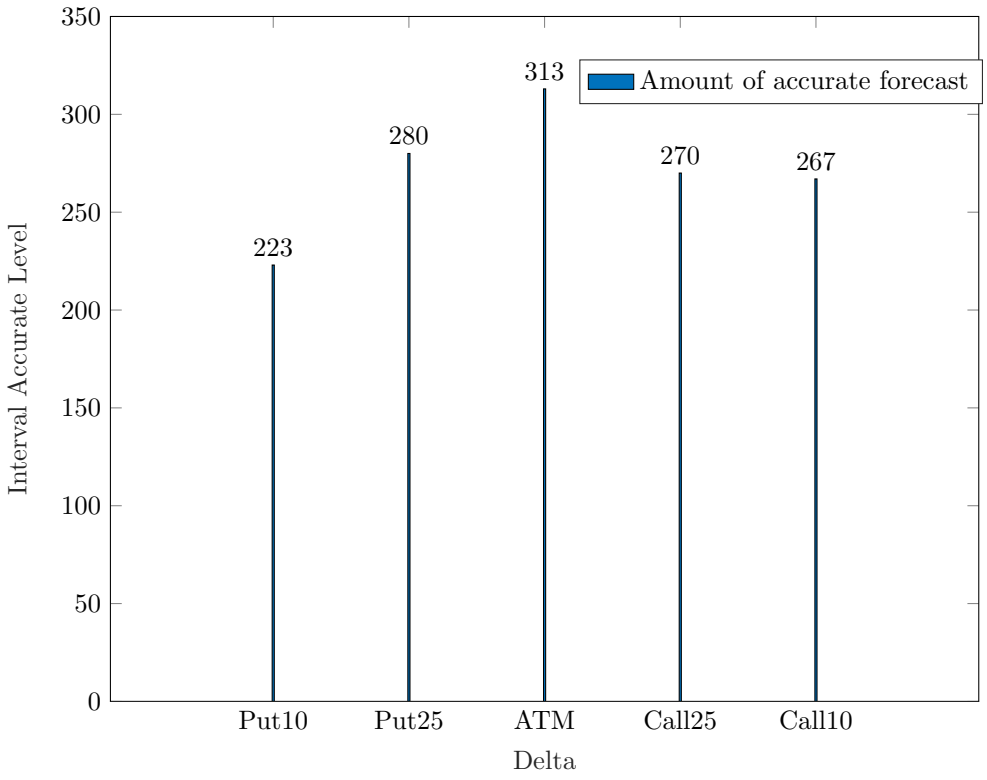
Notes: This table reports the overall eigen model forecasting accurate level. Result are reported delta and maturity respectively. 'Count' column report amount of corresponding delta/maturity. 'Interval Forecast' column report the accuracy of market price fall into eigen model bounds. 'Point Forecast' column report ratio of forecasts where error is less than 1(%).

2.6.2 Eigen Model Prediction Summary

In Table 2.1, I report the eigen model's forecasting summary by delta/maturity. For all five chosen deltas, accuracy of interval forecast range from 67.24% to 94.38%, with at-the-money option forecasting generating the highest accuracy level of 94.38%. With the extension of the Delta to the edge of surface, the forecast accuracy level gradually decreases to 67.24% and 80.66% for put10 and call10 respectively. At-the-money options are believed to be the most rational contracts in options trading; prices of at-the-money options are concentrated in a narrow distribution, which reduces the error made by choosing end-of-day price. Deep away-from-the-money options are not as rational as at-the-money contracts, as prices are largely affected by investors' personal expectations in relation to underlying assets. The wide distribution of the away-from-the-money option price makes both forecasting and comparison inaccurate; results indicate that put option prices are more sensitive to the delta affection.

Fig. 2.5 reports the negative skewness of the eigen model interval forecast where

Figure 2.5: Accuracy level of interval forecast by eigen model



Note: This figure presents the amount of accurate time series. I consider the average difference between model estimation and market quote lower than 1 as 'accurate time series'. All data reported by TRTH with frequency of end of day.

the number of forecasts are all the same at 345. Degree of accuracy for at-the-money interval forecast is 313, and the accuracy declines with delta especially for put option. The accuracy jumps from 280 to 223 when delta changes from 25 to 10 for put option; meanwhile the difference is 3 for call option. At-the-money is the key contract which determines the price level of the option; the eigen model's superior forecast performance provides solid backup for the accuracy of volatility surface forecasts.

Maturity is considered to be the other important factor influencing predictive power. The eigen model generates the best forecasting performance for 3-month maturity, as 91.13% of market prices are inside eigen bounds. The middle term, defined as the range from 2-month to 6-month, shows promising forecasting accuracy of around 90%, but for short-term(1 month) and long-term(1 year) maturity, the eigen model fails to fully capture the market price. For all 270 1-month options across deltas, 152 time series are in the eigen bounds, which is obviously less than other maturities. This deficiency comes mainly from investors' diverse expectations towards the underlying process for both short-term and long-term,especially 1-year, maturity.

In addition to interval forecast summary, Table 2.1 also reports the point forecast accuracy in the third column. In terms of delta, the at-the-money option generates the highest accuracy at 64.94%. Point forecast shows the same trend along volatility surface as interval forecasts, the accuracy decreases with the extending of deltas, with put10 having the lowest accuracy at 41.28%. In terms of maturity, although middle-term still outperforms short-/long- term in our study, there are slight difference compared with interval forecasts. Firstly, 6-month is the most accurate maturity for eigen model point forecasts but not 3-month in interval forecast. Secondly, the point forecast difference between 1-month maturity and middle-term maturity is not as significant as interval forecasts; 1 month maturity presents a relatively better performance in point forecast. The difference between point forecast accuracy degree for maturity is obviously smaller

than it is for interval forecast; point forecast is shown to be less affected than interval forecast in terms of maturity. For 1-month maturity options, point forecast results should carry more weight in forecasting the level of volatility surface.

2.6.3 Diebold-Mariano Test Summary

The eigen model is compared to the widely used GARCH model that is used as a benchmark. The Diebold-Mariano test is used to assess whether the eigen model outperforms the GARCH model in a statistically significant sense through deltas/maturities. The null hypothesis is that the GARCH model outperforms the eigen model.

The Diebold-Mariano(DM) test provides a comparison between two forecasting models. The test statistics DM is asymptotically N(0,1) distributed, so the DM statistic can give the confidence interval of one model outperforming the other. Negative statistic indicates that eigen value model has a higher possibility of outperforming the GARCH model. Three dataset(eigen model, GARCH model, market price) from Oct 2015 to Sep 2018 are used for the test.

In our case, two time series models are available for forecasting the option price. The Diebold-Mariano test provides a comparison between two time series models. Firstly I define the loss function as $e_{it} = \hat{y}_t - y_t, i = 1, 2$. The loss associated with forecast i is assumed to be a function of the forecast error, e_{it} and is denoted by $g(e_{it})$. $g(\dot{})$ is a loss function, that is a function such that $g(e_{it})$ 1)is never negative 2)increases in size as the errors become larger in magnitude. With known loss functions of both models, I denote $d_t = e_{1t} - e_{2t}$, then I have

$$DM = \frac{\bar{d}}{\sqrt{[y_0 + 2\sum_{k=1}^{h-1} y_k]/n}} \tag{2.52}$$

where $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$ and $y_k = \frac{1}{n}\sum_{t=k+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d})$. DM follows a standard normal

Table 2.2: Overall Diebold-Mariano test performance for eigen model and GARCH model (in percentage)

	Count	Eigen	GARCH	Similar
		Strategy		
'p10'	332	0.8313	0.1415	0.0271
'p25'	332	0.8915	0.0813	0.0271
'atm'	332	0.9337	0.0451	0.0210
'c25'	332	0.8765	0.1114	0.0120
'c10'	332	0.8614	0.1234	0.0150
		Maturity		
'1M'	320	0.9250	0.0656	0.0093
'1Y'	260	0.6423	0.3153	0.0423
'2M'	320	0.9406	0.0437	0.0156
'2Y'	250	0.8920	0.0840	0.0240
'3M'	255	0.9254	0.0588	0.0156
'6M'	255	0.9254	0.0549	0.0196

Notes: This table presents the overall Diebold-Mariano test summary for eigen model and GARCH model. 'Count' column report amount of corresponding delta/maturity. 'Eigen' column reports the ratio of eigen model outperforming GARCH model in DM test. 'GARCH' column reports the ratio of GARCH model outperforming eigen model in DM tests. 'Similar' column reports the ratio of two models do not present significant difference. Results are shown by corresponding delta/maturity.

distribution:$DM \in N(0,1)$, so by the DM statistic I can conclude which models perform better in forecasting. A DM statistic lower than 0 indicates that the first model has a higher possibility to outperform the second model.

The GARCH parameters are estimated by spot process from 1997 to 2018. A model is defined as outperforming the other with 95% confidence level. Critical statistics are -1.96 and 1.96 for eigen model and GARCH model respectively in our case. There are 1660 option price time series with different currency pairs, deltas and maturities; the eigen model is verified to outperform the GARCH model in 1459 tests(87.89%). Meanwhile, in 186(11.27%) tests the GARCH model performed better than the eigen model. For the remaining 14 tests(0.84%) there was no significant difference between the two models for corresponding time series.

Table 2.2 reports the DM test summary by listing the ratio of two models outperforming the other by delta and maturity. In the case of at-the-money delta, the eigen

model show the highest ratio with 93.37% (310/332) of eigen model forecasts in this delta verified to perform better than the GARCH model. The effect of away-from-the-money is also shown by the DM test summary, with the forecast accuracy decreasing in the away-from-the-money option. The eigen model shows lowest lead in the put10 option, where the 83.13%(276/332) eigen model forecast outperforms the GARCH model. The DM test difference between put10 option and at-the-money is not as significant as the point forecast comparison, which indicates that the forecasting performance of both models is affected by delta.

Consistently with the forecast accuracy summary, the eigen model has the largest advantage in middle term maturity. In all 320 DM tests for the 2-month maturity option, the eigen model outperformed the GARCH model in 301 cases, which generated the highest outperforming ratio of 94.06%. The worst performance of the eigen model is the 1-year maturity, where the GARCH model outperforming ratio was 31.53%. I can conclude that compared with the sharp decrease in the 1-year maturity predictive power of the eigen model, the GARCH model was as strongly affected. GARCH model forecasting for 1-year maturity is worth taking into consideration for adjusting eigen model forecasting.

In order to assess the graphical affection to two models' performance, I separated the option price process according to continent. Table 2.3 reports the DM test summary by maturity and delta combination for European countries. I also report the minimum and maximum DM statistic to show the confidence level of the test. The European economic area, having the most successful regional integration, provides sufficient liquid FX option trades within the continent where the amount of currency pairs for certain strategy ranges from 12 to 18. 1- and 2-month maturity options are the contracts with the highest level of liquidity; 18 different currency options are traded with each other. For the two liquid maturities, as shown in Table 2.2, the eigen model outperforms the GARCH model for most strategies. Call option with 25 delta generates the best performance for eigen

Table 2.3: Diebold-Mariano test performance for eigen model and GARCH model (Europe)

			Europe			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	18	0.7222	0.2778	0.0000	16.5219	-26.9926
p25_1M	18	0.7222	0.2778	0.0000	12.3349	-26.3993
atm_1M	18	0.8889	0.1111	0.0000	8.9472	-21.8198
c25_1M	18	0.9444	0.0556	0.0000	7.7845	-47.8155
c10_1M	18	0.8889	0.0556	0.0556	10.3045	-52.1315
p10_1Y	13	0.3077	0.6154	0.0769	25.4486	-19.4848
p25_1Y	13	0.6154	0.1538	0.2308	5.0622	-16.1851
atm_1Y	13	0.9231	0.0000	0.0769	-1.7943	-27.7167
c25_1Y	13	0.3077	0.6923	0.0000	226.3842	-34.8987
c10_1Y	13	0.0769	0.8462	0.0769	228.6671	-19.1670
p10_2M	18	0.7778	0.2222	0.0000	16.5498	-36.4356
p25_2M	18	0.8333	0.1667	0.0000	10.0830	-32.7695
atm_2M	18	0.8889	0.0556	0.0556	3.9961	-25.2440
c25_2M	18	0.9444	0.0556	0.0000	2.9341	-104.9686
c10_2M	18	0.8889	0.0556	0.0556	7.5845	-103.0454
p10_2Y	12	0.2500	0.5000	0.2500	21.4690	-22.6433
p25_2Y	12	0.7500	0.2500	0.0000	10.0274	-40.6065
atm_2Y	12	0.7500	0.1667	0.0833	17.6115	-37.9599
c25_2Y	12	0.9167	0.0000	0.0833	-1.9151	-118.6174
c10_2Y	12	1.0000	0.0000	0.0000	-4.3108	-114.1984
p10_3M	13	0.6923	0.3077	0.0000	17.3839	-33.8814
p25_3M	13	0.7692	0.1538	0.0769	8.2441	-32.7944
atm_3M	13	0.8462	0.0000	0.1538	1.1868	-27.9034
c25_3M	13	0.9231	0.0000	0.0769	-0.5854	-108.0858
c10_3M	13	0.9231	0.0769	0.0000	5.2464	-99.8500
p10_6M	13	0.6923	0.3077	0.0000	17.6768	-32.1492
p25_6M	13	0.7692	0.1538	0.0769	8.4480	-34.0209
atm_6M	13	0.8462	0.0769	0.0769	1.9788	-31.6372
c25_6M	13	0.9231	0.0000	0.0769	-1.4065	-117.5965
c10_6M	13	0.9231	0.0769	0.0000	4.4637	-119.7842

Notes: This table presents the Diebold-Mariano test summary for eigen model and GARCH model performance in European countries. Deltas and Maturities are separated. The 'Count' column reports the amount of corresponding delta/maturity. The 'Eigen' column reports the ratio of eigen model outperforming GARCH model in DM test. 'GARCH' column reports the ratio of GARCH model outperforming eigen model in DM tests. 'Similar' column reports the ratio of two models do not show significant difference. 'Max' column provides max Diebold-Mariano statistic for corresponding strategy. 'Min' column provides min Diebold-Mariano statistic for corresponding strategy.

value in both maturities; in 94.44%(17/18) of DM tests the eigen model outperforms the GARCH. At-the-money options are tested to be stable across maturities; the eigen model is verified to be a better choice in 75 cases out of 87(86.21%). The at-the-money option with 1-year maturity shows particularly high accuracy, with 12 out of 13 counts in the eigen model in this strategy outperforming the GARCH model. I discussed the eigen model's weak forecasting performance in the previous section, but I can conclude from the European DM table that the eigen model still holds obvious advantages for at-the-money contracts. The relatively low accuracy of the eigen model in 1-year maturity is mainly caused by the wing, especially call options.

In addition to the European summary table, I also tested the DM test summary by currency in a supplementary document. Outlines of the findings are as follows: 1) Currency liquidity is not a factor that affects the performance of eigen model forecasting. 2) As with the trend shown by the European continent table, at-the-money with 1-year maturity still performs well compared with the GARCH model. Maturity ffect is significant in deep away-from-the-money contracts. 3) In all single currency, eigen model is a better choice in option price forecasting. The differences between currencies are not significant.

To conclude, in this section I compare the performance of GARCH-type model and eigen model over a wide range of FX options. The GARCH-type model is one of the most popular parametric model in FX market, in my research I use historical data back to 1996 for the parameters estimating. It is clear from the result that eigen model outperform the well-estimated parametric model in almost all deltas and maturities.

The limitation of this chapter is that I only test the GARCH-type model performance as benchmark, market participants may work with other parametric models. More studies and comparisons will provide a much clearer picture of the eigen model predictive power. So in this chapter, the most reliable indicator for eigen model is the calculation error

from market bench.

2.7 Diebold-Mariano Test Summary (by Currency)

In order to get a precise comparison between the predictive power of the eigen and GARCH models, I list DM summary details by currency in this supplementary document from Table 5.3 to Table 5.17 in Appendix B. For each table, the 'Count' column reports the number of time series taken into comparison for the strategy in question. EUR and CHF are the currencies with the highest level of liquidity; there are 13-15 time series for each delta and maturity combination. The eigen model is considered to be a better choice for both currencies, especially with the middle-term strategy. The eigen model beats GARCH in 199 middle maturities out of 205(97.07%) for CHF, and 173 out of 200(86.50%) for EUR. At a fixed maturity, the at-the-money option shows the best performance across delta, for 1-month, 2-month, 3-month, 6-month maturity and 1-year maturity of CHF, the eigen model outperforms the GARCH model in all tested currency pairs.

I aim to provide a pricing model for the currency pairs which are not currently traded in the options market, so it is necessary to pay particular attention to the currencies with low levels of liquidity in our study. There are only 2 recorded options for HUF, MXN and ZAR. For MXN and ZAR, the eigen model is verified to be better for all 60 strategies; even the max DM statistics are much smaller than -1.96. HUF is also shown to be better forecast by the eigen model in 45 time series out of 60(75.00%). The eigen model is verified to outperform in all at-the-money options in particular. To conclude, the accuracy levels are stable across currency pairs, and liquidity of the underlying does not affect the predictive power of the eigen model. The Implementation of the eigen model will involve huge amounts of illiquid currency, and the stable performance in our study proves that it is reliable in application.

Secondly, as shown in the main document, the interval predictive power of the eigen model jumps to 64.23% for 1-year maturity, which is significantly lower than other maturities. With DM summary by currency, it is possible to gain a deep insight into the performance of the 1-year maturity option. As shown in the liquidity effect summary above, all at-the-money options are considered to be better predicted by the eigen model for illiquid currencies(HUF, MXN and ZAR). The relatively bad predictive power for 1-year maturity is mainly caused by the deep away-from-the-money option, so I can conclude from the currency summary that at-the-money is well predicted by the eigen model. For high liquidity currencies such as AUD,EUR and CHF, the winning rates of the eigen model for 1-year at-the-money are 90.00%(9/10),100%(14/14),and 100%(14/14). In other words, the eigen model is not affected by maturity when it is used to forecast at-the-money contracts, but the effect is obvious for the wings. For call option deltas 10 and 25, DKK,NOK and CAD with 1-year maturity, the eigen model is beaten by the GARCH model in all related currency pairs. For other currencies, the winning rate of the eigen model is just around 50%. The GARCH model shows an advantage in describing the shape of the volatility surface for 1-year maturity, but the eigen model still performs better in providing an accurate forecast of at-the-money contracts and determining level of volatility surface.

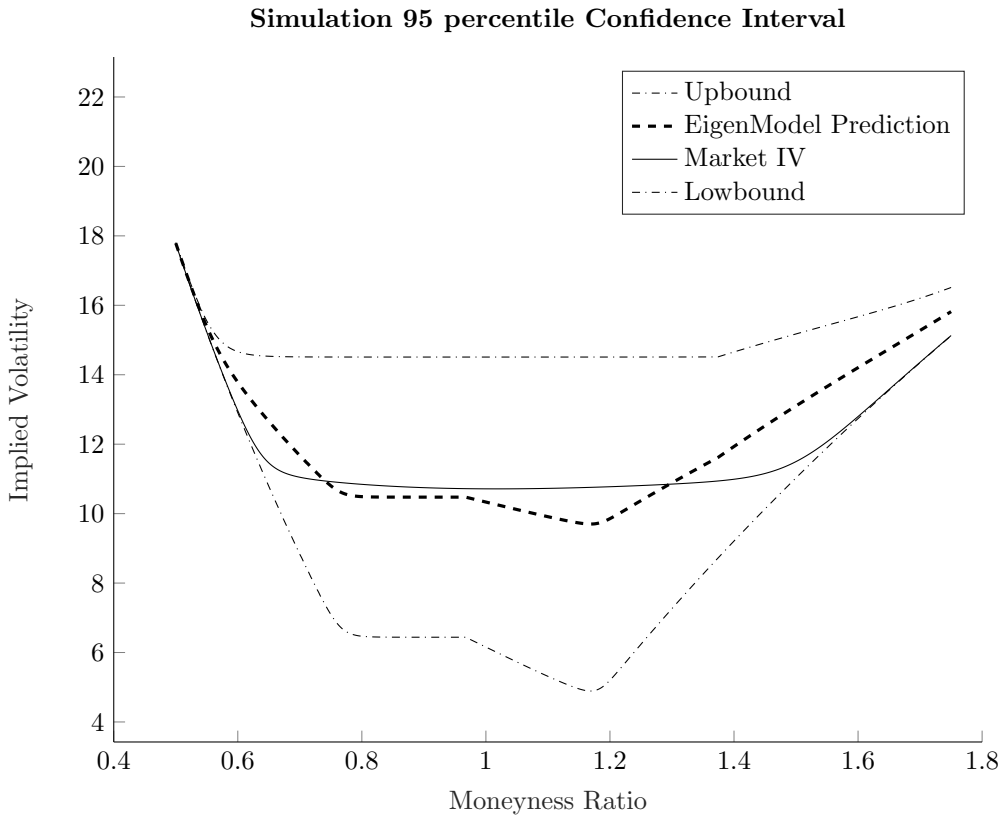
Thirdly, there is not a single currency for which GARCH model is an overall better choice in forecasting. As shown in all 15 currencies listed, the eigen model takes a higher winning rate. For particular currencies like MXN and ZAR, the eigen model is better across all maturities and deltas. The currency for which the two models have the closest performance is DKK, where tje eigen model is better in 129 time series out of 151(85.43%) and the GARCH model is better in 22(14.57%). Compared with the overall statistic where eigen value beats GARCH in 87.89% DM tests, I can conclude that the eigen model performance does not vary greatly with the currency.

To conclude, the eigen model is considered to be the obvious better choice for forecasting compared with the GARCH model. The underlying currency does not greatly influence the accuracy level of the eigen model; it provides stable and accurate forecasts for all currency pairs in the market. Deep away-from-money options with 1-year maturity are the only weak point of the eigen model, and most of the inaccurate forecasts come from these special strategies. But with reliable and accurate forecasts for 1-year at-the-money options, investors can still get the level of implied volatility, but more information is required to assess the shape of the volatility surface.

2.7.1 Eigen and GARCH Model Simulation

To assess the performance of the eigen model without effect from data selection, I implement a simulation based on GARCH parameters estimated from historical spot process. To generate the simulated spot process, firstly I randomly choose ten currencies versus USD (GBP, EUR, CHF, AUD, SEK, CAD, JPY, DKK, NZD and NOK in this simulation) and their cross currency pairs, then estimated GARCH parameters from historical spot process respectively. I generate a new set of spot process and option prices from estimated parameters, where option price is considered to be the 'fair price' used for the eigen model. In empirical study, 'fair price' is hard to obtain due to the wide distribution of market price. But in the simulation, I directly calculate it from parameters; this process eliminates the quote error and increases the eigen model's predictive power. Meanwhile in order to repeat the process GARCH model works, I estimate GARCH parameters from simulated spot process. For each turn of simulation, a new bunch of 'simulated parameters' is estimated from changing spot process. I use simulated parameters to derive the option price by using the GARCH model and compare the result with the eigen model result. By following the given steps, I simulate how the GARCH model derives volatility surface from the pre-estimated parameters. The error and irrational price are eliminated,

Figure 2.6: Simulation: GARCH and eigen model forecasts for CADSEK



Note: This figure presents hte simulation for CADSEK. Dash-point lines indicate the bounds of 95% simulated results for CADSEK FX option contracts by GARCH model, meanwhile dash lines indicate the eigen model calculation results and solid line indicate the market quote. Parameters of GARCH model are estimated by spot process from 2000 to 2018, reported by TRTH.

and both models generate ideal forecasts in simulation.

Fig. 2.6 shows the 2000 times simulation result for two models. The x-axis is the moneyness ratio for the option and y-axis is the implied volatility. So-called 'Market data' in the figure is the option data directly calculated from the original parameters; 'eigen model' plot reports the point forecast made by the eigen model along the moneyness ratio. 'Simulation distribution' give the interval forecast made by the GARCH model, where the darker zone indicates a lower percentile to the median.

I can conclude from the figure that the eigen model has a high level of accuracy across moneyness ratio. For deep away-from-the-money options, simulated eigen model point

forecast does not show a significant difference compared with at-the-money option, and still shows a high level of accuracy for volatility surface wings.

It is obvious from the simulation result that the eigen model is stable and extremely accurate for point forecast, but the interval forecast made by the GARCH model captures the 'market data' at only a 95% percentile. The high percentile indicates that GARCH is not a stable forecast model in the option pricing process; the forecasts of GARCH need to be adjusted and recalibrated with known market data. But for the market where there are not enough data to recalibrate with, such as the illiquid aim market of the eigen model, the model GARCH does not show convincing predictive power.

2.8 Concluding Remarks

In this chapter I have proposed a method for placing bounds on a volatility surface when direct market quotes are not available. I utilize the triangular arbitrage condition versus multiple leg currencies to define delta hedged portfolios of multiple currency legs and then use eigendecomposition to determine the bounds of the unknown pair. I then compare the synthetic quotes from this analysis with actual market data in a preliminary study. My objective is to evaluate the optimal trade-off between numbers of leg currencies and tightness of bounds. My final analysis will be an econometric evaluation when the observed implied correlation matrix is measured with pointwise error. In this case the observed matrix might not be positive definite, whereas the underlying, true matrix, is at least positive semi definite by construction. In this case reconditioning the stochastic matrix might be necessary to attain a new set of bounds. The effect is increasingly likely, as one increases the number of leg currencies in the matrix.

In the empirical section, I firstly show the accuracy level of the eigen model for interval forecast and point interval. In both cases the eigen model shows superior high predictive power. At-the-money option, as the most important indicator of volatility

surface, is perfectly forecasted by the eigen model; meanwhile it also generates a high level of accuracy for volatility surface wings across maturity. I also compare the eigen model with the widely used GARCH model by Diebold-Mariano test. The results show that for over 88% of strategies, eigen model is a significantly better choice for forecasting. The superior performance of the eigen model is not affected by currency pair, maturity or delta, the advantage is consistent across all strategies. In order to test the performance of both models in an ideal situation, I implement a simulation where the spot process is generated according to existing parameters. After eliminating the error made by irrational investors and improper quotes, the eigen model shows perfect forecasts for the market data.

To conclude, this paper provides a brand-new forecast model with a high accuracy level. Compared with the traditional GARCH model, the eigen model does not require market data to recalibrate with, and it is reliable for a wide range of currencies, especially for at-the-money and middle term contracts. Investors can use it to determine the volatility surface and guide investment.

Chapter 3

Eigenvalue Model Machine Learning Adjustment

3.1 Introduction

Following on from the eigensystem model described in the previous chapter, I will now conduct a machine learning exercise to improve the out of sample performance of the model. Machine learning also provides a way to rank the effect of predictors of the optimal location of the missing correlations in the viable set of eigenvalues about the maximal minimum eigenvalue.

3.1.1 Main Purpose

The primary contributions are two-fold. First, machine learning improves the eigen model's predictive accuracy in measuring risk premia of the foreign exchange options market. Risk premia are noisy to measure in options market; even if the underlying asset price process is perfectly observed, risk premia are still dominated by unforeseeable news and investors' behavior.

My study verifies that by taking related currency pairs' option trading information into consideration, the predictive power is significantly improved in two respects. The first is the direct overall predictive accuracy compared with original model. The machine learning adjusted eigen model generates an overall point forecast accuracy level of 68.27% versus the original result of 54.72%; the improvement mainly comes from the most notorious contracts: deep out-of-money options. I then introduce the out-of-sample R^2 to verify the promotion from the machine learning method. All results are shown separately by maturities and deltas to follow the format in the original model. Empirical performance indicates that machine learning provides a robust improvement across all different types of contracts and especially in those notorious contracts: deep out-of-money and mid-term maturity.

Second, I discover the relative important variables in the foreign exchange market by ranking them according to a notion of variable importance. The eigen model aims to calculate the missing implied volatility in the market by involving information for all related currency pairs; I can identify the importance of currency pairs involved by controlling the input variables. The study helps investors gain deeper understanding of covariates in foreign exchange market compared with traditional studies. The hidden relationships between currency pairs are explored and quantified through our model, which provides innovative ideas in investigating the currency market.

In this chapter, I will firstly introduce the basic conception of machine learning algorithms and its application in a financial context. I aim to combine the machine learning with my eigen model and improve the predictive power in FX option markets. In empirical study, regression tree is selected as the adjustment for raw eigen model, which is considered to be the best performance machine learning algorithm in forecasting.

As same as the raw eigen model, I use market quote as benchmark in the empirical study to test the forecast accuracy of eigen model with machine learning adjustment. In

order to present the improvement by regression tree, I will also compare the predictive power with raw eigen model by maturities and deltas.

Another question raised in the empirical study is that which currency pair is the most important predictor to other contracts' pricing. In the last section of this chapter, I will introduce the process to numerical the predictors' impact to other contracts by calculating R_{oos} reduction.

3.1.2 History of Machine Learning

The term 'Machine Learning' was identified by IBM computer scientist Arthur Samuel in 1959 as a subsection of artificial intelligence. The first implementation of machine learning was a computer game called 'checkers-playing' which could learn as it ran. With the development of related theories, a rift gradually emerged between AI and machine learning. Some scientists were interested in making machines learn from data by themselves, while others placed emphasis on the logical approach. In the 1960s, the research focused on inductive logic programming; machine learning was considered as a separate field in computer science and was abandoned.

Although several implements such as MADELINE(a neural network method which was used to predict the echo on phone lines) achieved great success in the 1970s, there was not much progress in the machine learning field. One of the most important reasons was that participants tended to save instructions and data in a single memory, which restricted the volume of data and made it harder for machine learning based methods to gain an advantage. Things changed in the 1980s when inspired by the way neurons work, John Hopfield raised a network with bidirectional lines in 1982, which is commonly known as the Hopfield Network. This innovative ideal emphasized the importance of data and information. Back propagation and other widely used machine learning algorithms were invented at around the same time. A remarkable milestone in machine learning

history was the invention of the IBM computer Deep Blue, which beat the world chess champion. The potential of machine learning has been highly valued by both academia and industry since then.

In the 21st century, especially since the 2010s, almost all leading businesses have realized that machine learning will increase calculation potential. Many projects such as DeepFace(by Facebook), DeepMind(by Google) and AmazonMachineLearningPlatform(by Amazon) have been established. These projects, whether open source or not, greatly promote research in related fields. In order to move beyond the limits of computational power, the major task of implementing machine learning has advanced from CPUs to GPUs.

With the explosive growth in machine learning related research, financial researchers and industrial investors a re paying more attention to the implement of this new technique in forecasting the rapidly changing market.

3.1.3 Why Apply Machine Learning in the Foreign Exchange Options Market?

There are several aspects of my study that make it perfectly suitable for analysis with machine learning methods:

1)Investors are eager to explore reasons behind difference in prices across option contracts. As mentioned in the introduction section above, risk premia are still a puzzle to some extent, especially in the options market. The options market is sensitive to the movement of underlying asset and related derivative contracts. As an algorithm designed for prediction tasks, machine learning is perfectly suited to the problem of pricing risk premia.

2)In my original model, the foreign exchange market is connected through the currency journey. The information for all related currency pairs is sufficiently organized to forecast

the price of missing contracts. The nature of currency markets decides that currency pairs have different weights in valuing a specific option. Dominant international currencies, eg. GBP,JPY,EUR, can play a major part in some pricing processes, meanwhile regional trading partnerships can be a core factor in currency pairs such as AUD and NZD. By applying machine learning, I can have a deep insight into how a certain options price is influenced by currency pairs in the leg journey.

3)My eigen model study covers more than 1500 different contracts. For any missing target option price, the factors involved could be different; the number of leg currencies range from three to ten. It is unrealistic to get a fixed formula to make adjustment for minimum eigen value error. Machine learning provides an methodology to make a improvement.

3.2 Literature

The application of machine learning is relatively thin in terms of asset pricing, but with the development of both technology and hardware, machine learning methods have appeared in the asset pricing literature. Rapach et al. [2013] used lagged returns as predictors to forecast the global equity market returns; the innovative lasso algorithm was implemented. Neural-networks as the most well-known machine learning technology, were first mentioned by Hutchinson et al. [1994]. This technology was also applied to the background of options pricing(Yao et al. [2000]). One of the most central question in financial study, ' risk premium measurement' was explained from the aspect of machine learning to some extent, when Gu et al. [2020a] tried to separate the risk into systematic risk compensation, idiosyncratic risk compensation and mispricing from a machine learning approach.

Recently, machine learning methods have been used to predict stock returns. Green et al. [2013] point out 330 predictive signals from previous literatures or drafts for stock

level prediction,Harvey and Liu [2015] studied the cross section of equity returns based on the bootstrap procedure; 316 factors were tested and their impact on equity return verified. Green et al. [2017] used macroeconomic variables to predict the return level of whole US market. Machine learning algorithm development provided the opportunity to apply tests across these high-dimensional problems. In addition to singe equity profit, Welch and Goyal [2008] analysed 20 variables for aggregate market return, comparing the in-sample and out-of-sample performance in a window period of 30 years.

Kelly et al. [2017a] and Giglio and Xiu [2019] applied machine learning to reduce dimension and to test factor pricing models. Freyberger et al. [2020] and Kozak et al. [2017] used shrinkage methods to approximate the stochastic discount factor for expected returns. In early implementations of machine learning, neural-networks were the most commonly used algorithm(Hinton et al. [2006],Sirignano et al. [2016]). The neural-network is also applied in portfolio management by Heaton et al. [2018].

Tree-regression was applied by Khandani et al. [2010] for credit default possibility estimation. Butaru et al. [2016] also implemented the regression tree in a risk management study of the credit card industryl; consumer tradeline, credit bureau and macroeconomic variables were considered as predictors. Moritz and Zimmermann [2016] used this algorithm for portfolio ranking. The major benefit of the regression tree is the interpretability; the regression tree methodology provides a combination of speed and good out-of-sample performance. Dietterich [2000] argued that the out-of-sample performance of the regression tree is lower than bagging algorithm in high noise settings, but the results indicate that the performance is case-dependent; tree decision and bagging algorithm are verified to be extremely similar in low noise setting studies. Campbell and Thompson [2008] compared the prediction from regression tree with the historical average return and verified the economical meaning for mean-variance investors.

In addition to the simple regression tree, boosting is a way to combine multiple over-

simplified trees and make them become a 'strong learner'. Schapire [1990] and Freund [1995] first described the idea of boosting to improve the performance of weak trees. Friedman et al. [2000] and Friedman [2001] expanded the boosting methodology and derived the gradient boosted regression tree.

A major negative point of machine learning is that it provides little help in understanding the economic mechanisms behind data; the improvement in prediction is only measurement. But in a recent study, Kelly et al. [2017b] add structure and introduce the machine learning algorithm subject to economic structure. By implementing instrumented principal component analysis they check the relationship between known latent factors for unobservable dynamic loading, and point out five factors that outperform existing economic variables. Feng et al. [2020] provide a new path to verify the value of factors. These studies help combining the data science technology with mechanisms analysis, providing a numerical way to make a selection between variables.

To conclude, the literature on the implementation of machine learning in financial study is rapidly expanding with the development of technology. But current applications mainly focus on the equity market and macro factors analysis.

3.3 Contribution of the Chapter

In this chapter, I further extend the raw eigen model with machine learning adjustment. As same as the raw model, this updated model also aims to provide accurate pricing methodology for illiquid currency pairs' option trading.

In the raw eigen model, the predictive power show significant difference between deltas. In practical investment, market participants ask for a accurate volatility surface to guide the trading and risk hedging, so deep away-from-money contracts are also important to them. These deep away-from-money contracts are harder to predict compared with at-the-money contracts. In raw eigen model's practical study, the eigen error between

maximized minimum eigen value and market quote matrix lead to the price error between calculation result and market quote, which indicate the fact that FX option market does not always follow assumption 'market tend to make itself the most stable'.

In order to eliminate the eigen error, I implement machine learning algorithm in this chapter. Implied volatility of FX options involved in the leg journey are considered as predictors in machine learning application, the eigen error is set to be aim variable. By taking advantage of related implied volatility level, regression tree provides a proper methodology to forecast the bias of minimum eigen value from idea assumption.

Practical results suggest that machine learning algorithm improve the forecast performance of eigen model significantly, especially those relative poor performance contracts in raw eigen model. Overall predictive power of eigen model rise 15% percent to 65%, prediction accuracy level difference between deltas are greatly reduced. Now the guidance for away-from-money contracts are more reliable.

The other outcome of machine learning application in this chapter is that I numerical the cross-country effect in FX option trading. I select 8 widely-traded global currency pair's option implied volatility as predictors and test the effect of them to other currency pairs. The empirical study provide a guidance for the investors in FX investment: which is the most effective currency pair need to be noticed in this market. This study provides a solution to this question from aspect of machine learning.

In the further study, the study could be further improved from two aspects: firstly, more machine learning algorithm can be applied to test the performance; secondly, the cross-country effect can be explained by many economic factors, I can connect it with the result from machine learning algorithm- a pure data-oriented outcome.

3.4 Methodology

This section firstly describes the reason why I want to apply machine learning in the eigen model. The specific regression tree algorithm is then introduced, and in the third part I introduce the method of ranking predictors.

3.4.1 Minimum Eigen Value Error and Implied Volatility

In the original eigen model, I assume the market tends to maximize the minimum eigen value. Empirical study indicates that there is a slight difference between the calculated eigen model's minimum eigen value and the benchmark. The spread between eigen value leads to the error in implied volatility. More formally, the process can be written as:

$$\hat{\sigma}(\Delta^{BS}, t, T, S_{i,j}) = \Phi(\hat{\lambda}, \Delta^{BS}, t, T, S_{(i,M_1)\ldots(M_k,j)}) \tag{3.1}$$

$$\hat{\lambda} = \tilde{\lambda}_{oem} + \epsilon_{S_{i,j}} \tag{3.2}$$

where $\hat{\sigma}$ is the benchmark option implied volatility, $\Phi(\cdot)$ is the eigen model, $\hat{\lambda}$ is the minimum eigen value for benchmark matrix, $\tilde{\lambda}_{oem}$ is the minimum eigen value from original eigen model,$\epsilon_{S_{i,j}}$ is the error I try to forecast by applying machine learning.

The aim of machine learning is to capture the ϵ from known currency journey information. The variables in machine learning are set to the option implied volatility for all currency involved in $\Phi(\cdot)$ versus USD.

A number of aspects of the implied volatility of options make them ideal to be predictors in machine learning:

1) FX options, as derivatives traded OTC, are mainly traded by institution to avoid the risk caused by holding a foreign asset. In order to lock the risk, the institution tends to trade a series of contracts contain a variety of underlying currency pairs. This

characteristic of the FX options market increases the covariance of derivatives' prices; the movement of implied volatility is significantly influenced by the related currency pairs. Meanwhile, as one of the key factors in the FX market, USD is widely traded with other currencies. The data of these USD option contracts are continuous in time series for almost all maturities.

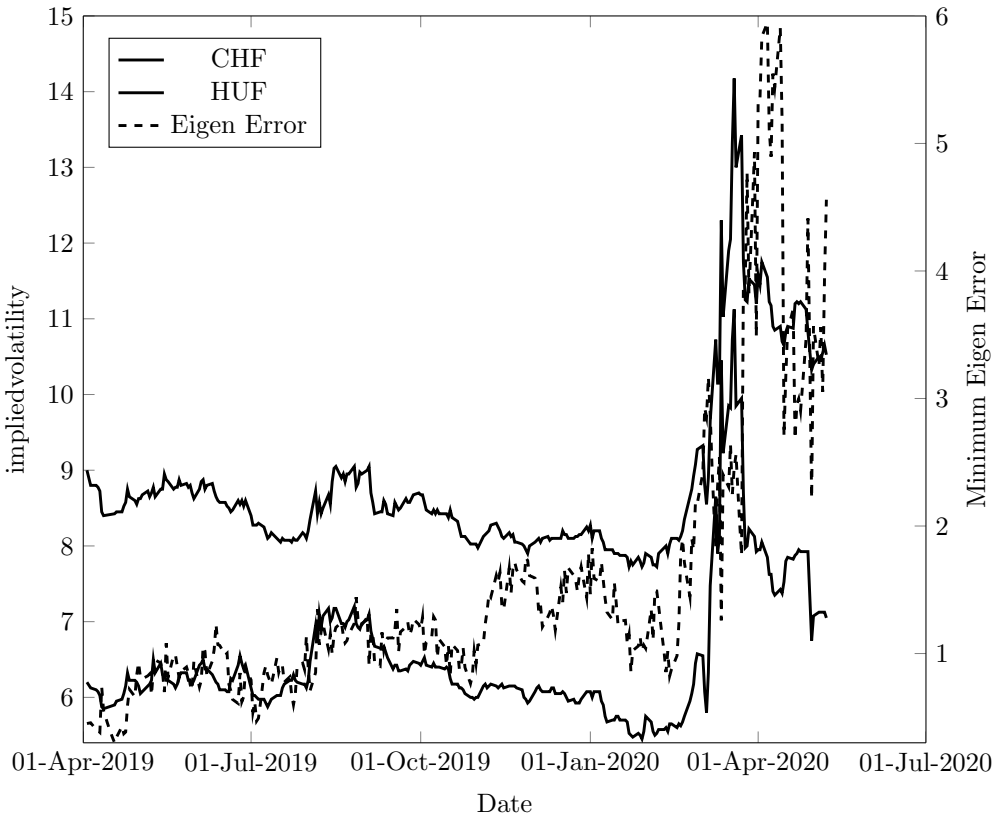
2)ϵ generates a high correlationship with the option's implied volatility. I can conclude from Fig. 3.1 that implied volatility of the two options(CHFUSD and HUFUSD) shared a similar movement in the windows period. For the time series before 2020, both options' implied volatilities are stable, and eigen value error remains at a relatively low level. Considering the complex background since 2020, especially after March, FX options' implied volatility jumped to a high level, meanwhile the error also generates an abnormal spike. Co-movement is observed across contracts. In the figure I only show the time-series for two related options.

In empirical study, I assess all currencies involved in the eigen model's currency journey versus USD. The scale of the market implied correlation matrix is dependent on the amount of currency in the leg journey for a chosen currency pair, so the amount of scale is also floating from three to ten. For different contracts, the machine learning model need to be re-trained. The model is built based on the regression tree method.

3.4.2 Regression Tree

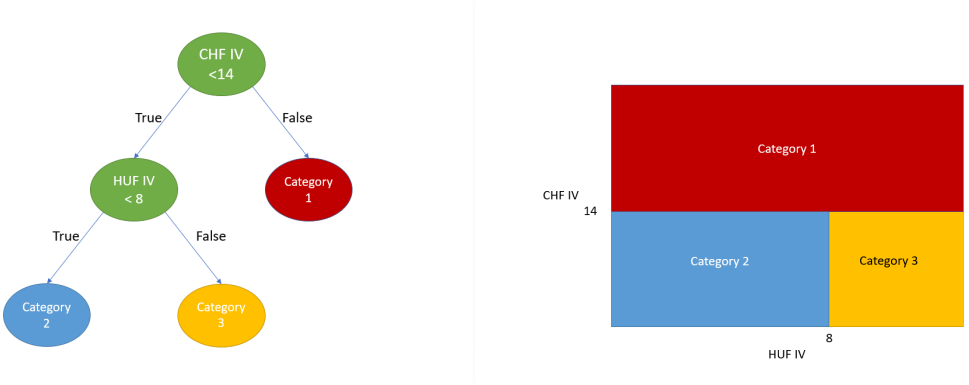
The decision tree is widely used in machine learning. It involves building a 'tree-like' model starting from observations and leading to target value, with the aim of forecasting the value based on input variables. It is known as a nonparametric supervised learning method used for both classification and regression tasks. Classification trees are decision trees where the target variable takes a discrete set of values; those where the target variable takes continuous values are regression trees. Compared with linear models, the

Figure 3.1: Implied Volatility and Eigen Value Error



Note: This figure presents the CHFUSD 1M, HUF 1M implied volatility and eigen error from Apr, 2019 to JUL, 2020 reported by TRTH. Eigen error indicate the difference between raw eigen model and market quote.

Figure 3.2: Regression Tree Example



Note: The figure presents a regression tree(left) and its representation(right) in the space of two characteristics (USDCHF implied volatility and USDCHF implied volatility). The terminal nodes of the tree are in red, blue and yellow respectively. The sample is divided into three categories.

regression tree is an attractive approach for incorporating multi-way predictor interactions. As a fully nonparametric method, the regression tree follows a basic logic of 'finding groups of observations with similar characteristics'.

Fig. 3.2 shows an example with two predictors:USDCHF implied volatility and USDHUF implied volatility. The left panel describes how the tree assigns each observation to a category based on its predictors. In the very beginning, observations are separated by USDCHF implied volatility; those above the breakpoint of 14 are assigned to category 1. Those with a lower level of implied volatility are then sorted by USDHUF implied volatility. Observations with a lower level of implied volatility than 8 are assigned to category 2, while those with a higher implied volatility than 8 go into category 3. In the end, forecasts for observations in a certain category are simply defined as average of the outcome variables' value in that that category.

Formally, a regression tree τ, with K categories, and depth of L, should be as follows:

$$g(z_{i,t};\theta,K,L)=\sum_{k=1}^{K}\theta_k 1_{z_{i,t}\in C_k(L)} \tag{3.3}$$

where $z_{i,t}$ is the predictor vector with i equities. $\dot{g()}$ is the flexible function for corresponding predictors. $C_k(L)$ is one of the K categories. θ_k is set to the mean of samples within the category.

To grow a regression tree, the main goal is to find the category that best discriminate among the potential outcomes. The method of split categories aim to minimize the forecast error at the start of each branch. At each new level l, I also need to maximize the discrepancy among average outcomes in each 'leaves'. But there is loss associated with the forecast error, which is called 'impurity', which describes similarity of observations on either side of the split. The trade off between impurity, forecast error and tree level L is the tricky part of tree growing.

Normally the basic regression tree growing algorithm is followed:

1) Start with a single node containing all points. Calculate RSS.

2) Search over all binary splits of variables for the one which reduces RSS as much as possible. If the largest decrease in RSS is less than pre-determined threshold δ or the resulting node contains less than q points, stop searching. Otherwise, take the split and create two new nodes.

3)In each new node, go back to step 1.

More formally, the process should be: $C_1(0)$ denotes the range of all elements, $C_l(d)$ denote the $l-th$ node of depth d. For level d from 1 to L, I need to update the node based on $C_l(d-1)$. For $i \in C_l(d-1), l = 1, \ldots, 2^{d-1}$, I follow the giving steps:

i) For predictors set $j = 1, 2, \ldots, P$,set threshold level α, define a split as $s = (j, \alpha)$ and divide $C_l(d-1)$ into C_{left} and C_{right}:

$$C_{left}(s) = \{z_j < \alpha\} \cap C_l(d-1); \tag{3.4}$$

and

$$C_{right}(s) = \{z_j > \alpha\} \cap C_l(d-1); \tag{3.5}$$

ii) Define the impurity function. In my study I choose the most popular impurity formula for each branch of the tree:

$$\Gamma(C, C_{left}, C_{right}) = \frac{\mid C_{left} \mid}{C} H(C_{left}) + \frac{\mid C_{left} \mid}{C} H(C_{left}) \tag{3.6}$$

where

$$H(C) = \frac{1}{\mid C \mid} \sum_{z_{i,t} \in C} (r_{i,t+1} - \theta)^2, \tag{3.7}$$

and $\theta = \frac{1}{|C|} \sum_{z_{i,t} \in C} r_{i,t} + 1$, $\mid C \mid$ denotes the number of observations in set C.

iii)Choose the optimal split

$$s^* = \underset{s}{\text{argmin}} \Gamma(C(s), C_{left}(s), C_{right}(s)). \tag{3.8}$$

iv)Update nodes:

$$C_{2l-1}(d) \leftarrow C_{left}(s^*), C_{2l}(d) \leftarrow C_{right}(s^*) \tag{3.9}$$

The output of regression tree is given by :

$$g(z_{i,t}; \theta, L) = \sum_{k=1}^{2^L} \theta_k 1\{z_{i,t} \in C_k(L)\}, \tag{3.10}$$

where $\theta_k = \frac{1}{|C_k(L)|} \sum_{z_{i,t} \in C_k(L)} r_{i,t+1}$.

3.4.3 Performance Evaluation and Variable Importance

To assess the predictive power of the machine learning adjusted eigen model for individual foreign exchange markets, firstly I simply compare the forecasts with market benchmark. Error threshold 1(%) is set to distinguish accurate and inaccurate forecasts, which is a widely accepted error tolerance level in the options market. In addition to accuracy level,

I calculate the out-of-sample R^2:

$$R_{oos} = 1 - \frac{\sum_{(i,t)\in\tau_{oos}}(r_{i,t+1} - \hat{r}_{i,t+1})^2}{\sum_{(i,t)\in\tau_{oos}} r^2_{i,t+1}}, \tag{3.11}$$

where τ_{oos} indicates that the performance evaluation will only be assessed on testing a subsample; these data will not be involved in the model estimation procedure.

In the empirical study, I also aim to identify the covariates that influence the pricing procedure while controlling for other predictors in the eigen model. For individual contracts, I rank the predictors according to a notion of variable importance. The notion of importance I use the reduction in panel predictive R^2_{oos} by setting all values of variable i to zero,while other variables remain fixed. The approach is also shown by Gu et al. [2020b].

In previous studies, cross-currency effects are studied based on several different aspects(currency liquidity, trading volume and geographical). The studies normally focus on a limited number of currency pairs, which makes it computationally infeasible to apply the adjustment for a wide range of contracts. The regression tree, as a fully nonparametric machine learning algorithm, can provide guidance on the cross-currency effect and reveal the hidden covariates in foreign exchange option markets.

The reason I select regression tree as algorithm in this chapter is that it is verified to be the best ML algorithm in prediction by Gu et al. [2020b]. But it is still worth trying algorithms mentioned in literature review. With more algorithm tested and incorporated with raw eigen model, the predictive power of eigen model maybe improved.

The cross-country effect study provides the machine learning methodology for this realm. This methodology overcome the short come of pre-determined predictors, the predictors number is improved. In my study, I select eight widely traded global currencies, but investors and researchers can add more currency pair data into the predictors.

3.5 Empirical Study of FX options

3.5.1 Dataset Description

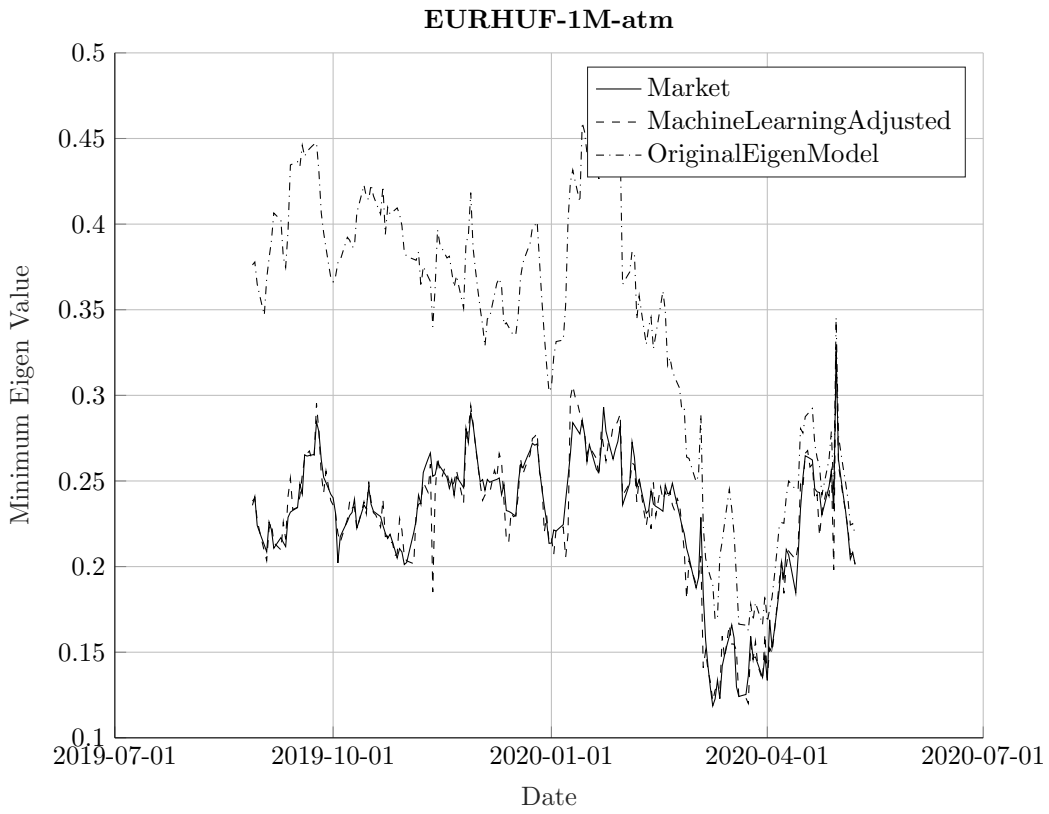
I obtain daily individual option contract implied volatility from foreign exchange options. FX option is traded by OTC market, contracts do not share a unique start date. I collect all FX option from their individual start date until May 2020. The deltas I select are five key points on volatility surface: put/call options with delta 10/25 and at-the-money contracts. These five points are used for constructing volatility surface and obtain the highest level of liquidity. For maturity, I select 1-month, 2-month, 3-month, 6-month, 1-year and 2-year contracts as our dataset, which are also the high liquidity derivatives in market. Although several currencies have extreme long maturity contracts(eg. 10-year contract for USDJPY), they do not generate meaningful results. Because market can not provides corresponding option in the procedure of currency journey in eigen model. So in the selection of contracts, liquidity is a key factor.

The number of contracts in our sample is 1870 across volatility surface and term structure (FX options information is listed in Appendix A). Three folds of data are included in this section. The first is obtained benchmark from FX option market. The second is the forecasts from original eigen model. The last part is the forecasts after machine learning adjustment.

3.5.2 Minimum Eigen Value Error Improvement

In this section, I will present the effect of machine learning on minimum eigen error. The currency pair I choose as example is the EURHUF one month contract. The reason I chose this as an example is that EURHUF is difficult to forecast. As a kind of illiquid currency, the trading with HUF in FX option is rare. The currencies involved in its leg journey are USDHUF, EURUSD USDCHF and CHFHUF, and the limited number of

Figure 3.3: EURHUF 1M ATM Minimum Eigen Error



Note: This figure presents the adjustment made by machine learning algorithm. Solid line indicate the market implied correlation matrix's minimum eigen value, dash-point line indicate the original eigen model's calculation and dash line indicate the result after machine learning adjustment.

Figure 3.4: EURHUF 1M Call10 Minimum Eigen Error

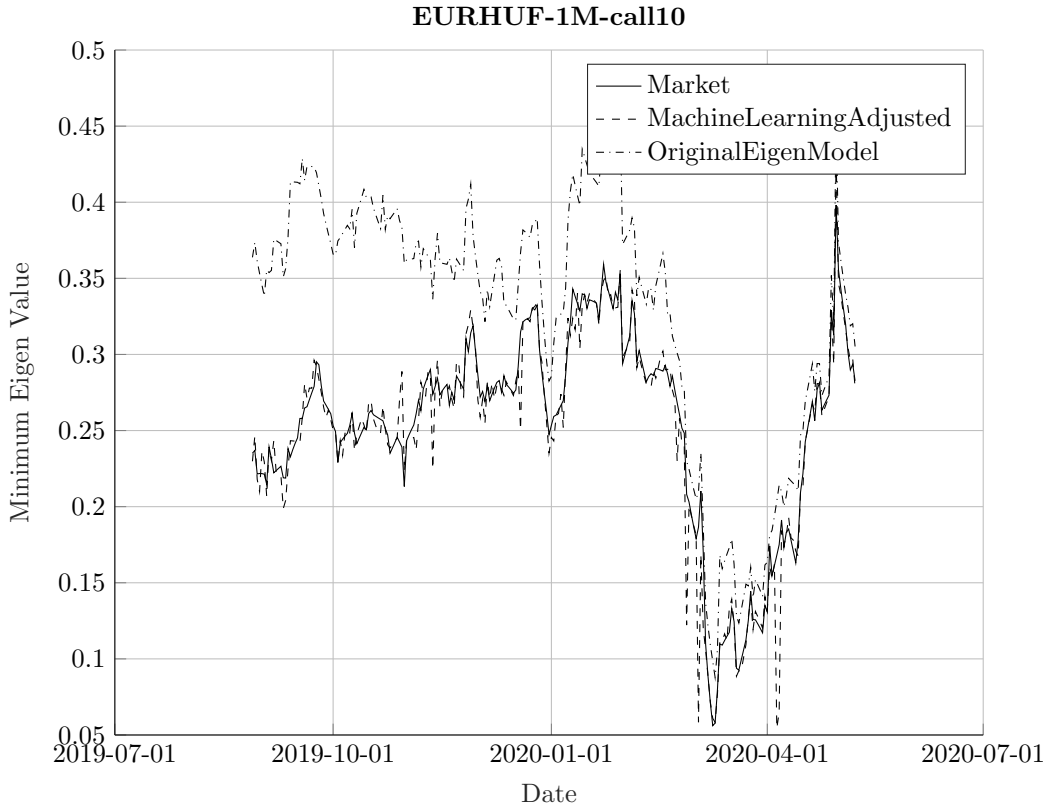


Figure 3.5: EURHUF 1M Call25 Minimum Eigen Error

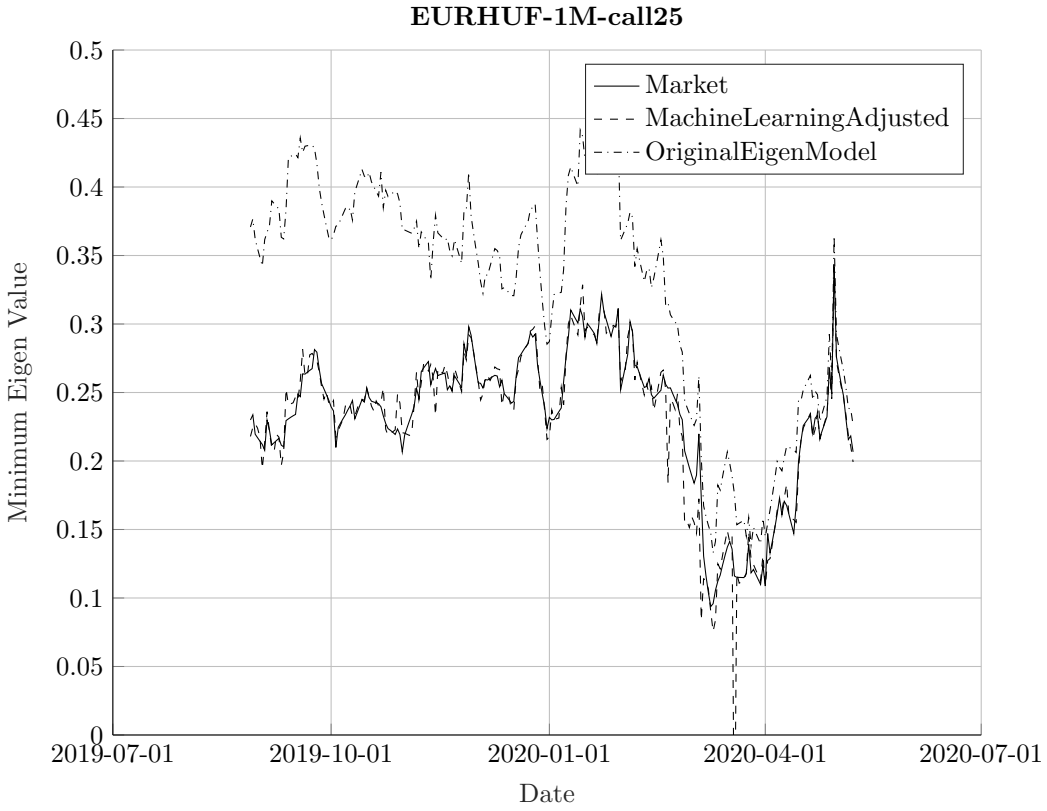


Figure 3.6: EURHUF 1M Put10 Minimum Eigen Error

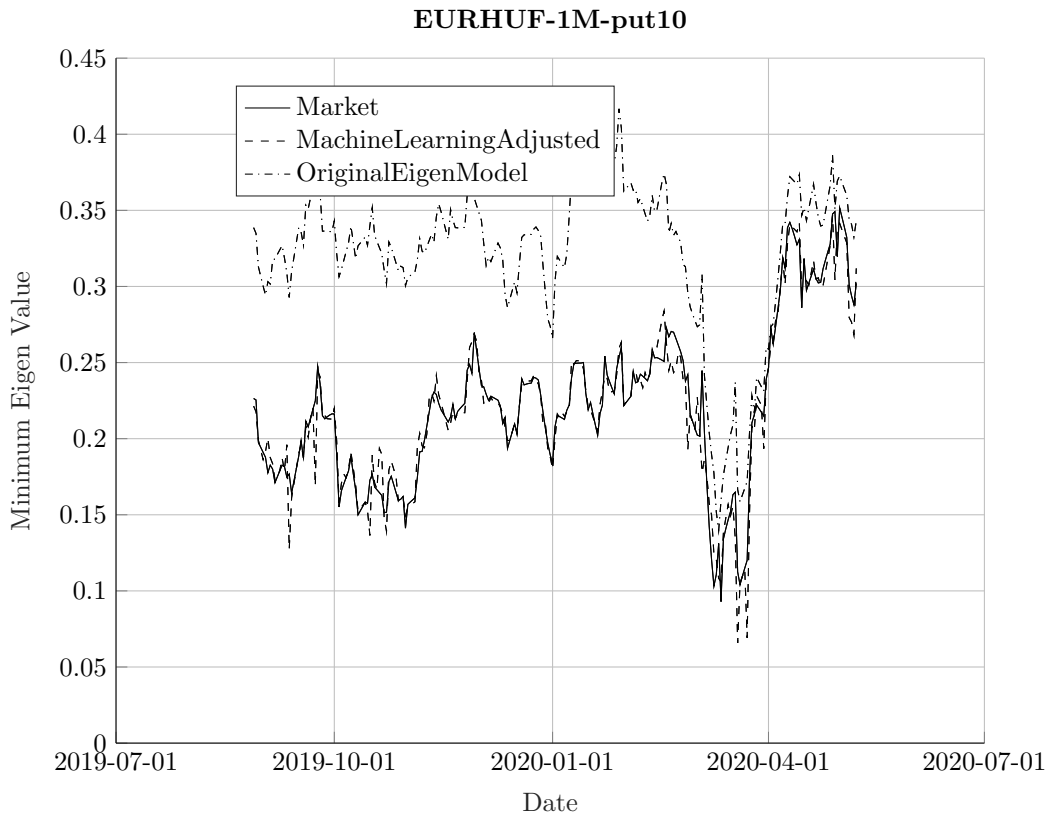
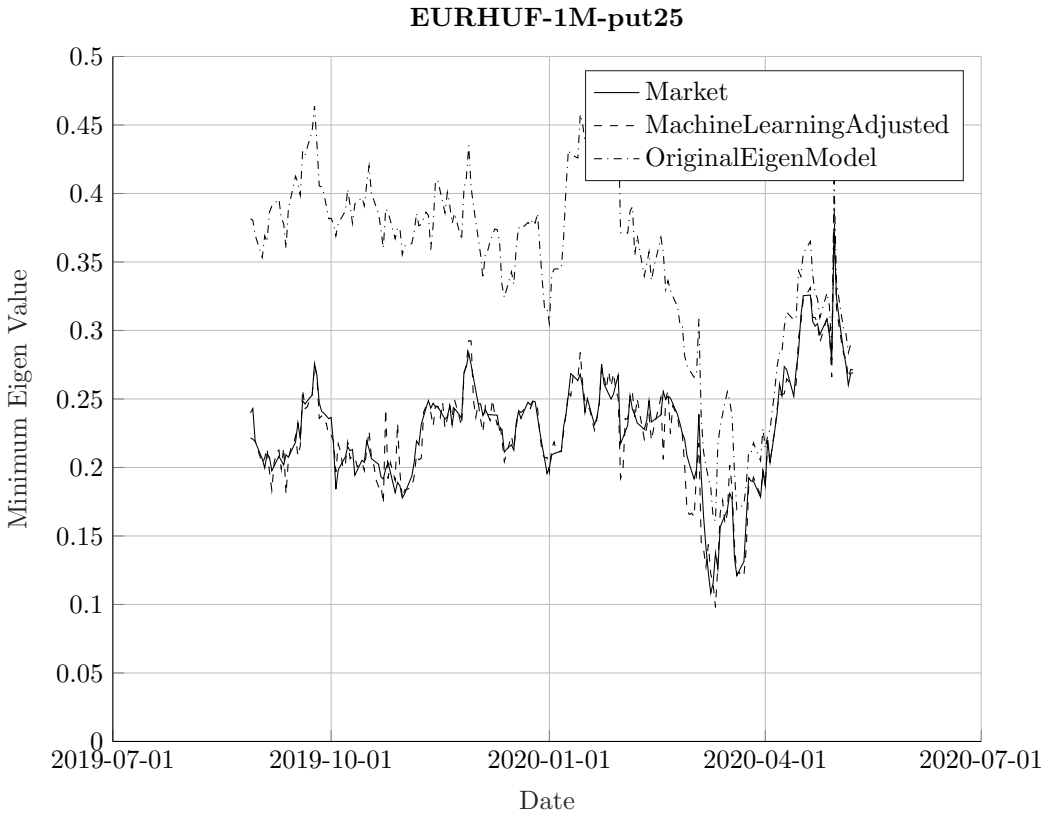


Figure 3.7: EURHUF 1M Put25 Minimum Eigen Error



currencies makes it hard to construct a proper implied correlation matrix for EURHUF. As shown in the previous section, error between prediction and market value reduces with the increment of leg currency amount.

In the raw eigen model, it is hard to get a precise minimum eigen value for EURHUF. The error leads to a bias in constructing implied correlation matrix, and finally the implied volatility is far away from the benchmark. In order to improve the forecast performance, I chose several predictors to reduce the minimum eigen value error between original eigen model result and benchmark.

Currency pairs in leg journeys are naturally considered as variables in machine learning, but for the illiquid currency HUF it is not enough to use only these variables. Considering the strong effect of EUR in European countries' trading, I take EURGBP, EURCAD, EURJPY and EURNOK option implied volatility as extra predictors. These extra predictors are either global currencies or currencies of important regional economies.

Fig. 3.3 to Fig. 3.7 present the machine learning promotion for the eigenvalue model across deltas. The results show that even for deep out-of-money contracts, machine learning reduce the error to a acceptable level. In later sections, I will show the promotion across all currency pairs, maturities and deltas.

3.5.3 Accuracy Level Improvement

3.5.3.1 Results by Delta

In Table 3.1, I report raw model and machine learning adjusted model's over all predictive power by delta. I can conclude from the table that for all deltas, machine learning improve the predictive power. As with the raw model, at-the-money contracts generate the highest accuracy level; 79.69% results from adjusted model are considered to be accurate in the market. Deep away-from-the-money contracts' forecast accuracy gradually reduces to 57.78% for put10 and 59.73% for call10 respectively.

Table 3.1: Eigen model forecasting accurate level summary by delta (in percentage)

		Raw Result	Machine Learning Result
	count	Mean Predictive Power	Mean Predictive Power
put10	374	0.4343	0.5778
put25	374	0.5123	0.7118
atm	374	0.6664	0.7969
call25	374	0.5910	0.7299
call10	374	0.5324	0.5973

Notes: This table presents the overall forecasting accurate level by delta. 'Raw Result' panel report information of raw eigen model. 'Machine Learning Result' panel report information of eigen model after machine learning adjustment. For each panel, 'Count' column report amount of time series for corresponding delta. 'Mean Predictive Power' column report ratio of forecasts where error is less than 1(%).

For at-the-money contracts, because the raw model already generates relatively strong predictive power, the improvement is not as significant as deltas on wings; the accuracy level rises from 66.64% to 79.69%. The machine learning algorithm greatly improves the pricing performance for away-from-the-money contracts, which are considered to be the most difficult to price in market trading. Call option with 10 delta predictive power rises from 53.24% to 59.73%.

In addition to the overall predictive power, market participants also pay attention to model performance across time series. In Table 3.2 I report the number of 'accurate time series by deltas. A contract is considered to be accurately predicted when the historical mean error is lower than 1(%). It can be concluded from the table that at-the-money contracts across currency pairs and maturities have the highest number of accurate predicted time series. 313 out of 374 at-the-money contracts are predicted; meanwhile, with the growth of delta to the volatility surface wings, the number of accurate contracts is reduced to 224. The result is consistent with Table 3.1, where the point forecast results are reported.

The machine learning result's adjustment in minimum eigen error greatly improves the performance of the eigen model for different contracts. 106 more contracts are accurately

Table 3.2: Eigen model forecasting accurate contracts summary by delta

	count	Raw Model Result Accurate Contract Amount	Machine Learning Result Accurate Contract Amount
put10	374	176	191
put25	374	281	288
atm	374	291	313
call25	374	284	311
call10	374	189	224

Notes: This table presents the overall forecasting accurate level by delta. The 'Raw Result' panel reports information from the raw eigen model. The 'Machine Learning Result' panel reports information from the eigen model after machine learning adjustment. For each panel, the 'Count' column reports number of time series for corresponding delta. The 'Mean Predictive Power' column reports ratio of forecasts where error is less than 1(%).

predicted by regression tree algorithm compared with the raw eigen model, especially for out-of-the-money contracts.

To conclude, machine learning algorithm adjustments improve the performance of the eigen model in all deltas. The predictive power is improved by training a regression tree and predicting the possible minimum eigen value error from implied volatility level. The improvement is more significant for the delta on volatility surface wings the relatively low-performance contracts in the raw model. With the involvement of the regression tree training set, the gap between at-the-money contract and deep out-of-money contracts is greatly reduced, and the eigen model could provide a more reliable forecast for the whole volatility surface.

3.5.3.2 Results By Maturities

As another important factor in option pricing, I report the effect of machine learning algorithm by maturities.Table 3.3 records the mean predictive power. I can conclude from Table 3.3 that the machine learning adjusted eigen model generates relatively strong predictive power for short term contracts. 71.40% calculation results of 1-month maturity have less than 1 error with market data; meanwhile with the increment of maturity, the

Table 3.3: Eigen model forecasting accurate level summary by maturity (in percentage)

	count	Raw Model Result Mean Predictive Power	Machine Learning Result Mean Predictive Power
1M	305	0.5910	0.7140
2M	305	0.6167	0.7184
3M	315	0.6023	0.6967
6M	315	0.6191	0.6765
1Y	315	0.4182	0.6435
2Y	315	0.5412	0.6493

Notes: This table presents the overall forecasting accurate level by maturity. 'Raw Result' panel report information of raw eigen model. 'Machine Learning Result' panel report information of eigen model after machine learning adjustment. For each panel, 'Count' column report amount of time series for corresponding maturity. 'Mean Predictive Power' column report ratio of forecasts where error is less than 1(%).

predictive power gradually reduces to 64.93%.

As shown in the delta section, machine learning improves the relatively low performance group considerably. The improvement for 1-year contract, which is known as the most difficult contract to forecast, is 22.53%. The accuracy level for 1-year contract is the same as other long term maturities. After the adjustment, the performance of the eigen model for all maturities is around or higher than 65%. The adjusted model is expected to be more stable across time series.

Table 3.4 reports the increment of accurate contracts amount. I can conclude from the table that for all contract maturities adjustment increases their performance. Especially for 1-year maturity contracts, 60 more contracts are considered to be well predicted after the adjustment. The difference between most well predicted and poorly predicted contracts is reduced from 101 to only 42.

3.5.4 R^2_{oos} Performance Improvement

In addition to answering most critical question: to what extent the model could be trusted as a reliable model for forecasting implied volatility, R^2_{oos} provides a deeper insight into

Table 3.4: Eigen model forecasting accurate contracts summary by maturity

	count	Raw Model Result Accurate Series	Machine Learning Result Accurate Series
1M	305	239	240
2M	305	227	235
3M	315	218	231
6M	315	206	218
1Y	315	138	198
2Y	315	193	205

Notes: This table presents the overall forecasting accurate level by maturity. 'Raw Result' panel report information of raw eigen model. 'Machine Learning Result' panel report information of eigen model after machine learning adjustment. For each panel, 'Count' column report amount of time series for corresponding maturity. 'Mean Predictive Power' column report ratio of forecasts where error is less than 1(%).

Table 3.5: Eigen model forecasting R^2_{oos} summary by delta

Delta	Raw Eigen Model R^2_{oos}	Machine Learning AdjustedR^2_{oos}
put10	0.9330	0.9363
put25	0.9471	0.9517
atm	0.9448	0.9548
call25	0.9191	0.9427
call10	0.8697	0.9029

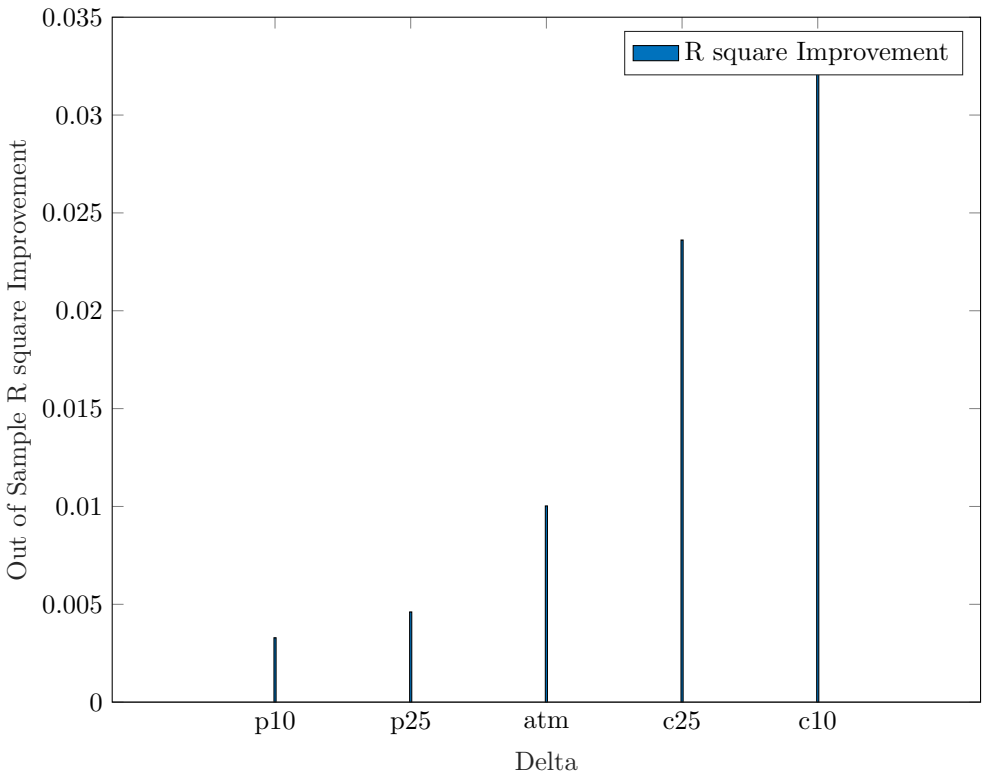
Notes: This table presents the aggregate level of R^2_{oos} by delta. 'Raw Result' panel report information of raw eigen model. 'Machine Learning Adjusted' panel report information of eigen model after machine learning adjustment. For each panel, the data presents the R^2 for all results in out-of-sample dataset in corresponding delta.

the performance for individual points among time series.

3.5.4.1 Results By Delta

Table 3.5 reports the R^2_{oos} for both original eigen model and results after regression tree adjustment. The forecast performance of model evaluated by R^2 generates a slight difference from the accuracy level in the previous section. As shown by Table 3.2, at-the-money contracts have the highest accuracy level. For out-of-sample R^2_{oos}, put option with delta 25 generates the best performance 0.9471 in raw eigen model, meanwhile at-the-money contracts is also well predicted with R^2_{oos} at 0.9448. Deep out-of-money call

Figure 3.8: R^2_{oos} Improvement By Delta



Note: This figure presents the out of sample R square improvement made by machine learning algorithm by delta. Data reported by TRTH from 1996 to 2020.

option are the hardest contracts to predict, which is consistent in both measurements, 0.8697 and 0.9191 R^2_{oos} are far behind other prediction performance.

But the benefit of machine learning is significant for these contracts with relatively low performance. Call option with 10 delta, the extreme contracts on the volatility surface, raise sharply to the R^2_{oos} level at 0.9029. For the other relevant call option, delta 25, the performance surpasses that of put10 contracts. The outcomes of machine learning are presented visually in Fig. 3.8. Machine learning adjustment reinforces the relatively weak predictive power of the raw eigen model in call options on volatility wings.

The slight difference of predictive power measured by two methods: accuracy level and R^2_{oos} arises for two main reasons: 1) The dataset chosen. Decided by the model-estimation

Table 3.6: Eigen model forecasting R^2_{oos} summary by maturity

Delta	Raw Eigen Model R	Machine Learning Adjusted R
1M	0.9490	0.9502
2M	0.9529	0.9543
3M	0.9210	0.9436
6M	0.9096	0.9434
1Y	0.9062	0.9392
2Y	0.8965	0.8999

Notes: This table presents the aggregate level of R^2_{oos} by Maturity. 'Raw Result' panel report information of raw eigen model. 'Machine Learning Adjusted' panel report information of eigen model after machine learning adjustment. For each panel, the data presents the R^2 for all results in out-of-sample dataset in corresponding delta.

process and out-of-sample set, the outcome with machine learning adjustment covers the relatively close window period from 2015-2020. Meanwhile the original method's forecast performance contains part of the previous benchmark. 2) As explained in section 3.2, in evaluating accuracy level I set a threshold value of 1(%). The value is identified to be reliable for investors' using in empirical study. R^2_{oos}, as statistic index, gives a deeper insight into individual points' performance among time series; outliers have a higher effect on the outcome compared with the accuracy level test in the previous section.

3.5.4.2 Results By Maturity

Table 3.6 reports the forecast performance by maturities, for both original eigen model and machine learning adjusted model. I can conclude from the table that for those with a good performance in the raw eigen model, the promotion from machine learning is not as significant as for contracts with a relatively low performance. For two short-term contacts, 1-month and 2-month, the R^2_{oos} rise from 0.9490 and 0.9529 to 0.9502 and 0.9543 respectively. But for mid term option, 3-month to 1-year, machine learning helps with a sharp increase in predictive power.

The R^2_{oos} increment is shown in Fig. 3.9. Middle term contracts benefit the most from

machine learning. The outcome follows our assumption: the regression trees promote the eigen model by incorporating option information on related currency pairs; middle-term contracts are the options most sensitive to market movement. The unforeseen jump in a certain option's implied volatility spreads quickly through market and influences investors' behavior in many aspects.

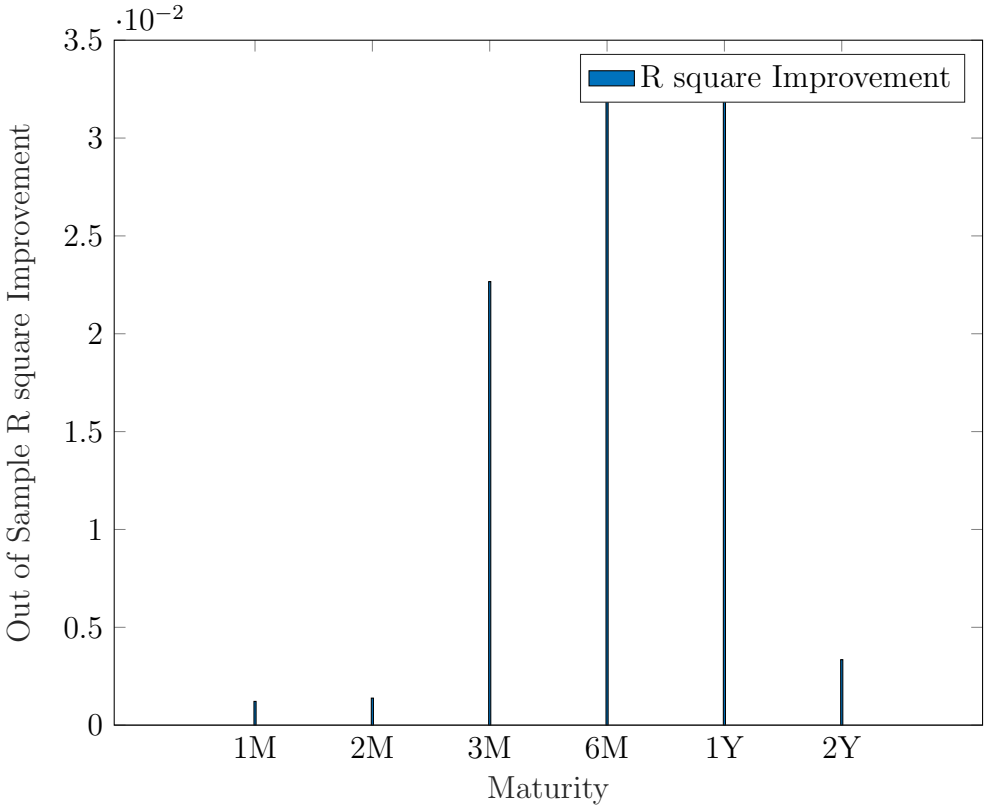
For short-term options (1-month and 2-month) the comparison indicates that they have already effectively used market information, as the predictors do not provide as high a promotion as they provide in middle-term options. Long-term options (2-year) are a specific type of contract; the movements in the market do not influence investors' choice in a long-term future. To have a better adjustment for long term FX options, it can be useful to have macroeconomic indicators as predictors.

3.5.5 Predictor Rank

I now investigate the relative importance of individual variables for the performance of promoting forecast ability. Firstly, I select the currency pair with highest level of liquidity described in previous sections as variables. Then for each aim currency's option price, I calculate the reduction in R^2_{oos} from setting all values of a given predictor to zero with in training sample and take the average value as importance measure for each variable. Variable importance within an individual training set are normalized to sum to one in order to give them relative importance.

Fig. 3.10 and Fig. 3.11 report overall rankings of variables for all aim option contracts from two aspects of data: GBPUSD and USDJPY. The reason I select these two is that they are the most important verified predictors in our study; the average normalized importance are absolute higher than other predictors. For each table, row name reports the machine learning aim currency pair, and column names indicate the variables used for corresponding aim currency pair. The color gradient within each column shows

Figure 3.9: R^2_{oos} Improvement By Maturity



Note: This figure presents the out of sample R square improvement made by machine learning algorithm by delta. Data reported by TRTH from 1996 to 2020.

the currency-pair-specific ranking of variables from least to most important(lightest to darkest). The result is also summarized in Table 3.7.

Fig. 3.12 to Fig. 3.19 report the numerical importance level of all predictors. For each predictor, I list ten aim currency pairs where the corresponding predictor has the highest level of importance. Fig. 3.10 and Fig. 3.11 demonstrate that the machine learning outcomes are in close agreement regarding the common sense. The option is highly affected by the prices of underlying currency versus USD, which indicates that investors take USD-related option price as an important reference in pricing. For example USDCAD is not tested to be a dominant global currency in our empirical study, but for the option trading related with CAD(CADCHF, CADJPY, CADNOK, CADSEK and CADSGD) it

Figure 3.10: Predictor Importance By GBPUSD

	GBPUSD	USDJPY	USDCHF	USDCAD	NZDUSD	AUDUSD	EURUSD	USDSGD
EURHKD	0.2849	0.0903	0.0708	0.0000	0.1407	0.0421	0.1092	0.1322
GBPSGD	0.2584	0.1211	0.1149	0.1395	0.1202	0.1067	0.0576	0.0818
GBPJPY	0.2555	0.1693	0.0948	0.0721	0.1304	0.0765	0.0830	0.1184
GBPNZD	0.2282	0.2267	0.1502	0.0917	0.1199	0.0590	0.0389	0.0853
GBPCHF	0.2268	0.2221	0.0946	0.1025	0.0787	0.0895	0.0860	0.0998
GBPHKD	0.2199	0.1066	0.1828	0.0000	0.0148	0.0635	0.1132	0.1231
EURNZD	0.2161	0.1217	0.1613	0.1212	0.0802	0.0610	0.1019	0.1364
CHFJPY	0.2014	0.0796	0.1324	0.1792	0.0698	0.1516	0.0872	0.0988
HKDJPY	0.1998	0.2080	0.1436	0.0000	0.0975	0.0487	0.1280	0.1176
AUDHKD	0.1868	0.0927	0.0444	0.0000	0.2359	0.1332	0.1639	0.0604
CHFSEK	0.1836	0.1553	0.0794	0.0969	0.0813	0.0585	0.0810	0.0000
EURJPY	0.1742	0.1081	0.1639	0.2043	0.0748	0.0739	0.0842	0.1166
AUDSEK	0.1738	0.1473	0.1233	0.0926	0.0585	0.1693	0.0683	0.0000
GBPNOK	0.1694	0.1679	0.1293	0.1265	0.1656	0.0435	0.0496	0.0000
AUDNOK	0.1660	0.0921	0.1005	0.1299	0.0849	0.1251	0.0875	0.0000
CHFSGD	0.1653	0.1232	0.1904	0.1582	0.1148	0.1282	0.0627	0.0572
CADJPY	0.1651	0.1382	0.1047	0.1757	0.1124	0.0724	0.0900	0.1414
HKDSGD	0.1622	0.0804	0.0892	0.0000	0.1089	0.1078	0.0573	0.2886
GBPCAD	0.1607	0.1648	0.1590	0.1396	0.0885	0.0752	0.1180	0.0943
GBPSEK	0.1541	0.0952	0.0850	0.2235	0.1128	0.1314	0.0365	0.0000
CADCHF	0.1425	0.1747	0.1436	0.0748	0.0692	0.1008	0.1833	0.1111
AUDSGD	0.1343	0.1109	0.1363	0.1283	0.1224	0.1306	0.1526	0.0846
EURSGD	0.1337	0.1761	0.1422	0.0838	0.1363	0.0719	0.0745	0.1816
CADNOK	0.1283	0.1200	0.1346	0.1759	0.0960	0.1196	0.1274	0.0000
CADSEK	0.1266	0.1176	0.1416	0.1719	0.0401	0.1043	0.1009	0.0000
EURCHF	0.1245	0.1110	0.1029	0.1273	0.1392	0.1993	0.0782	0.1175
CADSGD	0.1214	0.1137	0.1472	0.2307	0.0635	0.1173	0.0593	0.1468
EURGBP	0.1211	0.0813	0.1494	0.1133	0.1094	0.1059	0.1837	0.1360
EURAUD	0.1207	0.1256	0.1217	0.1236	0.1277	0.1087	0.1467	0.1253
GBPAUD	0.1193	0.1394	0.0953	0.1493	0.1319	0.1234	0.1255	0.1159
AUDJPY	0.1163	0.1093	0.1244	0.1080	0.2040	0.0606	0.0962	0.1812
AUDCHF	0.1140	0.2253	0.1175	0.0919	0.1129	0.1531	0.0684	0.1169
AUDCAD	0.1126	0.1980	0.1557	0.0994	0.0432	0.1016	0.1367	0.1529
CHFHKD	0.1103	0.0745	0.1696	0.0000	0.0838	0.1230	0.0665	0.2137
EURCAD	0.1082	0.0871	0.2254	0.0711	0.0982	0.0894	0.1974	0.1232
CHFNOK	0.0994	0.1618	0.1488	0.1195	0.1198	0.0704	0.0629	0.0000
AUDNZD	0.0890	0.1362	0.1169	0.1746	0.2069	0.1099	0.1010	0.0655
EURNOK	0.0724	0.1911	0.0907	0.0903	0.0933	0.1251	0.0364	0.0000
EURSEK	0.0620	0.1813	0.0916	0.1128	0.1408	0.0663	0.1818	0.0000

Note: Ranking of predictors by GBPUSD . Rows correspond to individual aim currency pair's option machine learning procedure. Columns correspond to chosen predictors and color gradients within each column indicate the most influential dark green) to least influential (yellow) predictors.

Figure 3.11: Predictor Importance By USDJPY

	USDJPY	GBPUSD	USDCHF	USDCAD	NZDUSD	AUDUSD	EURUSD	USDSGD
GBPNZD	0.2267	0.2282	0.1502	0.0917	0.1199	0.0590	0.0389	0.0853
AUDCHF	0.2253	0.1140	0.1175	0.0919	0.1129	0.1531	0.0684	0.1169
GBPCHF	0.2221	0.2268	0.0946	0.1025	0.0787	0.0895	0.0860	0.0998
HKDJPY	0.2080	0.1998	0.1436	0.0000	0.0975	0.0487	0.1280	0.1176
AUDCAD	0.1980	0.1126	0.1557	0.0994	0.0432	0.1016	0.1367	0.1529
EURNOK	0.1911	0.0724	0.0907	0.0903	0.0933	0.1251	0.0364	0.0000
EURSEK	0.1813	0.0620	0.0916	0.1128	0.1408	0.0663	0.1818	0.0000
EURSGD	0.1761	0.1337	0.1422	0.0838	0.1363	0.0719	0.0745	0.1816
CADCHF	0.1747	0.1425	0.1436	0.0748	0.0692	0.1008	0.1833	0.1111
GBPJPY	0.1693	0.2555	0.0948	0.0721	0.1304	0.0765	0.0830	0.1184
GBPNOK	0.1679	0.1694	0.1293	0.1265	0.1656	0.0435	0.0496	0.0000
GBPCAD	0.1648	0.1607	0.1590	0.1396	0.0885	0.0752	0.1180	0.0943
CHFNOK	0.1618	0.0994	0.1488	0.1195	0.1198	0.0704	0.0629	0.0000
CHFSEK	0.1553	0.1836	0.0794	0.0969	0.0813	0.0585	0.0810	0.0000
AUDSEK	0.1473	0.1738	0.1233	0.0926	0.0585	0.1693	0.0683	0.0000
GBPAUD	0.1394	0.1193	0.0953	0.1493	0.1319	0.1234	0.1255	0.1159
CADJPY	0.1382	0.1651	0.1047	0.1757	0.1124	0.0724	0.0900	0.1414
AUDNZD	0.1362	0.0890	0.1169	0.1746	0.2069	0.1099	0.1010	0.0655
EURAUD	0.1256	0.1207	0.1217	0.1236	0.1277	0.1087	0.1467	0.1253
CHFSGD	0.1232	0.1653	0.1904	0.1582	0.1148	0.1282	0.0627	0.0572
EURNZD	0.1217	0.2161	0.1613	0.1212	0.0802	0.0610	0.1019	0.1364
GBPSGD	0.1211	0.2584	0.1149	0.1395	0.1202	0.1067	0.0576	0.0818
CADNOK	0.1200	0.1283	0.1346	0.1759	0.0960	0.1196	0.1274	0.0000
CADSEK	0.1176	0.1266	0.1416	0.1719	0.0401	0.1043	0.1009	0.0000
CADSGD	0.1137	0.1214	0.1472	0.2307	0.0635	0.1173	0.0593	0.1468
EURCHF	0.1110	0.1245	0.1029	0.1273	0.1392	0.1993	0.0782	0.1175
AUDSGD	0.1109	0.1343	0.1363	0.1283	0.1224	0.1306	0.1526	0.0846
AUDJPY	0.1093	0.1163	0.1244	0.1080	0.2040	0.0606	0.0962	0.1812
EURJPY	0.1081	0.1742	0.1639	0.2043	0.0748	0.0739	0.0842	0.1166
GBPHKD	0.1066	0.2199	0.1828	0.0000	0.0148	0.0635	0.1132	0.1231
GBPSEK	0.0952	0.1541	0.0850	0.2235	0.1128	0.1314	0.0365	0.0000
AUDHKD	0.0927	0.1868	0.0444	0.0000	0.2359	0.1332	0.1639	0.0604
AUDNOK	0.0921	0.1660	0.1005	0.1299	0.0849	0.1251	0.0875	0.0000
EURHKD	0.0903	0.2849	0.0708	0.0000	0.1407	0.0421	0.1092	0.1322
EURCAD	0.0871	0.1082	0.2254	0.0711	0.0982	0.0894	0.1974	0.1232
EURGBP	0.0813	0.1211	0.1494	0.1133	0.1094	0.1059	0.1837	0.1360
HKDSGD	0.0804	0.1622	0.0892	0.0000	0.1089	0.1078	0.0573	0.2886
CHFJPY	0.0796	0.2014	0.1324	0.1792	0.0698	0.1516	0.0872	0.0988
CHFHKD	0.0745	0.1103	0.1696	0.0000	0.0838	0.1230	0.0665	0.2137

Note: Ranking of predictors by USDJPY. Rows correspond to individual aim currency pair's option machine learning procedure. Columns correspond to chosen predictors and color gradients within each column indicate the most influential dark green) to least influential (yellow) predictors.

is tested to be the most important factor in four of them; the normalized predictor ranges from 0.1719 to 0.2307. For other currency pairs, the implied volatility of the option with USD is verified to be important predictors, which indicate that USD-related currency pairs provide important information for options trading.

The second outcome of the study is exploring the 'global dominant currency' from a new perspective. By calculating the relative importance factors, I can get numerical information on how and to what extent a currency influences other currency pairs. GBPUSD and USDJPY are verified to be the top two variables in the options market. The global influences of GBPUSD and USDJPY are verified from two perspectives: i) the average numerical importance for GBPUSD and USDJPY rank in the top two among all selected predictors(GBP:0.1556, JPY:0.1371) ii) For any aim currency pairs' option, I consider the top 2 predictors as 'key predictors'; GBPUSD are key predictors in 20 currency pairs and USDJPY is key 14 times, which is the most frequent among all 8 factors.

The third finding is the regional effect, which is clearly explored in machine learning. As shown by numerical variable importance in Table 3.7, NZDUSD influences AUD trading with a variety of option prices versus currencies such as AUDHKD, AUDJPY; the effect from NZDUSD is even higher than global currency pairs. Meanwhile the numerical importance of NZDUSD is not significant for other currencies. The regional factor is mentioned in the previous literature as a result of trading and geographical effects; it is verified in our study by the machine learning algorithm. In addition, machine learning provides a numerical way to compare regional effect with global currency.

Fig. 3.20 presents key predictor frequency of all chosen predictors. We can conclude from the figure, as in the results from Fig. 3.12 to Fig. 3.19, that GBP and JPY are verified to be the 'global currency' in foreign exchange markets: GBPUSD is key predictor in 20 currency pairs, meanwhile USDJPY is key 14 times. The effect of GBPUSD and USDJPY is widely spread across almost all of FX options trading.

Table 3.7: Predictor Importance Table Summary

	GBPUSD	USDJPY	USDCHF	USDCAD	NZDUSD	AUDUSD	EURUSD	USDSGD
AUDCAD	0.1126	0.1980	0.1557	0.0994	0.0432	0.1016	0.1367	0.1529
AUDCHF	0.1140	0.2253	0.1175	0.0919	0.1129	0.1531	0.0684	0.1169
AUDHKD	0.1868	0.0927	0.0444	0.0000	0.2359	0.1332	0.1639	0.0604
AUDJPY	0.1163	0.1093	0.1244	0.1080	0.2040	0.0606	0.0962	0.1812
AUDNOK	0.1660	0.0921	0.1005	0.1299	0.0849	0.1251	0.0875	0.0000
AUDNZD	0.0890	0.1362	0.1169	0.1746	0.2069	0.1099	0.1010	0.0655
AUDSEK	0.1738	0.1473	0.1233	0.0926	0.0585	0.1693	0.0683	0.0000
AUDSGD	0.1343	0.1109	0.1363	0.1283	0.1224	0.1306	0.1526	0.0846
CADCHF	0.1425	0.1747	0.1436	0.0748	0.0692	0.1008	0.1833	0.1111
CADJPY	0.1651	0.1382	0.1047	0.1757	0.1124	0.0724	0.0900	0.1414
CADNOK	0.1283	0.1200	0.1346	0.1759	0.0960	0.1196	0.1274	0.0000
CADSEK	0.1266	0.1176	0.1416	0.1719	0.0401	0.1043	0.1009	0.0000
CADSGD	0.1214	0.1137	0.1472	0.2307	0.0635	0.1173	0.0593	0.1468
CHFHKD	0.1103	0.0745	0.1696	0.0000	0.0838	0.1230	0.0665	0.2137
CHFJPY	0.2014	0.0796	0.1324	0.1792	0.0698	0.1516	0.0872	0.0988
CHFNOK	0.0994	0.1618	0.1488	0.1195	0.1198	0.0704	0.0629	0.0000
CHFSEK	0.1836	0.1553	0.0794	0.0969	0.0813	0.0585	0.0810	0.0000
CHFSGD	0.1653	0.1232	0.1904	0.1582	0.1148	0.1282	0.0627	0.0572
EURAUD	0.1207	0.1256	0.1217	0.1236	0.1277	0.1087	0.1467	0.1253
EURCAD	0.1082	0.0871	0.2254	0.0711	0.0982	0.0894	0.1974	0.1232
EURCHF	0.1245	0.1110	0.1029	0.1273	0.1392	0.1993	0.0782	0.1175
EURGBP	0.1211	0.0813	0.1494	0.1133	0.1094	0.1059	0.1837	0.1360
EURHKD	0.2849	0.0903	0.0708	0.0000	0.1407	0.0421	0.1092	0.1322
EURJPY	0.1742	0.1081	0.1639	0.2043	0.0748	0.0739	0.0842	0.1166
EURNOK	0.0724	0.1911	0.0907	0.0903	0.0933	0.1251	0.0364	0.0000
EURNZD	0.2161	0.1217	0.1613	0.1212	0.0802	0.0610	0.1019	0.1364
EURSEK	0.0620	0.1813	0.0916	0.1128	0.1408	0.0663	0.1818	0.0000
EURSGD	0.1337	0.1761	0.1422	0.0838	0.1363	0.0719	0.0745	0.1816
GBPAUD	0.1193	0.1394	0.0953	0.1493	0.1319	0.1234	0.1255	0.1159
GBPCAD	0.1607	0.1648	0.1590	0.1396	0.0885	0.0752	0.1180	0.0943
GBPCHF	0.2268	0.2221	0.0946	0.1025	0.0787	0.0895	0.0860	0.0998
GBPHKD	0.2199	0.1066	0.1828	0.0000	0.0148	0.0635	0.1132	0.1231
GBPJPY	0.2555	0.1693	0.0948	0.0721	0.1304	0.0765	0.0830	0.1184
GBPNOK	0.1694	0.1679	0.1293	0.1265	0.1656	0.0435	0.0496	0.0000
GBPNZD	0.2282	0.2267	0.1502	0.0917	0.1199	0.0590	0.0389	0.0853
GBPSEK	0.1541	0.0952	0.0850	0.2235	0.1128	0.1314	0.0365	0.0000
GBPSGD	0.2584	0.1211	0.1149	0.1395	0.1202	0.1067	0.0576	0.0818
HKDJPY	0.1998	0.2080	0.1436	0.0000	0.0975	0.0487	0.1280	0.1176
HKDSGD	0.1622	0.0804	0.0892	0.0000	0.1089	0.1078	0.0573	0.2886

Note: Rankings of widely trading eight currency pairs. Rows correspond to individual aim currency pair's option machine learning procedure. Columns correspond to chosen predictors. Column rank follow the average importance for correspond predictors.

Figure 3.12: Variable Importance: GBPUSD

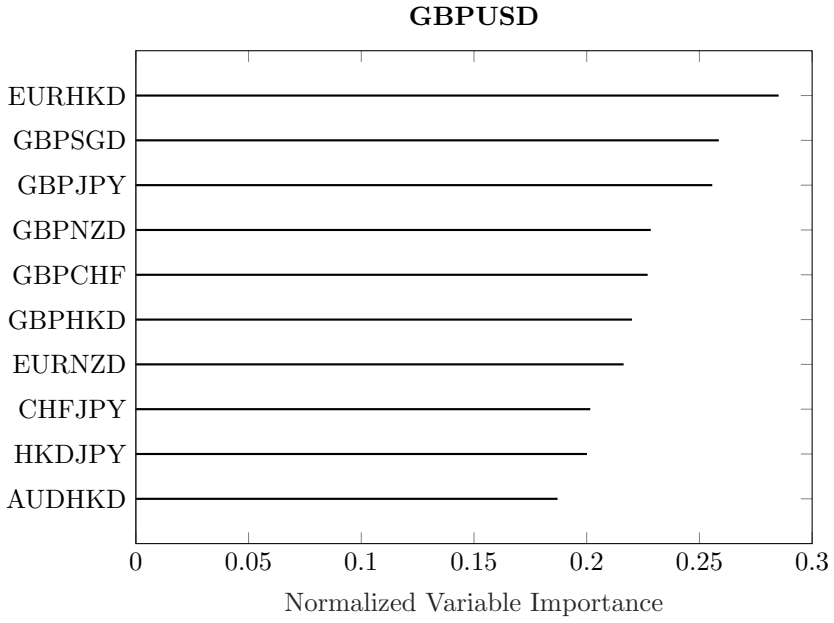


Figure 3.13: Variable Importance : EURUSD

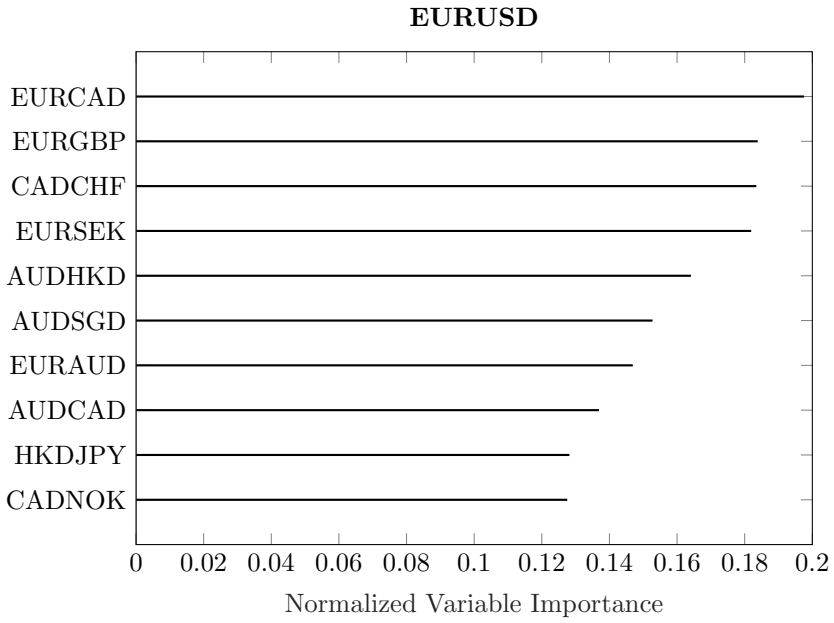


Figure 3.14: Variable Importance : AUDUSD

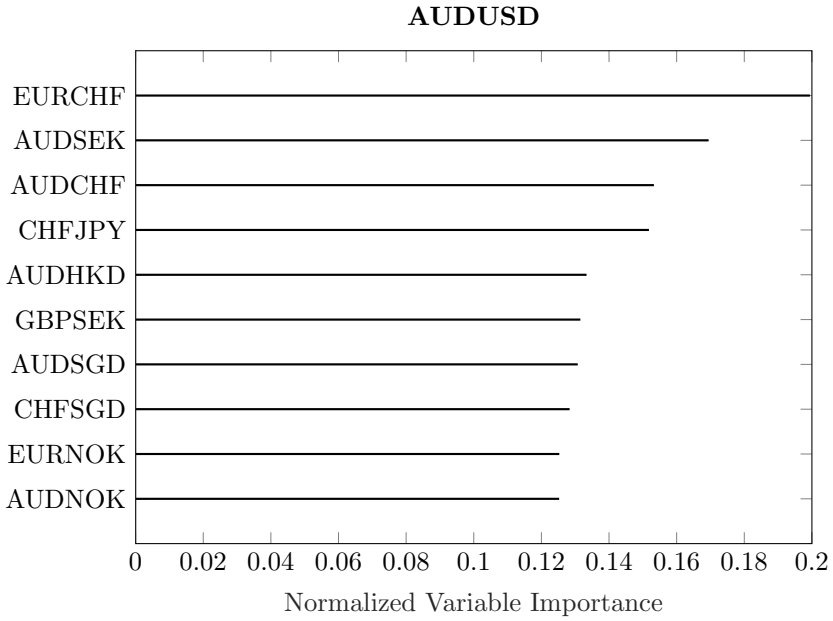


Figure 3.15: Variable Importance : USDSGD

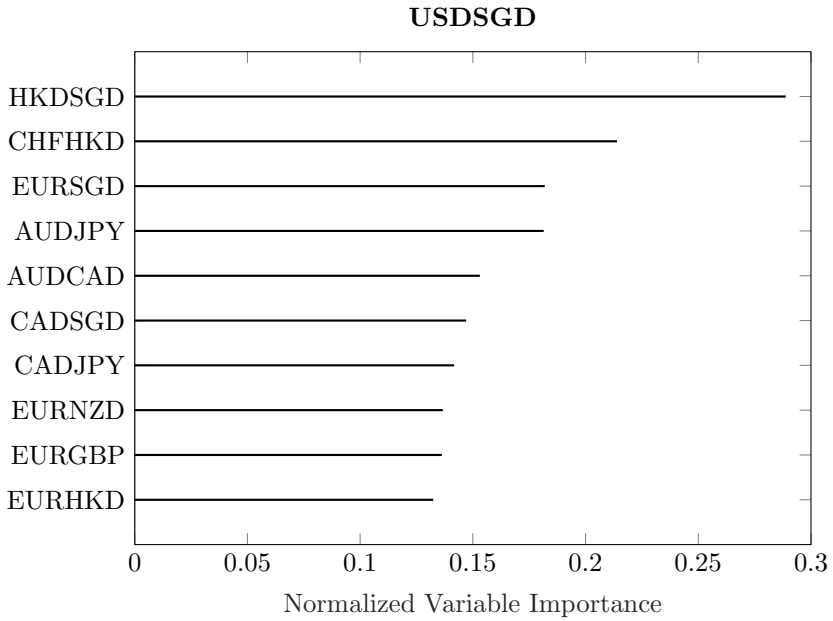


Figure 3.16: Variable Importance : USDJPY

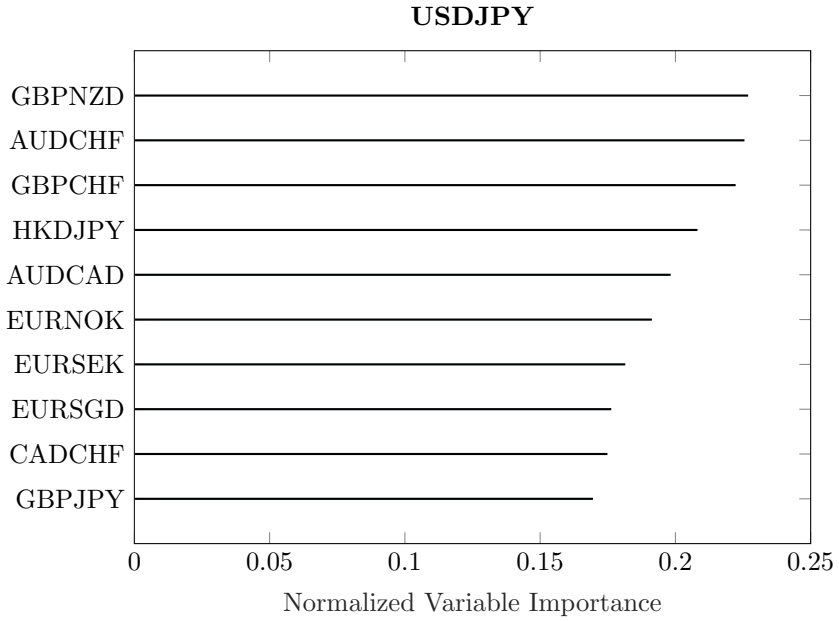


Figure 3.17: Variable Importance : USDCHF

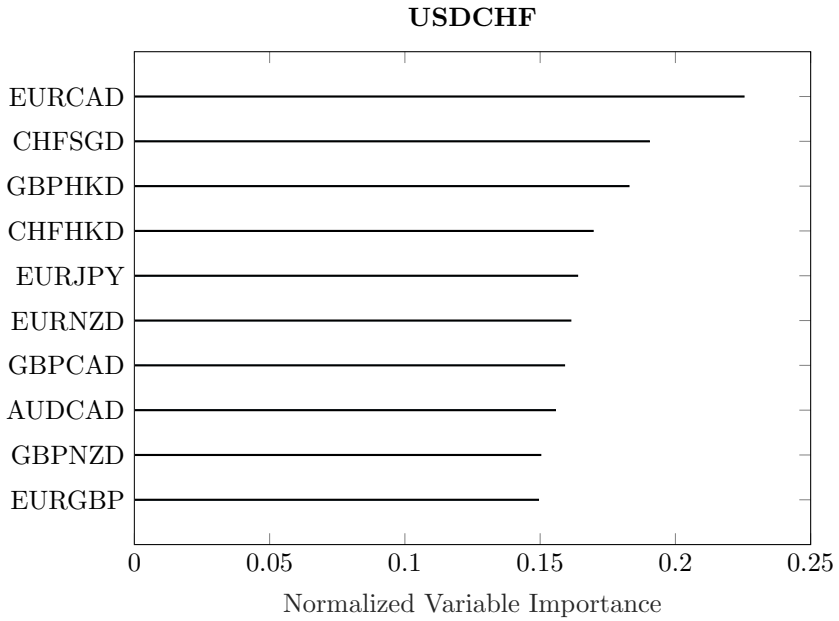


Figure 3.18: Variable Importance : USDCAD

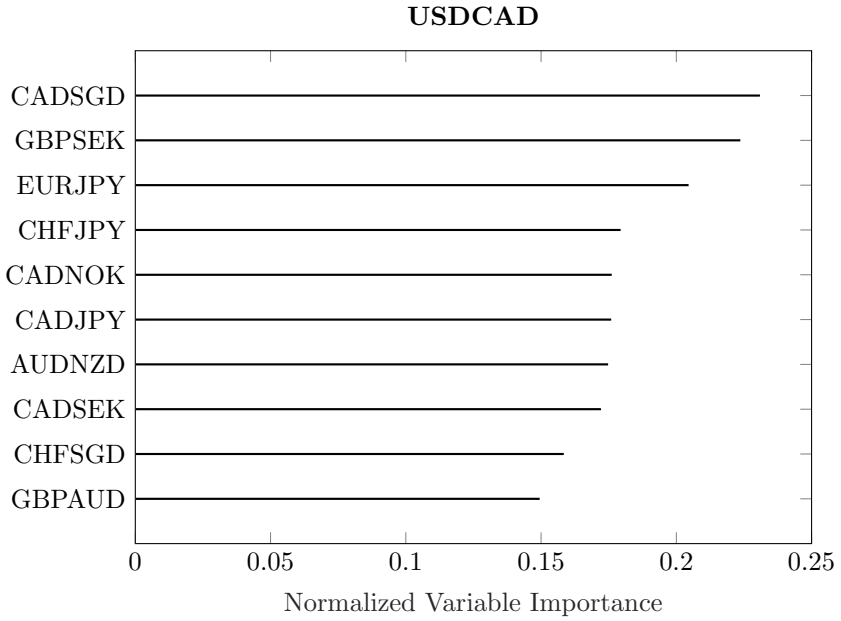


Figure 3.19: Variable Importance : NZDUSD

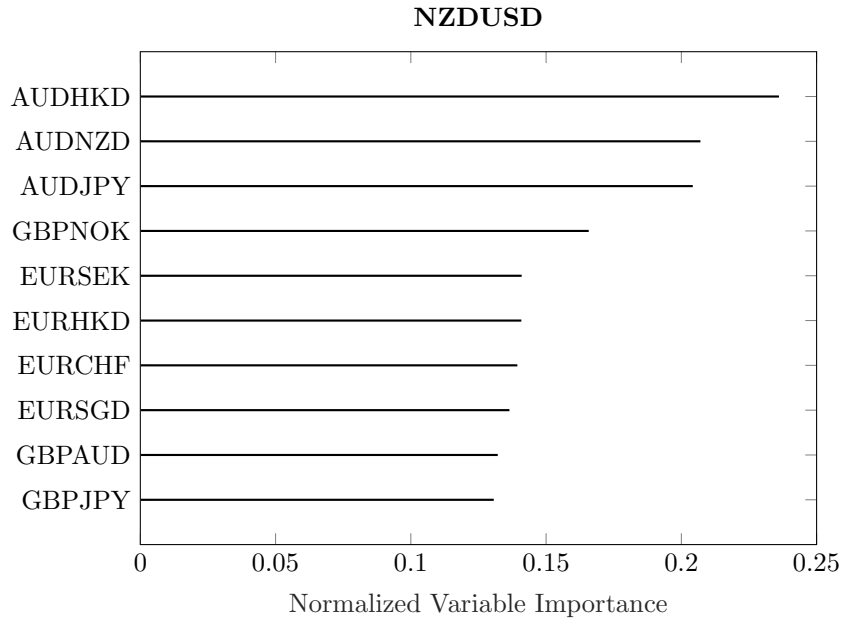
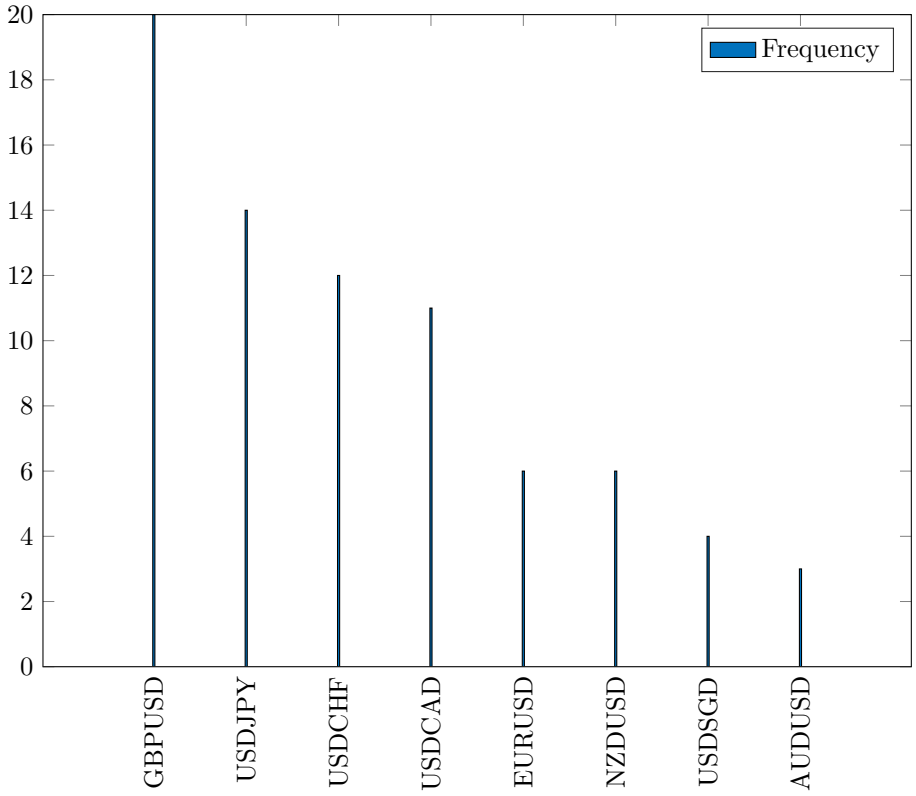


Figure 3.20: Key Predictors Frequency



Note: This figure presents frequency of key predictors for each currency. All deltas and maturities are taken into consideration. Data are reported by TRTH from 1996 to 2020.

For other regional predictors, they play a key role in the option 1) if the predictor is involved as participant and 2) in trading of a highly related country. USDCHF is verified to be a key predictor in 12 aim options; in addition to options with CHF, its effect is mainly shown in HUF and EUR related options. For NZDUSD, the regional effect is more significant, but the effect is limited in the NZD and AUD options.

The price of the option is a joint work of global indicators(GBP JPY) and powerful regional currencies. The algorithm behind it can be explained by the trading and currency liquidity. The proper use of related currency information will enable the pricing procedure to be more precise.

3.6 Concluding Remarks

To conclude, although the raw eigensystem is verified to be a reliable model to provide guidance for option pricing, especially for contracts with illiquid currency pairs as underlying assets, there are still slight errors between estimation and market quote from the aspect of minimum eigenvalue. By setting implied volatility of 'global currency pairs' based options as predictors, I apply the regression tree machine learning method to predict the eigenvalue error and cooperate this error estimation to improve the raw eigensystem model's prediction performance. Further, I test the predictors' importance level to rank global currencies and check their effect on option contracts, I verify the impact of GBPUSD and USDJPY on almost all FX option trades and the regional effect of other dominating currency pairs.

To verify the effect of the machine learning algorithm in reducing minimum eigenvalue error, I select FX option market quote as the data set to implement both raw eigenvalue model and machine learning adjusted based on raw model. I select option daily quotes from their individual start date until May 2020; delta and maturity are the same as for the raw eigenvalue model in Chapter 2. The window periods of contracts are case-dependent. For each aim contract with corresponding underlying assets, maturity and delta, I organize a 'currency journey' to derive the correlation matrix. The improvement of the results from using the machine learning algorithm is measured by comparison with the raw eigen model.

Results indicate that the promotion of machine learning implementation is consistent across maturities and deltas, especially for those with relatively low performance contracts. For the raw eigensystem model, put/call 10 delta contracts are the lowest performing contracts, and these deep away-from-money deltas are the contracts where machine learning improves the performance most significantly. With regard to maturity, 1-year contracts are hard to predict due to the effect of both long-term and short-term

variables, but machine learning also improves the predictive power to a similar level as other maturities. Generally speaking, implementing machine learning makes sure the performance after adjustment is consistent across volatility surface and maturities, which gives more reliable guidance to market participants.

In addition to point forecasting accuracy level, I also test the R_{oos} to provide a deeper insight into the improvement from machine learning. The R_{oos} improvement verifies the conclusion from point forecast: the improvement made by machine learning is more significant for those weak strategies in raw eigensystem estimation,as the differences between forecast accuracy level after machine learning adjustment is less than they are in the raw model. The machine learning adjustment makes the predictive power more consistent across deltas and maturities.

With the help of the regression tree method, I also numerical the importance of eight selected currency pairs(GBPUSD, USDJPY, USDCHF, USDCAD, NZDUSD, AUDUSD, EURUSD and USDSGD) to other option contracts. These currency pairs are involved in most of the currency journeys, which provides a chance to check their effect on our aim currencies. Cross-currency effect is an important question in the study of global markets, especially in derivatives trading such as option and forward, but previous studies usually focused on several countries with either geographical connection or significant trading partnerships. The global impact and the comparison between global currency and regional effect is hardly measured, so this chapter fill this gap. By setting currency pairs as predictors and checking their effects on options trading all over the world, I can obtain indicators for domestic and foreign investors.

I can generate from a set of results that the domestic currency with USD is always the important predictor in option trading. In addition to USD, GBP and JPY are two currencies with considerable impact on other contracts. The regional effect is verified in our study, as the currency of the foreign country closest in distance is tested to be an

important predictor in option pricing. My study verifies that the option pricing is case-dependent by underlying asset and provides a rank of predictors for market participants.

In further research, more machine learning algorithms could be considered. Although the regression tree is verified to be the most reliable algorithm for prediction, it is still worth trying to compare the performance between different methodologies. In my study, I use selected currency pairs as predictors, but more variables could be taken into consideration. More studies on the regional effect of currency pairs on each other would prove interesting; trading volume between nations and government policies might provide a deep insight into such an effect.

To conclude, the regression tree method raises the predictive power of the eigensystem model to a new level, filling the gap between maturities and deltas. It also helps point out the variables that need to be noted in options trading from the perspective of cross-currencies.

Chapter 4

Eigenvalue Analysis Based on S&P500

My third application of eigensystem analysis focuses on the construction of portfolios of stocks in the cross section. Previously I looked at the eigendecomposition of a positive definite matrix of the following form:

$$\boldsymbol{C}(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x})\boldsymbol{D}(\boldsymbol{x})\boldsymbol{V}'(\boldsymbol{x})$$

where $\boldsymbol{C}(\boldsymbol{x})$ is a positive definite matrix, $\boldsymbol{V}(\boldsymbol{x})$ is a columnwise matrix of eigenvectors, such that $\boldsymbol{V}(\boldsymbol{x}) = [\boldsymbol{v}_i(\boldsymbol{x})]_{i=1}^{N}$ and $\boldsymbol{D}(\boldsymbol{x})$ is a diagonal matrix of the form $\boldsymbol{D}(\boldsymbol{x}) = \text{diag}(\boldsymbol{d}(\boldsymbol{x}))$ where $\boldsymbol{d}(\boldsymbol{x})$ is a vector function of some state variable $\boldsymbol{x}$.

In these prior cases our interest focused on the smallest element of the vector $\boldsymbol{d}(x)$, the smallest eigenvalue and the domain for which it was greater than zero and hence preserving the positive definiteness if $\boldsymbol{C}(\boldsymbol{x})$ and it's corresponding correlation matrix $\boldsymbol{C}(\boldsymbol{x}) = \boldsymbol{R}(\boldsymbol{x}) \circ \sqrt{\boldsymbol{d}(\boldsymbol{x})}\sqrt{\boldsymbol{d}'(\boldsymbol{x})}$, where $\circ$ is the element-by-element product of two identical matrices.

The work in this chapter extends, via simulation the results found in Dovonon et al.

[2021], which look specifically at the eigenvalue structure $\boldsymbol{d}(\boldsymbol{x})$ when $\boldsymbol{C}(\boldsymbol{x})$ is either the spot or integrated covariance matrix directly estimated from high-frequency returns. My interest is in extending the results on the sport and integrated covariance matrix to specific columns of $\boldsymbol{V}(\boldsymbol{x})$, most specifically $\boldsymbol{v}_1(\boldsymbol{x})$, the eigenvector of the largest eigenvalue $d_1(\boldsymbol{x})$.

Because the results I am interested in are mostly in terms of $N \to \infty$, I shift from foreign exchange options to studying asset classes with larger cross sections, in this case equity returns.

4.1 Background

Following from my previous work I set out the following vector Brownian Semi Martingale in Grigelionis form:

$$\mathrm{d}\boldsymbol{x}(t) = \boldsymbol{\mu}(\boldsymbol{x})\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{w}(t) + Jumps$$

where $Jumps$ is an affine jump process, of the usual form following Dovonon et al. [2021]. It is important to note that our interest lies not specifically in $\boldsymbol{\Sigma}(\boldsymbol{x}, t)$, but $\boldsymbol{C}(\boldsymbol{x}, t) \approx \int_0^T \boldsymbol{\sigma}(\boldsymbol{x}, s)\sigma'(\boldsymbol{x}, s)\mathrm{d}s$. For all of the following analysis, the vector $\boldsymbol{x}$ is always ordered by the magnitude of the elements of f $\boldsymbol{v}(\boldsymbol{x})$ from largest to smallest.

4.1.1 Portfolio Composition

Let h be a fixed integer always smaller than the dimension N. Such that for a matrix decomposition of the form $\boldsymbol{C}(\boldsymbol{x}, t) = \boldsymbol{L}\boldsymbol{L} + \mathrm{diag}[\boldsymbol{e}]$, with $\boldsymbol{L}$ being of dimension $N \times h$ and $\boldsymbol{e}$ being an N length non-zero vector. I should note that $\boldsymbol{L}$ and $\boldsymbol{e}$ have set properties for a given sequence of spot covariance matrices of the time horizon t to T. As $T - t \to 0$ the measurement of the integrated covariance converges on the spot variance-covariance matrix.

The canonical results of Chamberlain [1983] posit an unlimited set of assets whereby $N \to \infty$, under this rule the limit of the estimated root $\hat{d}_1/N \to^d d_1/N \to \lambda_1$, when the number of assets increases and the largest eigenvalue converges on a constant. That is as the dimension increases the quotient of the largest eigenvalue by the dimension converges to a constant.

The confluence of results between the theoretical arguments in Chamberlain [1983] and the results of Dovonon et al. [2021], Chen et al. [2019] focus on the composition of the portfolio formed by using the eigenvector $\mathrm{d}\varpi(\boldsymbol{x}, t) = \boldsymbol{v}_1(\boldsymbol{x}, t)'\mathrm{d}\boldsymbol{x}$.

4.1.2 Statement of contribution

In a similar way as with the question posed in Chapter's 2 and 3, this research contribution seeks to ask: *how many assets are sufficient to construct a portfolio that approximates the asymptotic portfolio as* $N \to \infty$*?* Let $\boldsymbol{v}_1^{[J]}(\boldsymbol{x}, t)$ denote the first J elements of the largest eigenvector of $\boldsymbol{C}(\boldsymbol{x}, t)$. Hence $\boldsymbol{v}_1^{[N]}(\boldsymbol{x}, t) = \boldsymbol{v}_1(\boldsymbol{x}, t)$.

For a portfolio with weights $\boldsymbol{v}_1(\boldsymbol{x}, t)$, when $N \to \infty$, at what dimension J does the portfolio $\boldsymbol{v}_1^{[J]}(\boldsymbol{x}, t)'\mathrm{d}\boldsymbol{x}^{[J]} \to \boldsymbol{v}_1(\boldsymbol{x}, t)'\mathrm{d}\boldsymbol{x}$. Finding this out allows me to answer an important unanswered question from the literature on extracting Principal Components from cross sections of stocks, which is: *what dimension is the correct one to use when conducting an empirical principal component study on high-frequency data*?

The literature is quite sparse on this topic. For instance Chen et al. [2019] use 70 to 100 stocks to build portfolios. Dovonon et al. [2021] use exactly 100 from a potential set of 500 assets to construct approximate factors of a cross section of larger stocks.

I will use the data set of Dovonon et al. [2021] and combine this with simulation evidence to illustrate that the dimension N plays an important role in the standardized results from previous studies on the degree of tracking of a portfolio. My results indicate that for a standard collection of stocks such as the S&P 500, the minimum number of

stocks to include is around 50. Once the number of assets in the cross section exceeds 200, the scaling indicates that further additional assets no longer significantly changes the structure of $\mathrm{d}\varpi^{[200]}(\boldsymbol{x},t) = \boldsymbol{v}_1^{[200]}(\boldsymbol{x},t)^{[200]\prime}\mathrm{d}\boldsymbol{x}$. The remainder of this chapter introduces the empirical study and explains the simulation conditions. I then outline the key results and illustrate the decay in the variation in the eigenvector with respect to dimension. I then apply my analysis to equity data and re-evaluate the results of Dovonon et al. [2021] and Chen et al. [2019] in light of my findings. I show that while both studies are robust, there are significant trade-offs to be made in terms of the choice of N and the sample frequency $T-t$.

In this chapter, I will firstly show the simulate the first eigen value movement with the increment of portfolio dimension. Simulation aim to provide the eigen value trend in ideal situation. Then empirical study based on S&P500 equity data from 1996 to 2020 are presented. The largest eigen value structure are listed versus portfolio dimension and time series.

Convergence of the first principal component is defined in this chapter, I use vector angel between corresponding eigen vectors as the indicator. In order to have a deeper insight in to principal component structure, I set two thresholds for eigen vector angels.

Rest of this chapter is organized as follows: firstly I will present the simulation process, then S&P500 data is introduced for empirical research. Then the critical point of convergence is discussed, a distribution of convergence portfolio dimension is presented in the last section of this chapter.

4.2 Contribution of the Chapter

In this chapter, I aim to find the minimum asset number to apply the principal component analysis. Although PCA is a widely used analysis methodology in financial study, the minimum asset number required in high-frequency data is still a puzzle. I use S&P500

constituents market data back to 1998 as empirical data set.

The key finding in this chapter is the convergence of the first principal component. In empirical study, I estimate the variance-covariance matrix monthly. For certain window period, I check the normalized largest eigen value movement versus matrix dimension. The convergence is significant and consistent across time series.

Convergence of the first principal component indicate that adding more assets into portfolio provides no more information. This result provides the guidance for PCA in high-frequency data, a portfolio with 150-250 assets is enough to describe S&P500 market.

4.3 Empirical study

In the empirical study, I focus on the eigenvalue/vectors structure for a high-frequency S&P 500 equity data set (stock tickers are listed in Table 5.1 in Appendix A. By increasing the scale of variance matrix, I aim to test the critical point where the eigen structure turns to stable, and does not fluctuate with market information. The sections are organized as follows: data selection and cleaning procedures are introduced, then I generate eigenvalue structure based on data set after cleaning. In the final section, I present the eigenvectior convergence process.

4.3.1 Dataset description

As the largest and most liquid equity market, S&P 500 provides sufficient data to implement the eigenvalue analysis over time. In order to show a continuous time series of egien vector/eigenvalues, the data period need to be carefully selected. Data providers such as Bloomberg and Thomson Reuters have historical data back to 1998 for S&P 500 equities. In a twenty two year window period, a large number of stocks have been delisted, so it is important to find a relatively stable period to apply eigenvectior structure analysis.

In practical study, I would consider that an equity scale larger than 300 is sufficient to speculate the convergence of eigenvector/values, which is identified in the practical study. The window period I selected for empirical study ranges from 2015 to 2020. There are over 450 equities with high-frequency tick historical data for the given period.

Data cleaning takes place as follows:

1) Remove values that are five times higher than weighted mean value. I would consider these irrational data as incorrect inputs by the data providers. The eigenvector analysis relies on the realized co-variance matrix and taking abnormal high/low data into the model will greatly influence the result.

2) I select five minutes frequency to evaluate realized correlation. Five minutes frequency is widely used in a wide range of studies, and is verified to provide a reliable estimation of correlation. I also show results with fifteen minutes estimation for part of data, it can be derived that the difference between five and fifteen minutes are slight enough to be ignored. As for window period, one-week and one-month is both tested, the result will be shown in empirical study.

3) All equities are ranked by trading frequency. I aim to find the equities that provide the highest degree of price informativeness, and construct the portfolio based on these stocks. With the increment of covariance matrix, I would add less liquid equity into portfolio. Threshold is pre-determined to detect if the equity is rarely traded in a certain window period(1-week or 1-month).

To conclude, the priority in selecting equities is decided by two rules: firstly I evaluate the frequency in the window period; equity with higher levels of liquidity enjoys a higher rank in our portfolio. 2) For every re-estimation circulation(weekly or monthly), a slight adjustment will be made to eliminate the equities delisted or non-traded.

4.3.2 Eigenvalue Simulation

As noted above, my study aims to provide guidance on the eigenvalue structure for a large empirical dataset, in this case high-frequency returns on a large cross section of equities. Prior literature sets the portfolio dimension to between 90 and 50, in my study I present the convergence of principal components and in particular the eigenvector weights for scaling the dimension of the cross section up to 500 or more.

Before moving on to the empirical results, I firstly apply the simulation for eigenvalue structure. I have a set of factors to simulate the complex market environment:

$$\mathrm{H}_t = [H_t^0, H_t^1, H_t^2 \dots H_t^n]$$

, the factors amount n is changed with simulation turns.

For each amount of H_t, I generate a new set of factors and loading, the time series are generated by :

$$\boldsymbol{S}_t = \boldsymbol{H}_t \boldsymbol{F}_t + \boldsymbol{\epsilon}_t \tag{4.1}$$

$\mathrm{F_t}$ are generated data loading to construct time series. ϵ_t is the stochastic error with no correlation to each other. All the data mentioned follow normal distribution. S_t is the matrix containing the simulated equity price process.

The simulation is organized according to the following steps: 1) For any fixed amount of $\boldsymbol{H}_t$, I apply 1000 turns of simulation. In a turn of simulation, I generate a data set by 4.1.To meet the requirement of full rank and simulate high frequency trading background, the$t$ is set to 10000. This is the frequency much higher than the 1-week or 1-month market data (the window period I use for actual market data).

2) Then for generated dataset $\boldsymbol{S}_t$, I select a different dimension of equities to construct our portfolio, the minimum scale I set is 20 and the maximum is 500. For each selected portfolio, I implement the eigenstructure analysis to collect normalized eigenvalues and

eigenvectors.

3) Increasing $\boldsymbol{H}_t$ by 1, then go back to step1.

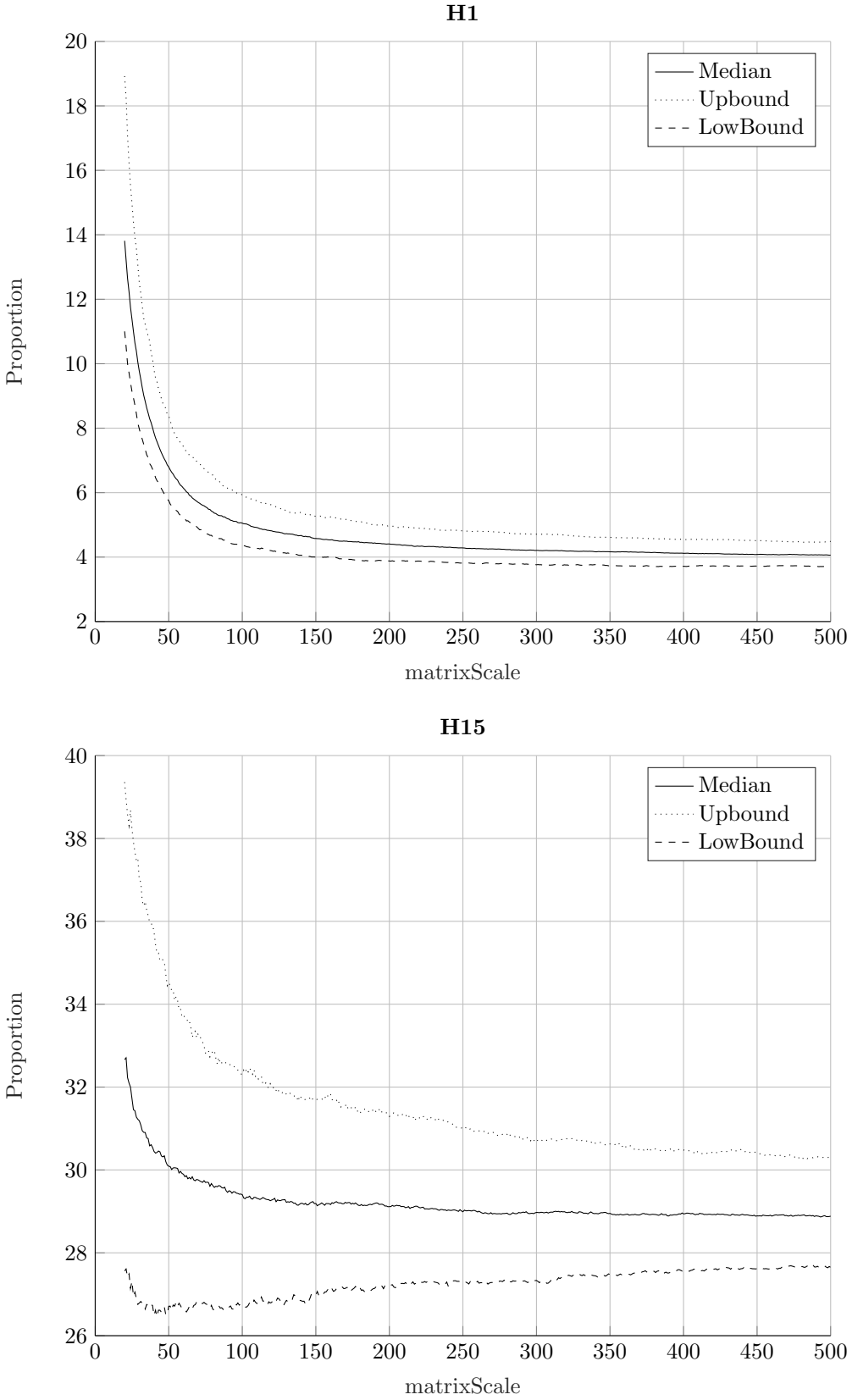
As such, I aim to verify eigenstructure under different numbers of potential factors. In my study, the H range is from 1 to 500 to simulate different complexities in the underlying factor structure for the market. For each factor set, I generate thousands of data sets as simulated 'market data' and select different numbers of 'equities' from the market data to construct portfolios for eigenstructure analysis, the number of equities in a portfolio ranges from 20 to 500.

Fig. 4.1 and Fig. 4.2 report several key points for H(1,15,51,201). I plot the percentage of the total variation explained by the first principal components versus portfolio scale. For all five critical points I select from 1 to 201. The explanatory power of the first principal component will gradually decrease with the increment of matrix dimension. For example, when H is fixed to 1, the proportion falls sharply to a level lower than 6 when the matrix dimension is larger than 100. With the increment of H, the proportion of the first principal component is increasing, the convergence is consistent over all H. I can generate from figures that when the matrix scale is larger than 150, there is little movement in the explanatory power of the first principal component.

I also check the impact of H on the proportion in Fig. 4.3 and Fig. 4.4. For each subplot, I fix the scale of portfolio and plot the proportion versus the number of underlying factors, H. The plots indicate that the principal component explanatory power shows the same trend: it increases with the H factors. It also shows convergence when the factor amount reaches the threshold of around 250. Another interesting thing the plot shows is that with the increment of matrix scale, the 95% distribution is narrowed; the simulation is considered to be more precise.

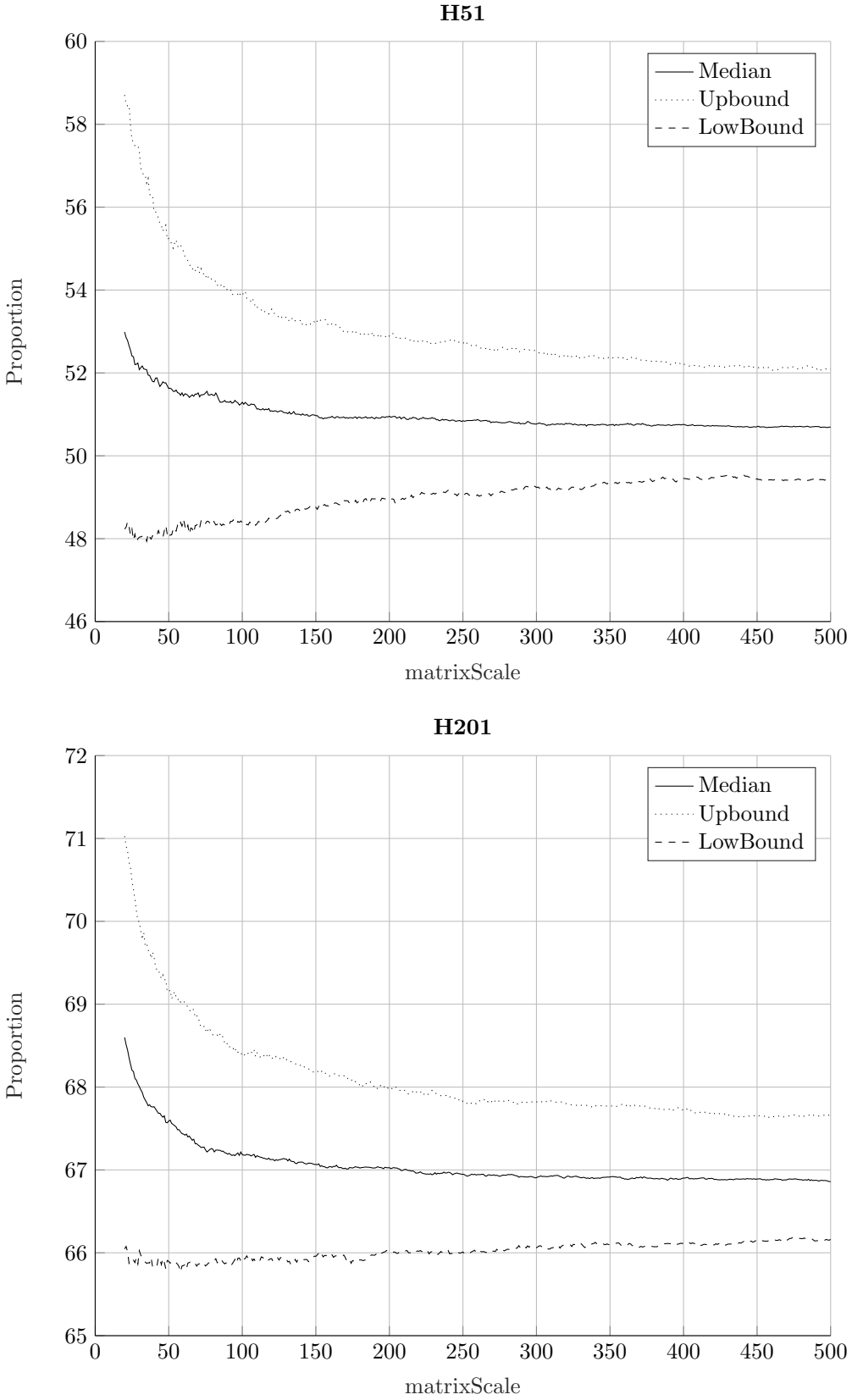
To conclude, I can derive from Fig. 4.1, Fig. 4.2, Fig. 4.3 and Fig. 4.4 that the convergence of eigenvalue is significant from both aspects. So the questions are: 1) is the

Figure 4.1: principal component proportion by matrix Scale, $H = 1, 15$



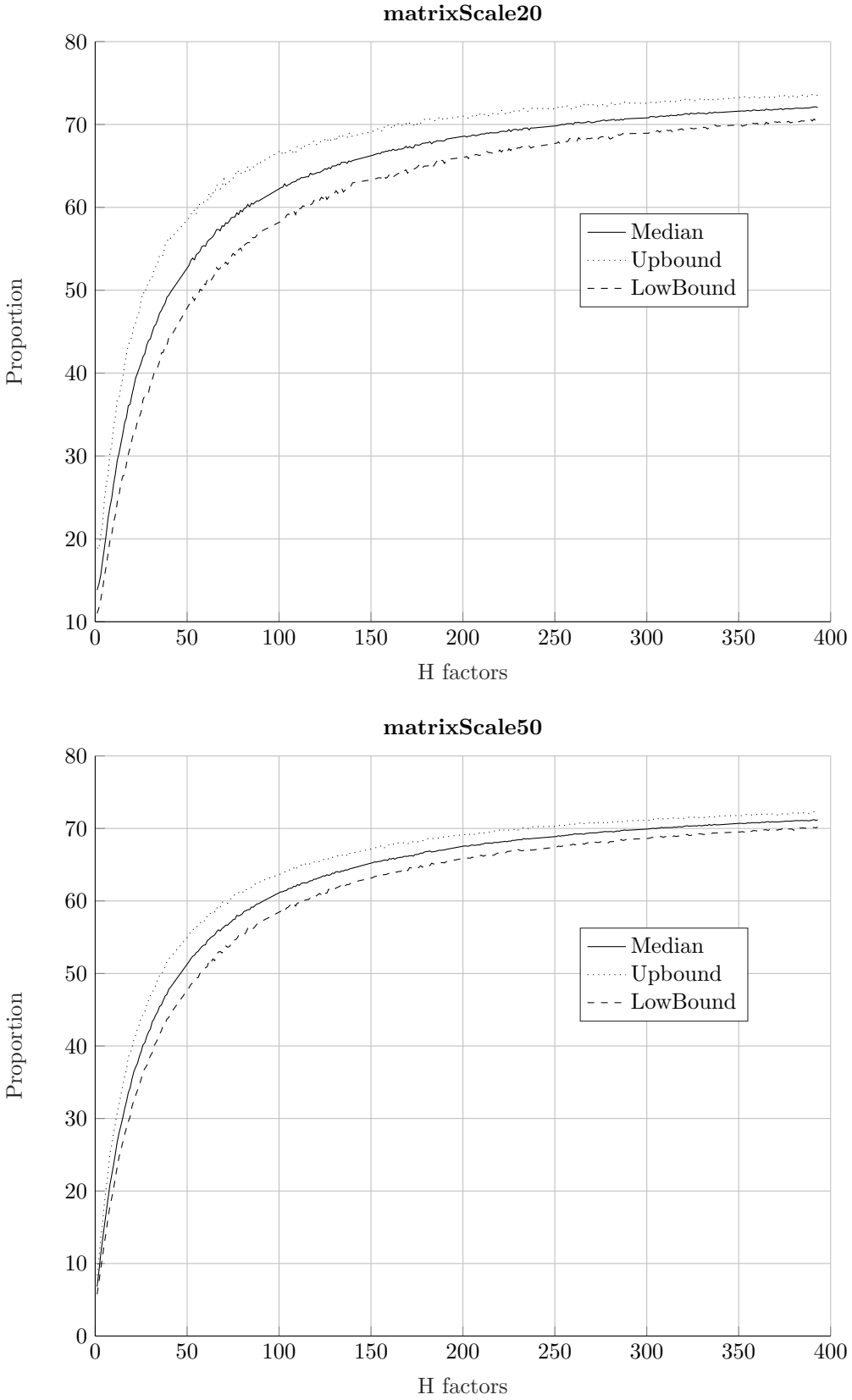
Note: This figure presents the simulation distribution for corresponding H. The bound selected is chosen by 95% confidence interval, median is used to show the point forecast.

Figure 4.2: principal component proportion by matrix Scale, $H = 51, 201$



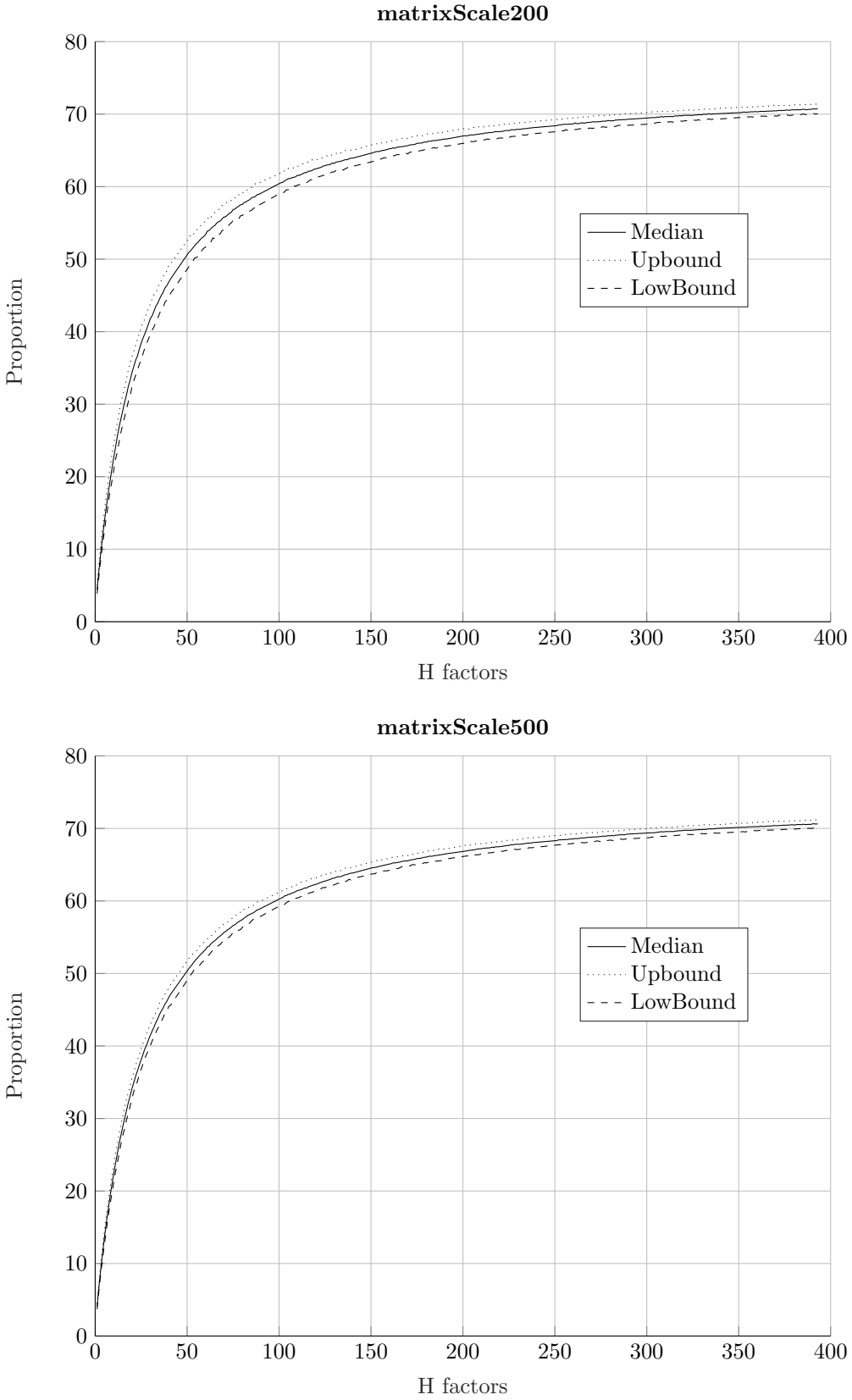
Note: This figure presents the simulation distribution for corresponding H. The bound selected is chosen by 95% confidence interval, median is used to show the point forecast.

Figure 4.3: principal component proportion by H, $Scale = 20, 50$



Note: This figure presents the simulation distribution for corresponding matrix scale. The bound selected is chosen by 95% confidence interval, median is used to show the point forecast.

Figure 4.4: principal component proportion by H, $Scale = 200, 500$



Note: This figure presents the simulation distribution for corresponding matrix scale. The bound selected is chosen by 95% confidence interval, median is used to show the point forecast.

Table 4.1: Top ten long equity list from Apr.1996 to Mar.2020

Date	18-Apr-1996	24-Jan-2002	17-Jul-2008	12-Feb-2015	05-Mar-2020
Rank1	WFCN	BRKbN	DISCBO	AVGOO	DISCBO
Rank2	IBMN	MMMN	GOOGO	SRCLO	CTASO
Rank3	UTXN	TXTN	BRKbN	ALXNO	VRSNO
Rank4	CCIN	ETNN	MAN	SIALO	CMEO
Rank5	PGN	IBMN	QCOMO	AMGNO	AMZNO
Rank6	MON	ACSN	ALXNO	WBAO	PCARO
Rank7	BACN	DOWN	URBNO	NTRSO	NTRSO
Rank8	MSFTO	CSCN	SPLSO	ISRGO	ISRGO
Rank9	ALTRO	CRMN	CMCSAO	XRAYO	MNSTO
Rank10	YHOOO	FLIRO	CINFO	PDCOO	FFIVO

Notes: this table presents the top ten long equity ticker for corresponding time. The window period for matrix is one week, the frequency is 5 minutes for all estimations. Data are reported by TRTH from 1996 to 2020.

convergence also significant in the empirical dataset or does it only occur in simulation study? and 2) how to set a proper measurement for convergence, how to define and predict the threshold for this convergence. To answer these questions, I use S&P500 equity market historical data to implement the eigenstructure analysis mentioned in the previous section.

4.3.3 Historical weight

By implementing eigenvectors analysis to S&P500 historical data from 1996 to 2020, I derive a set of asset weights. The weights are re-estimated weekly, I select five weeks from historical data to present the changes of assets weight over time. Table 4.1 and Table 4.2 show the top 10 stocks for long and short respectively.

In the long stock list Table 4.1 , software/internet companies(IBM, Microsoft,Altair and Yahoo) make up a large portion of the top ten in 1996. Finance(Wells Fargo and Bank of America) is another important company sector in the 1990s. In top 10 companies for 2002, IBM is the only computer science company left. The trend has shifted to manufacturing. The highest weighted company(Berkshire Hathaway) comes from financial services sector, but the rest of the companies mainly belong to manufacturing which

Table 4.2: Top ten short equity list from Apr.1996 to Mar.2020

Date	18-Apr-1996	24-Jan-2002	17-Jul-2008	12-Feb-2015	05-Mar-2020
Rank1	TSCOO	PSXN	DGN	FTRN	CHKN
Rank2	GRMNO	TEGN	ADTN	WINN	JCPN
Rank3	TRIPO	ADTN	PSXN	ODPN	NEN
Rank4	NTAPO	MAN	MJNN	GTN	DNRN
Rank5	BBTN	DPSN	TRIPO	MUN	SWNN
Rank6	CRMN	CBGN	CFNN	AMDN	ODPN
Rank7	GMCRO	MSN	MMIN	HASN	AKSN
Rank8	HANSO	TRIPO	MNKN	ADIN	NBRN
Rank9	FISVO	FTRN	MNSTO	NVLSO	RRCN
Rank10	NTRSO	MMIN	KMIN	BEAMN	QEPN

Notes: this table presents the top ten short equity ticker for corresponding time. The window period for matrix is one week, the frequency is 5 minutes for all estimations.

cover electronic, machine, chemical and so on. With the development of technology, there is a significant trend that the top stocks feature high tech companies(Google/Alphabet, MasterCard, and Qualcomm). In 2015 and 2020, retails and e-commerce companies' stocks are considered as good quality assets.

These trends are significantly presented in our results, with the investors' interest moving from the internet in the late 1990s to manufacturing. Then new companies based on cloud, big data and artificial intelligence are highly ranked after the 2010s. These stocks take the highest level of weight in different periods.

In the short list in Table 4.2, the differences between periods are not as significant as for the long list. But the sectors of companies in the short list are similar to those in the long list over time: financial service companies(Truist Financial, Northern Trust etc.) retailers(Tesco, ODP corporation etc.) are listed in different periods. A unique characteristics of short list is that energy corporations(oil, gas, and electronic) appeared in all periods.

Fig. 4.6 presents the historical Spearman rank correlation coefficient from 1996 to 2020. The window period is the same as the historical weight estimation, coefficients are

re-estimated and recorded weekly. Estimations follow the traditional definition:

$$\rho_{sp} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \tag{4.2}$$

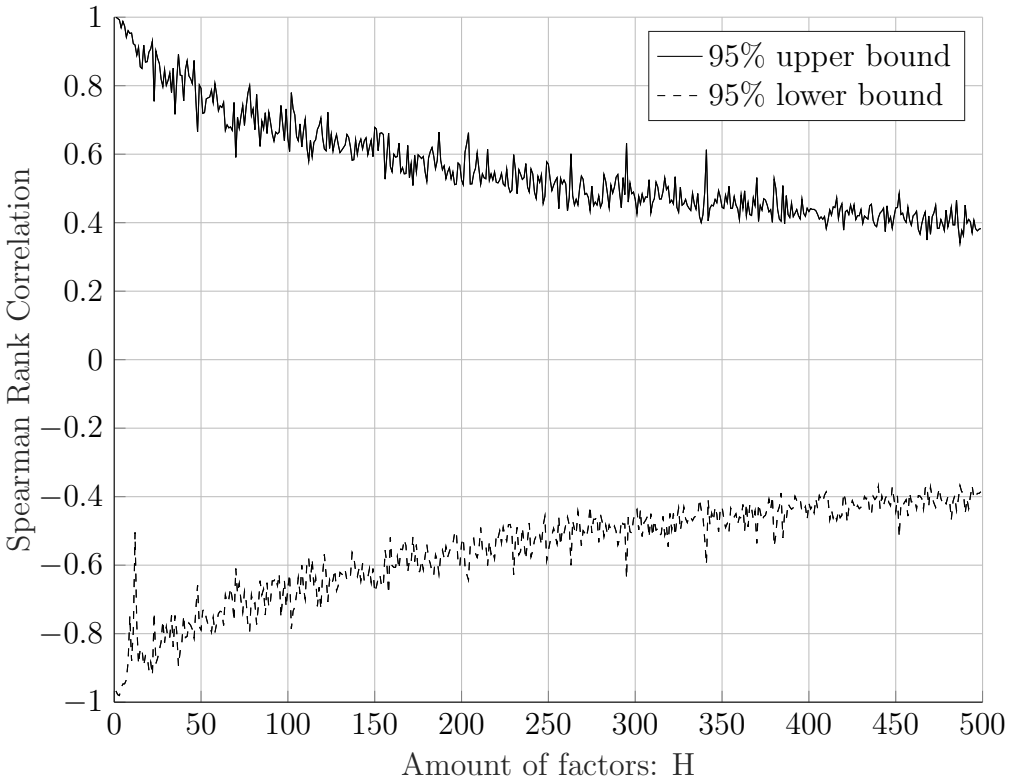
By implementing rank correlation coefficient estimation, I aim to find the stability of the selected portfolio. The Spearman rank correlation coefficient ranges from -1 to 1; a coefficient around 0 indicates that there is little or no relationship between two vectors. Fig. 4.5 presents the simulation results of Spearman rank correlation of a corresponding amount of factors. In my simulation I select factors up to 500, the figure shows that the distribution of rank correlation coefficient is narrower with the increment of factors. When the scale of H is lower than 20, the correlation coefficient nearly covers all possible outcomes from -1 to 1, which indicates that there exists the possibility of perfect correlation. With the increase of factors, the rank correlation between series is turning to relatively weak. The 95% distribution bound is converged to around 0.4 for a dataset with 500 factors.

Fig. 4.6 presents the time series of rank correlation coefficient from 1996 to 2020. Historical coefficients gather around 0, which generates the fact that the equity portfolio is quite unstable, investors need to change the assets in the portfolio to keep high levels of return. But we can still find that the correlation coefficients are time-dependent, the market presents a constant high correlation in several periods: early 1996, 2006, 2008 and early 2020. For other periods such as 2017-2019, the correlation coefficients fluctuate around zero.

4.3.4 Eigenvalue analysis

In this section I introduce the eigenvalue structure based on an empirical data set from the S&P 500 from 2015 to 2020. I focus on the convergence phenomenon of eigenvalue, especially the largest eigenvalue, with the increment of covariance matrix, which indicates

Figure 4.5: Spearman rank correlation simulation



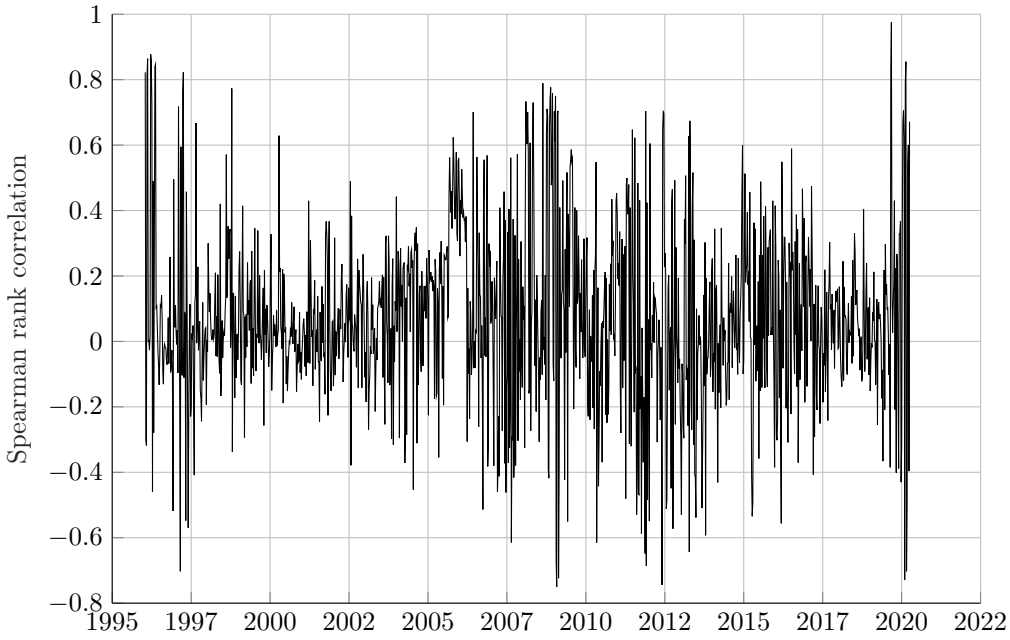
Note: This figure presents the simulated Spearman correlation coefficients bounds for corresponding H. The bound selected is chosen by 95% confidence interval.

the number of stocks taken into consideration.

Fig. 4.7 presents the trend of largest eigenvalue for co-variance matrix corresponding to the increment of matrix scale. I can conclude from the figure that the largest eigenvalue has its peak value when the matrix scale is set to 20. It gradually decrease before the matrix scale reaches 180, which indicate 180 stocks with the highest level of liquidity are taken into calculation. After the decline in fluctuation, the observed eigenvalue increases slightly and converges to a relatively stable value. The critical point for when the largest eigenvalue begin to converge is the key point to decide if the portfolio has begun to stabilize.

With $\Delta_\tau = 5$ minutes, Fig. 4.7 generates the result that the weekly-estimation and

Figure 4.6: Spearman rank correlation from 1996 to 2000

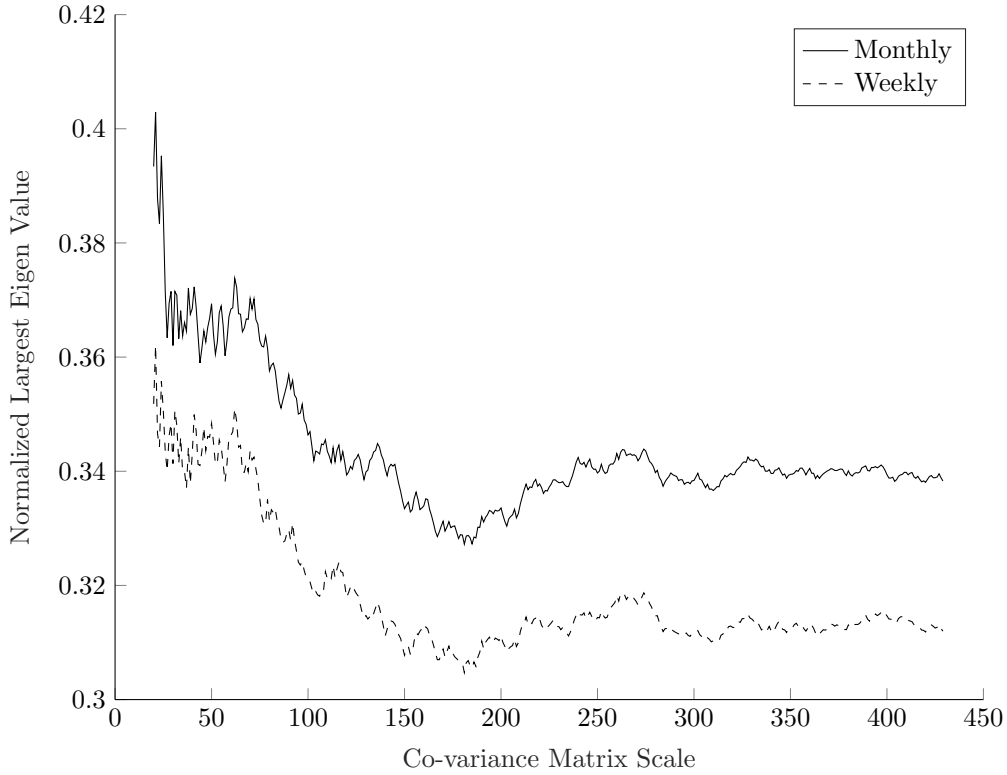


Note: This figure presents the Spearman correlation coefficient for S&P500 market from 1998 to 2020. Correlation coefficients are updated weekly, the frequency is 5 minutes for all estimations.

monthly-estimation show a unique trend and share a similar critical point. Although the level of the two eigenvalues is not the same, it is obvious from the figure that both of them begin their convergence at 180 matrix scale, then flatten after 300.

I apply the analysis through the window period. Generally speaking, the level of eigenvalue structure is time-dependent, as shown by Fig. 4.7, normalized eigenvalue range from 0.302 to 0.405 and get a relative stable value around 0.31 when co-variance matrix scale is large enough. The distribution of eigenvalue structure would change with time, but the trend is uniform in many respects: the first eigenvalue would be high for a small set of equities, then the value would rapidly reduce to a relatively low level and remain stable around this level.

Figure 4.7: The largest eigenvalue for S&P 500 portfolio at 31th Jan 2015,Δ_τ= 5 minutes

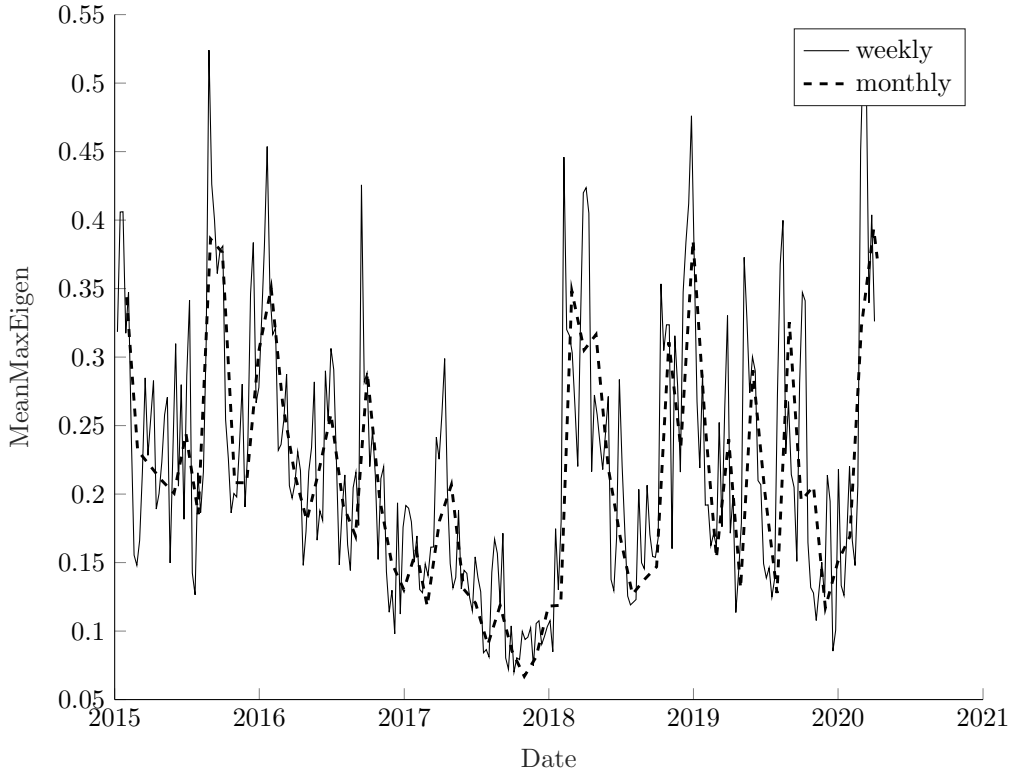


Note:This figure presents the average normalized largest eigen value versus the matrix dimension, two estimation window periods are set: weekly and monthly. All data come from the TRTH at 31th Jan 2015, data frequency is fixed to every 5 minutes.

Fig. 4.8 reports the average of the first eigenvalue over time. I can conclude from the figure that the mean level of 1st eigenvalue could be as low as 0.04 and the peak value could reach 0.53. The weekly-estimation eigenvalue structure shows the same movement trend as the monthly-estimation value, although it fluctuate significantly more compared with the monthly-estimation result. The synergy movement of two estimation period results are consistent in two respects: firstly for a chosen date, the two first eigenvalues show the same trend corresponding to portfolio dimension. Secondly the average values are consistent over time.

In addition to the largest eigenvalue structure, I generate the percentages of total variation explained by principal components corresponding to the first four eigenvalues in Fig. 4.9. The scale of matrix is time-dependent; I follow the rule to ensure the chosen scale

Figure 4.8: Mean 1st eigenvalue when Δ_τ= 5 minutes



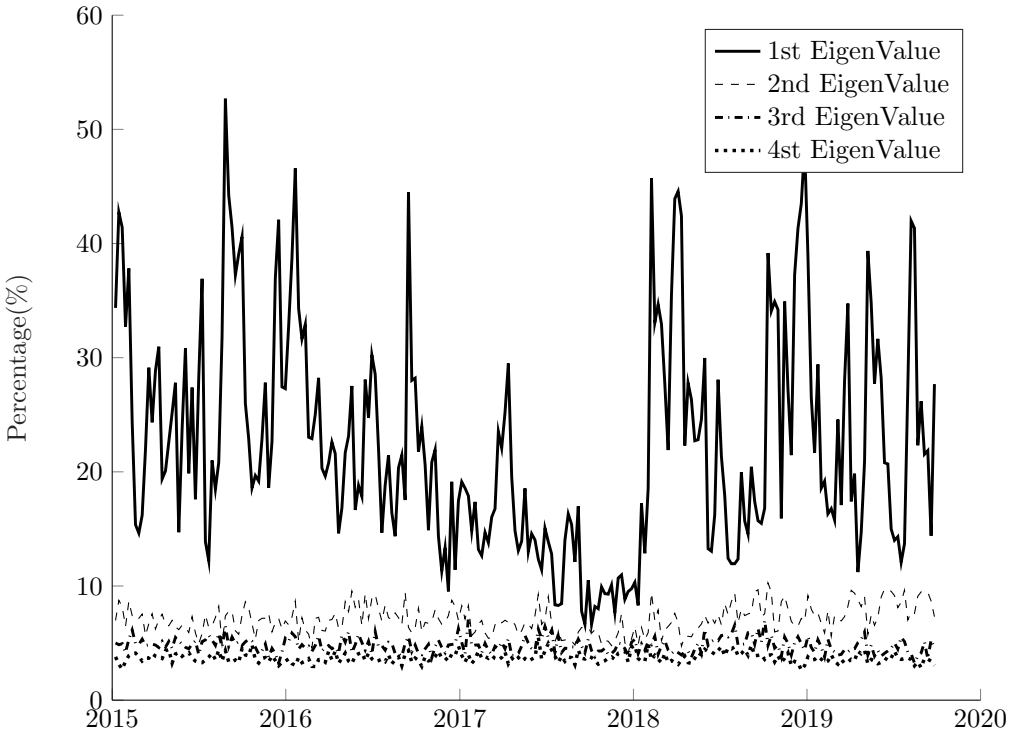
Note:This figure presents the average normalized largest eigen value over time series., two estimation window periods are set: weekly and monthly. All data come from the TRTH from Jan 2015 to Apr 2020.

has been converged and keep constant to a stable level. The measurement is introduced in a later section. I conclude from the figure that the first principal fluctuates sharply, but 2nd-4th principal keep at a relatively stable level over time. The difference between 2nd-4th principal components is not as significant as between 1st principal component and others.

4.3.5 Eigenvector Analysis

In this section, I introduce the definition for 'critical point' to measure the convergence of eigenvalue structure. Before moving to the empirical results, I firstly generate how I define the 'convergence procedure'. To better understand the eigenvalue structure, I

Figure 4.9: Percentage of the Total Variation Explained by Principal Components



Note:This figure presents the percentage of variation explained by the principal components corresponding to the 1st-4th eigenvalues of S&P 500 during January 2015 - April 2020.

would derive it from aspects of the eigenvector.

$$\theta = cos^{-1}\left(\frac{\mathrm{v}_t \cdot \sum^n \frac{\mathrm{v}_{tail}}{n}}{\| \mathrm{v}_t \| \cdot \| \sum^n \frac{\mathrm{v}_{tail}\|}{n}}\right) \tag{4.3}$$

Where v_t indicates the eigenvector correspond to first principal eigenvalue for the co-variance matrix scale t. v_{tail} indicates the eigenvector when co-variance matrix scale reach its limit. Such as for a dataset contains 450 equities, I would select the the matrix scale from 440 to 450.

Angle θ is used to determine if the eigenvalue structure is beginning to converge.Fig. 4.10 presents an example for θ: weekly- and monthly estimation, θ rapidly reduces to a low level. I can conclude from figure that compared with weekly-estimation eigenvectior an-

gle, the monthly-estimation results are smoother with a lower level of volatility. This characteristic is consistent over time series.

I define two thresholds : 0.05 and 0.1; the angle smaller than these thresholds is considered to be 'stable' for corresponding eigenvalue.In Fig. 4.11, I present the matrix scale over time for both critical points. I can conclude from thefigure that although both dimensions fluctuate with time, they share a similar trend. When θ critical point is set to be 0.05, the matrix scale ranges from 50 to 262. Meanwhile for 0.1, the matrix scale could be as low as 20 to be considered stable, the max scale for this critical point is 191.

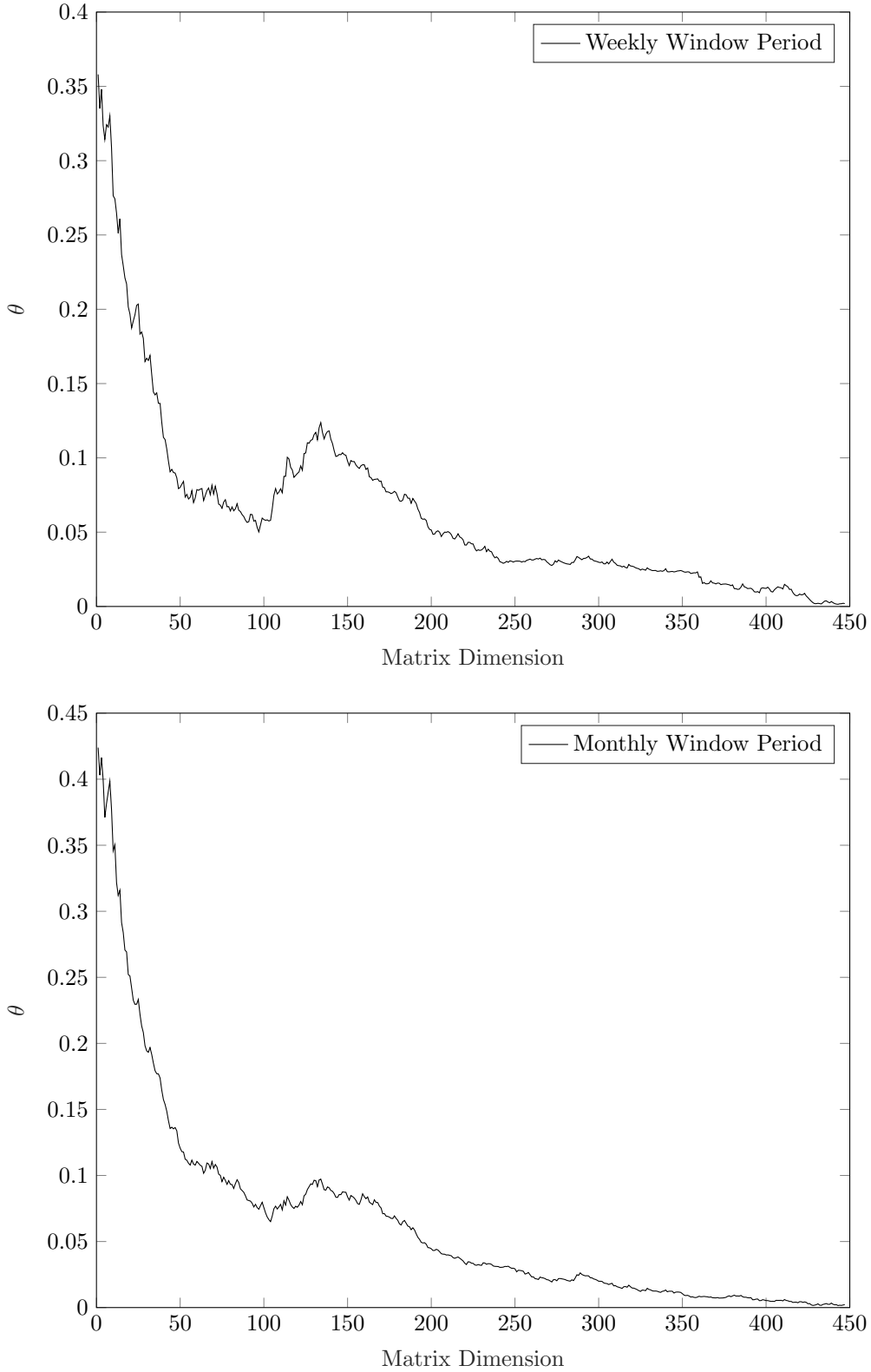
Fig. 4.12 shows the statistical distribution for both critical points. I can conclude from the figure that when I set the threshold to 0.05, a relatively high standard, the peak number of equities is around 170. Although the scale distribution ranges from 20 to 270, most of the time series are considered to enter into a stable process when the matrix scale is promoted to 150-200. When I set the threshold to 0.1, the distribution of critical point is narrower compared with 0.05 threshold. Most eigenstructures are considered to be stable when the matrix scale is between 20 and 60, the peak dimensions are between 30 and 40. Only small number of matrix scale exceed 100.

4.4 Conclusion

In this chapter, I have focused on the eigenvalue and eigenvector structure. Previous literature noted the phenomenon of first principal component convergence, but due to the limit of the data set scale, the eigenvalue trend over time series was not suspected. I select the equity market as the empirical data set in this study, S&P500 historical data back to the 1990s was collected and cleaned to drive the time series for eigenvalue and eigenvector structure.

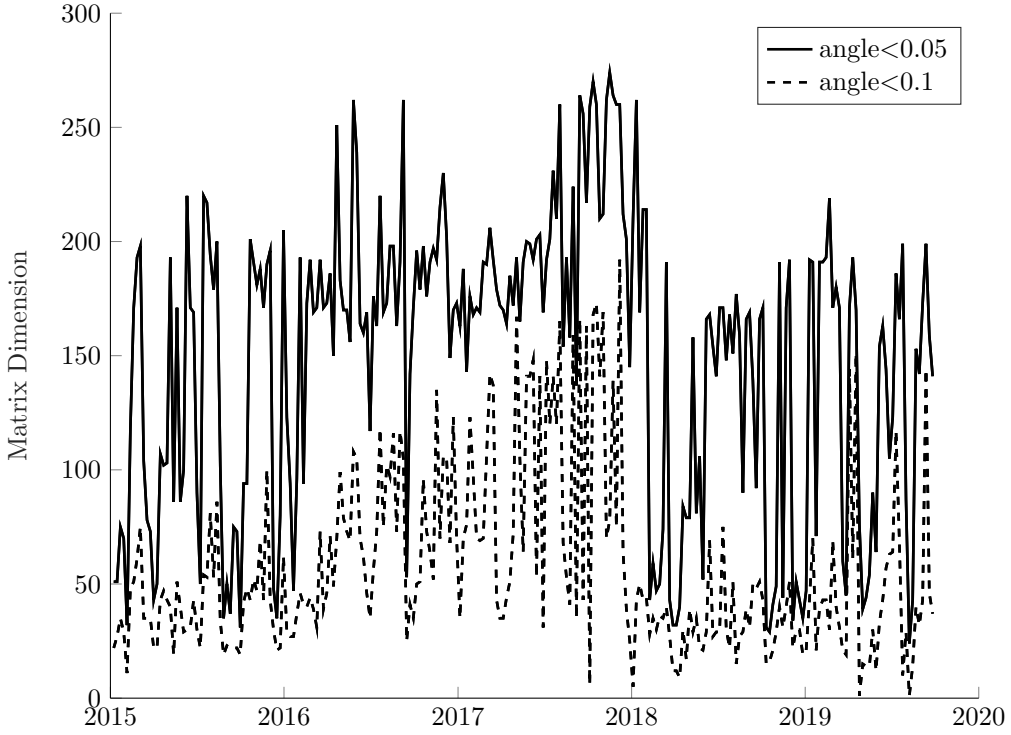
I verify that the effects of window period and data frequency in principal component performance are not significant enough to sway the conclusion. Window periods of co-

Figure 4.10: θ in 22-OCT-2015



Note: This figure presents the θ follow 4.3 at 22-OTC-2015. The eigenvector selected correspond to first principal component.

Figure 4.11: Matrix Dimension For Two critical point: 0.1 and 0.05



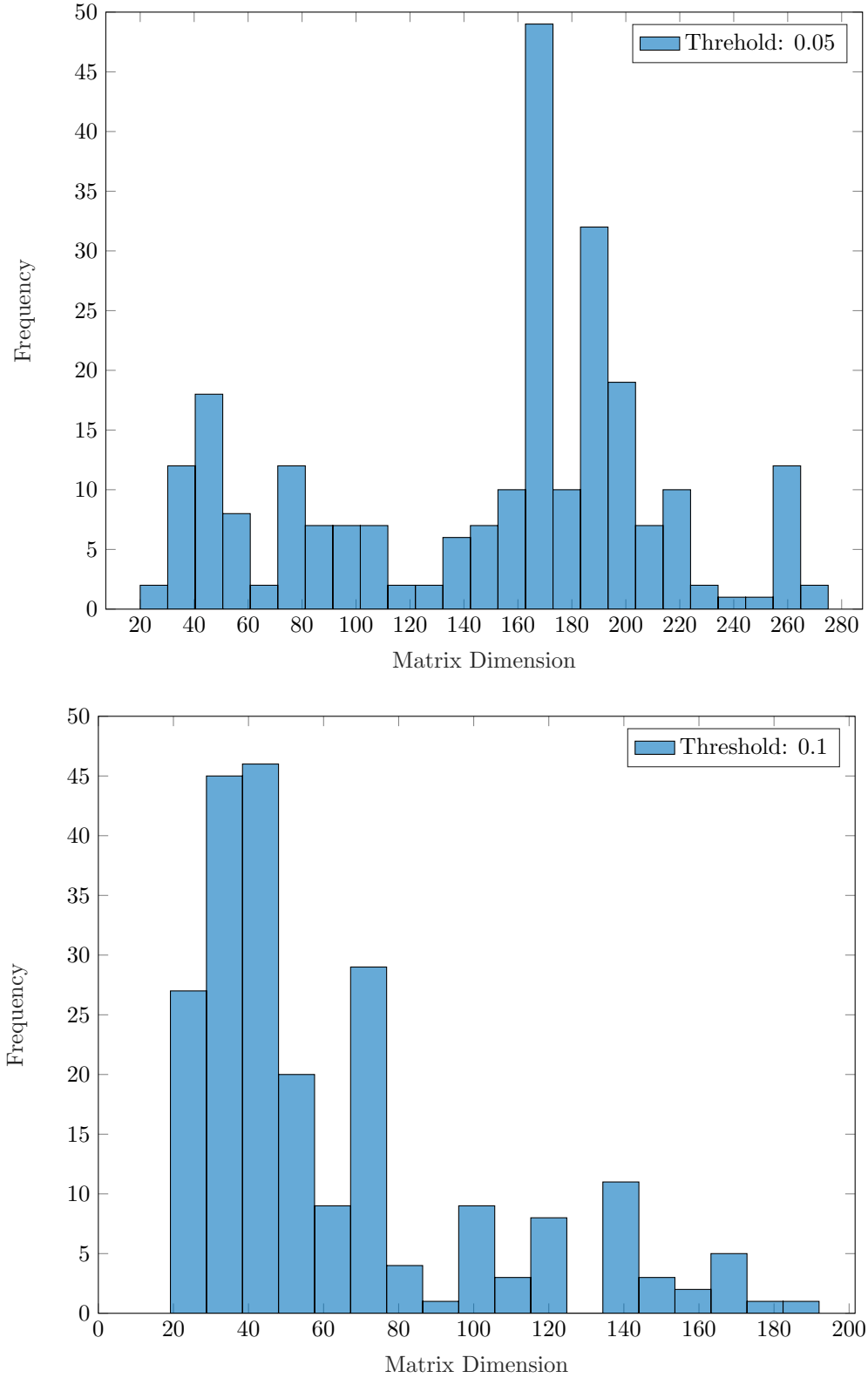
Note: This figure presents the critical dimension for the largest eigen value convergence over time series from Jan 2015 to Apr 2020. The data are collected from TRTH.

variance matrix used are one week and one month. Compared with weekly estimated co-variance matrix, the eigenvalue structures of monthly estimated co-variance matrix are smoother.

I use angle of eigenvector as the measurement to decide if eigenvalue is stable. The angle θ lower than the pre-determined critical point is considered to be in 'stable' process, I select two different critical points for θ to generate the statistical summary: 0.1 and 0.05. I generate from the summary that the eigenvalue structure will fluctuate slightly when it firstly reduces to the threshold 0.1, but when it reduces to levels lower than 0.05, the principal component will become extreme stable after that point.

Our empirical study organizes massive high-frequency time series from S$P500 his-

Figure 4.12: Critical points distribution



Note: This figure presents the scale distribution for critical point. Bar height indicates the frequency of matrix scale for certain portfolio scale, x-axis show the corresponding amount of equities in co-variance matrix.

torical data. The matrix scale is increased to more than 400 equities, which allows us to observe the trend of eigenvalue. For each selected date, I firstly sort the stock by liquidity level, then gradually increase the number of equities taken into the co-variance matrix construction. The distribution of the matrix scale indicates that the critical point of scale gather from 30 to 50 when the threshold is 0.1 and gathers from 150-200 when the threshold is set to 0.05.

In addition to the empirical study, I also apply Monte Carlo method to simulate the eigenvalue structure under idea market assumption. I define two dimensions to verify the consistent of eigenvalue and eigenvector convergence behavior: 1) determine factors of equity time series 2) the matrix scale. Result verify that the simulation summary follow the empirical study in both matrix scale distribution and the trend of eigenvalue. Simulation and empirical co-variance matrix converge their principal component at a similar scale.

Further study could focus on the structure of other principal components. I could generate from the previous study that, although as not significant as the first principal analysis, 2nd-4th eigenvalue also show a unique trend across time series. In my thesis, I use the eigenvectors' angle as measurement to decide if the eigenvalue is converged, more measurement methodologies could be raised to make the procedure clearer.

To conclude, this study focuses on the structure of principal components, I expand the data set from relatively small (around 100 equities) to larger than 400 in our study. The percentage of total variation explained by principal components is tested over time series and it is tested to be stable when the equity amount reaches threshold. This study has provided a deep insight into the principal component structure analysis and aim to provide guidance to portfolio management.

Chapter 5

Conclusions

5.1 Summary and remarks

This thesis provides an extensive analysis of the use of eigensystems and eigenfunctions in finance. Eigenfunctions within specific closed sets, such as those that constrain bounds on specific eigenvalues (such as the smallest being positive) provide useful guidance for consistent pricing in options and derivatives markets. Analysis of the random variation of eigenvalues and largest eigenvectors provides insight into the stochastic structure of large cross sections of equities.

In both cases the problem is quite similar, answering the question how many redundant assets are needed to correctly price the object of interest? In FX this is the number of leg currencies to narrow down the bounds for the implied correlation function against delta (hence the eigenfunction constraining the smallest eigenvalue to preserve positive definiteness). In contrast for equity the largest eigenvalue and the vector space provide an answer to an important question, for approximate factors extracted from asset returns how many assets should we include to extract the factor structure? This is a question which has not been addressed in the core literature on the topic.

Supplementing the primary objectives is a series of diagnostic tests and applications

of machine learning to deliver better predictions. In FX the eigensystem analysis shrinks the forecasting problem for the entire strike price and maturity surface to a single time series, bounded by the eigenvalue limits from the leg currencies. Through a combination of simulation and empirical data analysis the equity analysis again provides a low bound on the number of assets needed to be included, using the stability of the eigenvector of the largest eigenvalue as a guide. I show using rank correlations of the weightings of assets that the resulting portfolio is quite volatile for substantial periods, indicating that careful attention is needed when conducting studies on realized principal components. This sheds new light on the varying results from the literature on realized principal components and suggests that a lot of variation is driven by the asset dimension choice.

5.2 Further Research

This thesis can be improved in many respects. For the FX market, the predictive power is not consistent over currency pairs, such as HUF, and it is hard to provide accurate predictions for their option trading. Due to the limitation of currency pairs amount in leg journey, the error between minimum eigenvalue and market quote is hard to predict, which lead to a further spread between calculated implied volatility and market quote. For machine learning, it is worth importing more methodologies to compare with the regression tree. Although the regression tree is verified by the literature to be the most reliable methodology in prediction, it would provide more support to our prediction if more methods could be imported.

References

Ait-Sahalia, Y. et al. (2008). Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics 36*(2), 906–937. 33

Aït-Sahalia, Y., D. Amengual, and E. Manresa (2015). Market-based estimation of stochastic volatility models. *Journal of Econometrics 187*(2), 418–435. 33

Ballotta, L., G. Deelstra, and G. Rayée (2017). Multivariate fx models with jumps: Triangles, quantos and implied correlation. *European Journal of Operational Research 260*(3), 1181–1199. 34

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies 9*(1), 69–107. 39

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics 94*(1-2), 181–238. 33

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy 81*(3), 637–654. 30

Bossens, F., G. Rayée, N. S. Skantzos, and G. Deelstra (2010). Vanna-volga methods applied to fx derivatives: from theory to market practice. *International Journal of Theoretical and Applied Finance 13*(08), 1293–1324. 31

Butaru, F., Q. Chen, B. Clark, S. Das, A. W. Lo, and A. Siddique (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance 72*, 218–239. 76

Campa, J. M. and P. K. Chang (1998). The forecasting ability of correlations implied in foreign exchange options. *Journal of International Money and Finance 17*(6), 855–880. 34

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies 21*(4), 1509–1531. 76

Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. *Econometrica: Journal of the Econometric Society*, 1305–1323. 115

Chen, D., P. A. Mykland, and L. Zhang (2019). The five trolls under the bridge: Principal component analysis with asynchronous and noisy high frequency data. *Journal of the American Statistical Association*, 1–18. 115, 116

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society 53*(2), 385–407. 29, 32, 33, 36

Da Fonseca, J., M. Grasselli, and F. Ielpo (2014). Estimating the wishart affine stochastic correlation model using the empirical characteristic function. *Studies in Nonlinear Dynamics & Econometrics 18*(3), 253–289. 33

Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine learning 40*(2), 139–157. 76

Dovonon, P., A. Taamouti, and J. Williams (2021). Testing the eigenvalue structure of spot and integrated covariance. *Journal of Econometrics*. 113, 114, 115, 116

Du, W., A. Tepper, and A. Verdelhan (2017). Deviations from covered interest rate parity. Technical report, National Bureau of Economic Research. 34

Duffie, D. and P. Glynn (2004). Estimation of continuous-time markov processes sampled at random time intervals. *Econometrica 72*(6), 1773–1808. 33

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376. 33

Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance 75*(3), 1327–1370. 77

Freund, Y. (1995). Boosting a weak learning algorithm by majority. *Information and computation 121*(2), 256–285. 77

Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies 33*(5), 2326–2377. 76

Friedman, J., T. Hastie, R. Tibshirani, et al. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Annals of statistics 28*(2), 337–407. 77

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232. 77

Garman, M. B. and S. W. Kohlhagen (1983). Foreign currency option values. *Journal of International Money and Finance 2*(3), 231–237. 30, 31, 32, 47

Giglio, S. and D. Xiu (2019). Asset pricing with omitted factors. *Chicago Booth Research Paper* (16-21). 76

Green, J., J. R. Hand, and X. F. Zhang (2013). The supraview of return predictive signals. *Review of Accounting Studies 18*(3), 692–730. 75

Green, J., J. R. Hand, and X. F. Zhang (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies 30*(12), 4389–4436. 76

Gu, S., B. Kelly, and D. Xiu (2020a). Autoencoder asset pricing models. *Journal of Econometrics*. 75

Gu, S., B. Kelly, and D. Xiu (2020b). Empirical asset pricing via machine learning. *The Review of Financial Studies 33*(5), 2223–2273. 85

Hansen, L. P. and R. J. Hodrick (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of political economy 88*(5), 829–853. 31

Harvey, C. R. and Y. Liu (2015). Lucky factors. *Available at SSRN 2528780*, 980. 76

Heaton, J. B., N. G. Polson, and J. H. Witte (2018). Deep learning in finance. 76

Heston, S. L. (1993a). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies 6*(2), 327–343. 33

Heston, S. L. (1993b). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*(2), 327–343. 38

Hinton, G. E., S. Osindero, and Y.-W. Teh (2006). A fast learning algorithm for deep belief nets. *Neural computation 18*(7), 1527–1554. 76

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The journal of finance 42*(2), 281–300. 33

Hutchinson, J. M., A. W. Lo, and T. Poggio (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance 49*(3), 851–889. 75

Kelly, B., S. Pruitt, and Y. Su (2017a). Some characteristics are risk exposures, and the rest are irrelevant. *Unpublished Manuscript, University of Chicago*. 76

Kelly, B., S. Pruitt, and Y. Su (2017b). Some characteristics are risk exposures, and the rest are irrelevant. *Social Science Research Network*. 77

Khandani, A. E., A. J. Kim, and A. W. Lo (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance 34*(11), 2767–2787. 76

Kozak, S., S. Nagel, and S. Santosh (2017). Shrinking the cross section. Technical report, National Bureau of Economic Research. 76

Moritz, B. and T. Zimmermann (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. *Available at SSRN 2740751*. 76

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of financial economics 63*(1), 3–50. 33, 38, 39

Rapach, D. E., J. K. Strauss, and G. Zhou (2013). International stock return predictability: What is the role of the united states? *The Journal of Finance 68*(4), 1633–1662. 75

Rubinstein, M. (1976). The valuation of uncertain income streams and the pricing of options. *The Bell Journal of Economics 7*(2), 407–425. 31, 32

Schapire, R. E. (1990). The strength of weak learnability. *Machine learning 5*(2), 197–227. 77

Shreve, S. E. (2004). Stochastic calculus for finance ii: Continuous-time models. *Springer-Verlag New York Inc. 1*, 550. 31, 32

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics 102*(1), 111–141. 33

Sirignano, J., A. Sadhwani, and K. Giesecke (2016). Deep learning for mortgage risk. *arXiv preprint arXiv:1607.02470*. 76

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of financial economics 5*(2), 177–188. 32

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508. 76

Yao, J., Y. Li, and C. L. Tan (2000). Option price forecasting using neural networks. *Omega 28*(4), 455–466. 75

Yu, J. (2007). Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the chinese yuan. *Journal of Econometrics 141*(2), 1245–1280. 33

Appendix A

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
AAPL_O	02-Jan-1996	09-Apr-2020	3,669,526	ALXN_O	28-Feb-1996	09-Apr-2020	2,145,565
AA_N	02-Jan-1996	09-Apr-2020	2,221,518	AMAT_O	02-Jan-1996	09-Apr-2020	2,821,397
ABBV_N	02-Jan-2013	09-Apr-2020	712,387	AMD_N	02-Jan-1996	09-Apr-2020	1,912,969
ABC_N	30-Aug-2001	09-Apr-2020	1,807,644	AME_N	02-Jan-1996	09-Apr-2020	1,889,937
ABT_N	02-Jan-1996	09-Apr-2020	2,268,216	AMGN_O	02-Jan-1996	09-Apr-2020	2,758,504
ACE_N	02-Jan-1996	14-Jan-2016	1,441,181	AMG_N	21-Nov-1997	09-Apr-2020	1,879,610
ACN_N	03-Jan-1996	09-Apr-2020	1,896,081	AMP_N	02-Jan-1996	09-Apr-2020	1,593,161
ACS_N	02-Jan-1996	05-Feb-2010	947,814	AMT_N	02-Jan-1996	09-Apr-2020	1,946,736
ACT_N	24-Jan-2013	09-Apr-2020	261,264	AMZN_O	15-May-1997	09-Apr-2020	3,193,488
ADBE_O	02-Jan-1996	09-Apr-2020	2,618,703	ANF_N	26-Sep-1996	09-Apr-2020	2,067,576
ADI_N	02-Jan-1996	09-Apr-2020	1,585,531	ANR_N	19-Mar-1997	17-Jul-2015	1,118,753
ADM_N	02-Jan-1996	09-Apr-2020	2,143,481	ANTM_N	03-Dec-2014	09-Apr-2020	524,012
ADP_O	21-Oct-2008	09-Apr-2020	1,381,458	AN_N	02-Jan-1996	09-Apr-2020	2,110,813
ADSK_O	01-May-1996	09-Apr-2020	2,403,810	AON_N	01-Dec-2009	09-Apr-2020	1,014,779
ADS_N	08-Jun-2001	09-Apr-2020	1,721,158	APA_N	02-Jan-1996	09-Apr-2020	2,144,448
ADT_N	02-Jan-1996	09-Apr-2020	600,834	APC_N	02-Jan-1996	08-Aug-2019	2,053,934
AEE_N	02-Jan-1998	09-Apr-2020	2,017,015	APD_N	02-Jan-1996	09-Apr-2020	2,124,120
AEP_N	02-Jan-1996	09-Apr-2020	2,152,151	APH_N	02-Jan-1996	09-Apr-2020	1,915,150
AES_N	16-Oct-1996	09-Apr-2020	2,095,153	APOL_O	02-Jan-1996	01-Feb-2017	1,946,639
AET_N	02-Jan-1996	28-Nov-2018	2,036,869	ARG_N	02-Jan-1996	23-May-2016	1,522,112
AFL_N	02-Jan-1996	09-Apr-2020	2,114,566	ATI_N	02-Jan-1996	09-Apr-2020	2,098,614
AGN_N	02-Jan-1996	09-Apr-2020	2,077,095	AVB_N	05-Jun-1998	09-Apr-2020	1,859,245
AIG_N	02-Jan-1996	09-Apr-2020	2,257,675	AVGO_O	27-Sep-2000	09-Apr-2020	1,327,634
AIV_N	02-Jan-1996	09-Apr-2020	1,894,877	AVP_N	02-Jan-1996	03-Jan-2020	2,164,158
AIZ_N	02-Jan-1996	09-Apr-2020	1,638,713	AVY_N	02-Jan-1996	09-Apr-2020	2,084,055
AKAM_O	29-Oct-1999	09-Apr-2020	2,304,276	AXP_N	02-Jan-1996	09-Apr-2020	2,280,966
AKS_N	02-Jan-1996	13-Mar-2020	1,952,840	AYE_N	01-Oct-1997	25-Feb-2011	1,097,592
ALLE_N	02-Dec-2013	09-Apr-2020	617,439	AZO_N	02-Jan-1996	09-Apr-2020	2,060,039
ALL_N	02-Jan-1996	09-Apr-2020	2,244,917	A_N	23-May-1996	09-Apr-2020	2,040,477
ALTR_O	02-Jan-1996	09-Apr-2020	2,345,175	BAC_N	02-Jan-1996	09-Apr-2020	2,309,659

Notes: this table presents the ticker list used in S&P500 equity study. The 'Ticker' column shows the ticker for equity. The 'Start Date' column shows the first recorded date for trades in database of my study. The 'End Date' column shows the last recorded date for trades in database of my study. The 'Number of Quotes' column shows the number of trades in my database for corresponding equity.

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
BAX_N	02-Jan-1996	09-Apr-2020	2,171,172	CLF_N	02-Jan-1996	09-Apr-2020	1,839,530
BA_N	02-Jan-1996	09-Apr-2020	2,274,856	CLX_N	02-Jan-1996	09-Apr-2020	2,171,045
BBBY_O	02-Jan-1996	09-Apr-2020	2,459,181	CL_N	02-Jan-1996	09-Apr-2020	2,231,253
BBT_N	02-Jan-1996	06-Dec-2019	1,987,627	CMA_N	02-Jan-1996	09-Apr-2020	2,144,472
BBY_N	02-Jan-1996	09-Apr-2020	2,190,650	CMCSA_O	02-Jan-1996	09-Apr-2020	2,428,563
BCR_N	02-Jan-1996	29-Dec-2017	1,813,253	CME_O	14-Jul-2008	09-Apr-2020	1,371,192
BDK_N	02-Jan-1996	12-Mar-2010	1,105,923	CMG_N	26-Jan-2006	09-Apr-2020	1,367,851
BDX_N	02-Jan-1996	09-Apr-2020	2,134,359	CMI_N	08-Sep-1997	09-Apr-2020	1,628,658
BEAM_N	04-Oct-2011	09-Apr-2020	251,240	CMS_N	02-Jan-1996	09-Apr-2020	2,041,274
BEN_N	02-Jan-1996	09-Apr-2020	2,121,540	CNP_N	01-Oct-2002	09-Apr-2020	1,708,266
BFb_N	02-Jan-1996	09-Apr-2020	1,982,935	CNX_N	30-Apr-1999	09-Apr-2020	1,836,060
BHI_N	02-Jan-1996	03-Jul-2017	1,949,428	COF_N	02-Jan-1996	09-Apr-2020	2,126,940
BIG_N	18-Aug-2006	09-Apr-2020	1,333,212	COG_N	02-Jan-1996	09-Apr-2020	1,882,393
BIIB_O	13-Nov-2003	09-Apr-2020	2,013,304	COH_N	05-Oct-2000	30-Oct-2017	1,587,334
BJS_N	02-Jan-1996	28-Apr-2010	1,162,251	COL_N	02-Jan-1996	26-Nov-2018	1,931,153
BK_N	02-Jan-1996	09-Apr-2020	2,211,943	COP_N	19-Nov-1996	09-Apr-2020	1,760,202
BLK_N	02-Jan-1996	09-Apr-2020	1,781,945	COST_O	06-Feb-1997	09-Apr-2020	2,529,781
BLL_N	02-Jan-1996	09-Apr-2020	2,019,069	COV_N	14-Mar-2001	26-Jan-2015	800,622
BMC_O	26-Oct-2009	10-Sep-2013	480,119	CPB_N	02-Jan-1996	09-Apr-2020	2,187,649
BMS_N	02-Jan-1996	11-Jun-2019	1,938,625	CPWR_O	02-Jan-1996	15-Dec-2014	1,782,873
BMY_N	02-Jan-1996	09-Apr-2020	2,294,662	CRM_N	02-Jan-1996	09-Apr-2020	1,566,365
BRCM_O	17-Apr-1998	29-Jan-2016	2,104,917	CSCO_O	03-Jan-1996	09-Apr-2020	3,224,065
BRKb_N	09-May-1996	09-Apr-2020	1,693,073	CSC_N	02-Jan-1996	31-Mar-2017	1,899,394
BSX_N	02-Jan-1996	09-Apr-2020	2,188,840	CSX_N	02-Jan-1996	09-Apr-2020	1,926,547
BTU_N	22-May-2001	09-Apr-2020	1,675,096	CTAS_O	02-Jan-1996	09-Apr-2020	2,315,710
BWA_N	02-Jan-1996	09-Apr-2020	1,961,159	CTL_N	02-Jan-1996	09-Apr-2020	2,072,991
BXP_N	18-Jun-1997	09-Apr-2020	1,888,179	CTSH_O	19-Jun-1998	09-Apr-2020	2,229,101
CAG_N	02-Jan-1996	09-Apr-2020	2,180,497	CTXS_O	02-Jan-1996	09-Apr-2020	2,535,043
CAH_N	02-Jan-1996	09-Apr-2020	2,143,968	CVC_N	07-Dec-1999	21-Jun-2016	1,549,302
CAM_N	02-Jan-1996	01-Apr-2016	1,674,898	CVG_N	13-Aug-1998	05-Oct-2018	1,826,766
CAT_N	02-Jan-1996	09-Apr-2020	2,250,247	CVH_N	03-Apr-1998	06-May-2013	1,128,300
CA_O	28-Apr-2008	05-Nov-2018	1,240,887	CVS_N	16-Oct-1996	09-Apr-2020	2,153,645
CBE_N	02-Jan-1996	30-Nov-2012	1,374,414	CVX_N	10-Oct-2001	09-Apr-2020	1,807,457
CBG_N	07-Nov-1997	19-Mar-2018	1,368,851	C_N	02-Jan-1996	09-Apr-2020	2,271,779
CBS_N	02-Jan-1996	04-Dec-2019	1,567,657	DAL_N	02-Jan-1996	09-Apr-2020	2,075,944
CB_N	02-Jan-1996	09-Apr-2020	2,175,400	DD_N	02-Jan-1996	09-Apr-2020	2,132,829
CCE_N	02-Jan-1996	06-Nov-2018	1,942,843	DELL_O	02-Jan-1996	29-Oct-2013	2,015,701
CCI_N	02-Jan-1996	09-Apr-2020	2,023,245	DE_N	02-Jan-1996	09-Apr-2020	2,187,527
CCL_N	02-Jan-1996	09-Apr-2020	2,147,111	DFS_N	02-Jan-1996	09-Apr-2020	1,608,753
CEG_N	03-May-1999	12-Mar-2012	1,164,994	DF_N	02-Jan-1996	12-Nov-2019	1,938,649
CELG_O	02-Jan-1996	20-Nov-2019	2,409,271	DGX_N	14-Jan-1997	09-Apr-2020	1,952,166
CEPH_O	03-Jan-1996	14-Oct-2011	1,341,735	DG_N	02-Jan-1996	09-Apr-2020	1,848,020
CERN_O	02-Jan-1996	09-Apr-2020	2,287,061	DHI_N	02-Jan-1996	09-Apr-2020	1,978,307
CFN_N	09-Feb-1996	17-Mar-2015	685,796	DHR_N	02-Jan-1996	09-Apr-2020	2,066,197
CF_N	02-Jan-1996	09-Apr-2020	1,924,500	DISCA_O	21-Jul-2005	09-Apr-2020	1,707,163
CHK_N	02-Jan-1996	09-Apr-2020	1,974,371	DISCB_O	21-Jul-2005	09-Apr-2020	825,438
CHRW_O	15-Oct-1997	09-Apr-2020	2,143,100	DIS_N	02-Jan-1996	09-Apr-2020	2,296,810
CIEN_N	23-Dec-2013	09-Apr-2020	617,016	DLPH_N	17-Nov-2011	09-Apr-2020	814,834
CINF_O	02-Jan-1996	09-Apr-2020	2,184,680	DLTR_O	02-Jan-1996	09-Apr-2020	2,333,572
CI_N	02-Jan-1996	09-Apr-2020	2,199,521	DNB_N	02-Jan-1996	08-Feb-2019	1,878,981

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
DNR_N	08-May-1997	09-Apr-2020	1,789,199	FIS_N	01-Feb-2006	09-Apr-2020	1,388,771
DOV_N	02-Jan-1996	09-Apr-2020	2,105,004	FITB_O	02-Jan-1996	09-Apr-2020	2,375,099
DOW_N	02-Jan-1996	09-Apr-2020	2,076,036	FLIR_O	02-Jan-1996	09-Apr-2020	2,069,682
DO_N	02-Jan-1996	09-Apr-2020	2,147,971	FLR_N	02-Jan-1996	09-Apr-2020	2,088,430
DPS_N	17-Jul-1998	09-Jul-2018	998,388	FLS_N	23-Jul-1997	09-Apr-2020	1,880,938
DRI_N	02-Jan-1996	09-Apr-2020	2,059,474	FMC_N	02-Jan-1996	09-Apr-2020	2,057,789
DTE_N	02-Jan-1996	09-Apr-2020	2,096,381	FOSL_O	02-Jan-1996	09-Apr-2020	2,084,197
DTV_O	03-Dec-2007	24-Jul-2015	920,482	FOXA_O	01-Jul-2013	09-Apr-2020	795,548
DUK_N	02-Jan-1996	09-Apr-2020	2,179,894	FPL_N	02-Jan-1996	09-Apr-2020	1,526,980
DVA_N	09-Oct-2000	09-Apr-2020	1,815,901	FRX_N	08-Oct-1999	30-Jun-2014	1,390,588
DVN_N	13-Oct-2004	09-Apr-2020	1,519,726	FSLR_O	17-Nov-2006	09-Apr-2020	1,773,622
DV_N	02-Jan-1996	23-May-2017	1,645,741	FTI_N	14-Jun-2001	09-Apr-2020	1,766,695
DYN_N	07-Jul-1998	09-Apr-2018	1,716,591	FTR_N	02-Jan-1996	09-Apr-2020	574,603
D_N	02-Jan-1996	09-Apr-2020	2,135,979	F_N	03-Jan-1996	09-Apr-2020	2,231,903
EA_O	20-Dec-2011	09-Apr-2020	1,040,832	GAS_N	02-Jan-1996	30-Jun-2016	1,642,536
EBAY_O	02-Jan-1996	09-Apr-2020	2,783,791	GCI_N	02-Jan-1996	09-Apr-2020	2,180,581
ECL_N	02-Jan-1996	09-Apr-2020	2,062,930	GD_N	02-Jan-1996	09-Apr-2020	2,130,823
ED_N	02-Jan-1996	09-Apr-2020	2,140,587	GENZ_O	02-Jan-1996	08-Apr-2011	1,376,583
EFX_N	02-Jan-1996	09-Apr-2020	2,067,271	GE_N	02-Jan-1996	09-Apr-2020	2,325,726
EIX_N	05-Feb-1996	09-Apr-2020	2,109,407	GGP_N	02-Jan-1996	28-Aug-2018	1,627,547
EL_N	02-Jan-1996	09-Apr-2020	2,031,938	GILD_O	02-Jan-1996	09-Apr-2020	2,638,467
EMC_N	02-Jan-1996	06-Sep-2016	1,903,910	GIS_N	02-Jan-1996	09-Apr-2020	2,166,899
EMN_N	02-Jan-1996	09-Apr-2020	2,125,874	GLW_N	02-Jan-1996	09-Apr-2020	2,205,813
EMR_N	03-Jan-1996	09-Apr-2020	2,217,214	GMCR_O	02-Jan-1996	03-Mar-2016	1,419,045
EOG_N	02-Jan-1996	09-Apr-2020	2,048,174	GME_N	03-Jan-1996	09-Apr-2020	1,732,877
EP_N	31-Dec-2001	24-May-2012	1,012,625	GM_N	03-Jan-1996	09-Apr-2020	2,109,282
EQR_N	03-Jan-1996	09-Apr-2020	1,978,649	GNW_N	25-May-2004	09-Apr-2020	1,547,212
EQT_N	02-Jan-1996	09-Apr-2020	1,925,753	GOOGL_O	03-Apr-2014	09-Apr-2020	961,708
ERTS_O	02-Jan-1996	19-Dec-2011	1,522,287	GOOG_O	18-Aug-2004	09-Apr-2020	2,420,334
ESRX_O	02-Jan-1996	20-Dec-2018	2,211,830	GPC_N	02-Jan-1996	09-Apr-2020	2,082,753
ESS_N	03-Jan-1996	09-Apr-2020	1,794,719	GPS_N	02-Jan-1996	09-Apr-2020	2,213,385
ESV_N	02-Jan-1996	30-Jul-2019	2,101,464	GRMN_O	02-Jan-1996	09-Apr-2020	2,036,438
ES_N	15-Nov-2007	09-Apr-2020	1,028,230	GR_N	02-Jan-1996	26-Jul-2012	1,325,039
ETFC_O	27-Dec-2006	09-Apr-2020	1,566,814	GS_N	04-May-1999	09-Apr-2020	2,004,379
ETN_N	02-Jan-1996	09-Apr-2020	2,091,651	GT_N	02-Jan-1996	09-Apr-2020	1,662,446
ETR_N	02-Jan-1996	09-Apr-2020	2,093,237	GWW_N	02-Jan-1996	09-Apr-2020	2,060,499
EW_N	03-Apr-2000	09-Apr-2020	1,826,973	HAL_N	02-Jan-1996	09-Apr-2020	2,271,199
EXC_N	02-Jan-1996	09-Apr-2020	1,917,471	HANS_O	02-Jan-1996	06-Jan-2012	931,805
EXPD_O	03-Jan-1996	09-Apr-2020	2,193,606	HAR_N	02-Jan-1996	10-Mar-2017	1,620,588
EXPE_O	10-Nov-1999	09-Apr-2020	2,034,736	HAS_N	23-Jun-1999	09-Apr-2020	1,187,549
FAST_O	02-Jan-1996	09-Apr-2020	2,259,506	HBAN_O	02-Jan-1996	09-Apr-2020	2,234,148
FB_O	18-May-2012	09-Apr-2020	1,455,790	HCBK_O	02-Jan-1996	30-Oct-2015	1,491,786
FCX_N	02-Jan-1996	09-Apr-2020	2,092,799	HCN_N	02-Jan-1996	27-Feb-2018	1,660,846
FDO_N	02-Jan-1996	06-Jul-2015	1,573,720	HCP_N	02-Jan-1996	04-Nov-2019	1,878,390
FDX_N	02-Jan-1996	09-Apr-2020	2,178,218	HD_N	02-Jan-1996	09-Apr-2020	2,271,196
FE_N	10-Nov-1997	09-Apr-2020	2,040,899	HES_N	09-May-2006	09-Apr-2020	1,366,377
FFIV_O	04-Jun-1999	09-Apr-2020	2,289,549	HIG_N	02-Jan-1996	09-Apr-2020	2,171,192
FHN_N	21-Apr-2004	09-Apr-2020	1,564,114	HNZ_N	02-Jan-1996	07-Jun-2013	1,522,589
FII_N	14-May-1998	31-Jan-2020	1,861,654	HOG_N	15-Aug-2006	09-Apr-2020	1,338,192
FISV_O	02-Jan-1996	09-Apr-2020	2,363,549	HON_N	02-Jan-1996	09-Apr-2020	2,200,455

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
HOT_N	02-Jan-1996	23-Sep-2016	1,678,450	KSS_N	02-Jan-1996	09-Apr-2020	2,133,149
HPQ_N	06-May-2002	09-Apr-2020	1,755,199	KSU_N	02-Jan-1996	09-Apr-2020	2,025,538
HP_N	02-Jan-1996	09-Apr-2020	2,076,917	K_N	02-Jan-1996	09-Apr-2020	2,162,828
HRB_N	02-Jan-1996	09-Apr-2020	2,106,876	LB_N	02-Dec-2013	09-Apr-2020	622,066
HRL_N	02-Jan-1996	09-Apr-2020	1,952,920	LEG_N	02-Jan-1996	09-Apr-2020	2,021,440
HRS_N	02-Jan-1996	28-Jun-2019	1,987,621	LEN_N	02-Jan-1996	09-Apr-2020	2,002,687
HSP_N	25-May-2000	03-Sep-2015	1,357,020	LH_N	02-Jan-1996	09-Apr-2020	1,940,136
HST_N	18-Apr-2006	09-Apr-2020	1,371,924	LIFE_O	02-Jan-1996	09-Apr-2020	1,098,997
HSY_N	02-Jan-1996	09-Apr-2020	2,139,033	LLL_N	19-May-1998	28-Jun-2019	1,848,569
HUM_N	02-Jan-1996	09-Apr-2020	2,082,803	LLTC_O	02-Jan-1996	10-Mar-2017	2,161,769
IBM_N	02-Jan-1996	09-Apr-2020	2,298,859	LLY_N	02-Jan-1996	09-Apr-2020	2,310,786
ICE_N	16-Nov-2005	09-Apr-2020	1,404,895	LMT_N	02-Jan-1996	09-Apr-2020	2,202,793
IFF_N	02-Jan-1996	09-Apr-2020	2,091,009	LM_N	02-Jan-1996	09-Apr-2020	1,999,676
IGT_N	02-Jan-1996	09-Apr-2020	2,049,894	LNC_N	02-Jan-1996	09-Apr-2020	2,150,317
INTC_O	03-Jan-1996	09-Apr-2020	3,259,077	LOW_N	02-Jan-1996	09-Apr-2020	2,184,577
INTU_O	02-Jan-1996	09-Apr-2020	2,511,024	LO_N	10-Jun-2008	12-Jun-2015	687,618
IPG_N	02-Jan-1996	09-Apr-2020	2,142,489	LRCX_O	02-Jan-1996	09-Apr-2020	2,537,020
IP_N	02-Jan-1996	09-Apr-2020	2,266,209	LSI_N	02-Jan-1996	09-Apr-2020	1,895,872
IRM_N	26-Apr-1999	09-Apr-2020	1,831,510	LTD_N	02-Jan-1996	29-Nov-2013	1,525,757
IR_N	02-Jan-1996	09-Apr-2020	2,161,653	LUK_N	02-Jan-1996	23-May-2018	1,677,733
ISRG_O	13-Jun-2000	09-Apr-2020	1,960,366	LUV_N	02-Jan-1996	09-Apr-2020	2,209,479
ITT_N	02-Jan-1996	09-Apr-2020	1,964,119	LVLT_N	20-Oct-2011	01-Nov-2017	590,527
ITW_N	02-Jan-1996	09-Apr-2020	2,168,574	LXK_N	02-Jan-1996	29-Nov-2016	1,786,175
IVZ_N	24-May-2007	09-Apr-2020	1,261,434	LYB_N	14-Oct-2010	09-Apr-2020	928,822
JAVA_O	27-Aug-2007	26-Jan-2010	277,271	L_N	02-Jan-1996	09-Apr-2020	1,669,718
JBL_N	05-May-1998	09-Apr-2020	2,053,368	MAC_N	02-Jan-1996	09-Apr-2020	1,849,187
JCI_N	02-Jan-1996	09-Apr-2020	2,112,705	MAR_N	02-Jan-1996	09-Apr-2020	1,655,981
JCP_N	02-Jan-1996	09-Apr-2020	2,185,135	MAS_N	02-Jan-1996	09-Apr-2020	2,147,133
JDSU_O	06-Jul-1999	03-Aug-2015	1,898,220	MAT_O	28-Sep-2009	09-Apr-2020	1,290,404
JEC_N	02-Jan-1996	09-Dec-2019	1,916,851	MA_N	02-Jan-1996	09-Apr-2020	1,378,698
JNJ_N	02-Jan-1996	09-Apr-2020	2,296,920	MBI_N	02-Jan-1996	09-Apr-2020	2,120,931
JNPR_N	29-Oct-2009	09-Apr-2020	1,024,662	MCD_N	02-Jan-1996	09-Apr-2020	2,250,045
JNS_N	02-Jan-1996	26-May-2017	1,424,064	MCHP_O	02-Jan-1996	09-Apr-2020	2,439,960
JOYG_O	01-Aug-2001	05-Dec-2011	1,056,235	MCK_N	02-Jan-1996	09-Apr-2020	2,110,387
JOY_N	06-Dec-2011	05-Apr-2017	521,891	MCO_N	23-May-1996	09-Apr-2020	1,870,824
JPM_N	02-Jan-1996	09-Apr-2020	2,267,853	MDLZ_O	02-Oct-2012	09-Apr-2020	974,535
JWN_N	10-Jun-1999	09-Apr-2020	1,950,631	MDP_N	02-Jan-1996	09-Apr-2020	1,945,280
KBH_N	02-Jan-1996	09-Apr-2020	2,048,350	MDT_N	02-Jan-1996	09-Apr-2020	2,238,390
KEY_N	02-Jan-1996	09-Apr-2020	2,167,379	MEE_N	02-Jan-1996	01-Jun-2011	962,804
KFT_N	13-Jun-2001	25-Jun-2012	1,063,160	MET_N	23-Sep-1997	09-Apr-2020	1,913,739
KG_N	23-May-2000	28-Feb-2011	1,013,006	MFE_N	01-Jul-2004	28-Feb-2011	651,845
KIM_N	02-Jan-1996	09-Apr-2020	1,886,252	MHFI_N	14-May-2013	27-Apr-2016	289,262
KLAC_O	02-Jan-1996	09-Apr-2020	2,633,101	MHK_N	16-Dec-1997	09-Apr-2020	1,908,025
KMB_N	02-Jan-1996	09-Apr-2020	2,228,730	MHP_N	03-Jan-1996	13-May-2013	1,442,341
KMI_N	08-Oct-1999	09-Apr-2020	1,545,872	MHS_N	20-Aug-2003	30-Mar-2012	840,997
KMX_N	04-Feb-1997	09-Apr-2020	1,851,269	MIL_N	02-Jan-1996	12-Feb-2016	1,356,647
KORS_N	15-Dec-2011	31-Dec-2018	687,523	MI_N	02-Jan-1996	05-Jul-2011	1,107,732
KO_N	02-Jan-1996	09-Apr-2020	2,290,887	MJN_N	17-Jan-2002	15-Jun-2017	903,622
KRFT_O	02-Oct-2012	02-Jul-2015	322,867	MKC_N	25-May-1999	09-Apr-2020	1,879,911
KR_N	03-Jan-1996	09-Apr-2020	2,151,941	MLM_N	02-Jan-1996	09-Apr-2020	1,946,616

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
MMC_N	02-Jan-1996	09-Apr-2020	2,180,988	ODP_N	02-Jan-1996	09-Apr-2020	1,760,912
MMI_N	02-Jan-1996	09-Apr-2020	726,650	OI_N	03-Jan-1996	09-Apr-2020	2,002,438
MMM_N	02-Jan-1996	09-Apr-2020	2,236,930	OKE_N	02-Jan-1996	09-Apr-2020	1,983,066
MNK_N	08-Feb-1999	09-Apr-2020	664,365	OMC_N	02-Jan-1996	09-Apr-2020	2,125,336
MNST_O	01-May-2003	09-Apr-2020	1,605,548	ORCL_N	15-Jul-2013	09-Apr-2020	660,786
MOLX_O	02-Jan-1996	09-Dec-2013	1,570,855	ORLY_O	02-Jan-1996	09-Apr-2020	2,148,506
MON_N	10-Dec-1997	07-Jun-2018	1,657,348	OXY_N	02-Jan-1996	09-Apr-2020	2,174,665
MOS_N	25-Oct-2004	09-Apr-2020	1,508,979	PAYX_O	02-Jan-1996	09-Apr-2020	2,449,028
MO_N	02-Jan-1996	09-Apr-2020	2,323,517	PBCT_O	02-Jan-1996	09-Apr-2020	1,975,144
MPC_N	01-Jul-2011	09-Apr-2020	859,484	PBG_N	31-Mar-1999	26-Feb-2010	929,642
MRK_N	02-Jan-1996	09-Apr-2020	2,292,172	PBI_N	02-Jan-1996	09-Apr-2020	2,154,627
MRO_N	02-Jan-1996	09-Apr-2020	2,160,427	PCAR_O	03-Jan-1996	09-Apr-2020	2,356,138
MSFT_O	02-Jan-1996	09-Apr-2020	3,309,159	PCG_N	02-Jan-1996	09-Apr-2020	2,128,396
MSI_N	04-Jan-2011	09-Apr-2020	906,893	PCLN_O	30-Mar-1999	26-Feb-2018	2,112,002
MS_N	02-Jan-1996	09-Apr-2020	1,447,941	PCL_N	02-Jan-1996	19-Feb-2016	1,543,526
MTB_N	19-Sep-1996	09-Apr-2020	1,886,130	PCP_N	02-Jan-1996	29-Jan-2016	1,546,273
MUR_N	02-Jan-1996	09-Apr-2020	2,001,979	PCS_N	02-Jan-1996	30-Apr-2013	1,084,949
MU_N	02-Jan-1996	09-Apr-2020	1,487,777	PDCO_O	02-Jan-1996	09-Apr-2020	2,137,652
MWV_N	30-Jan-2002	01-Jul-2015	1,301,478	PEG_N	02-Jan-1996	09-Apr-2020	2,123,829
MWW_N	10-Nov-2008	01-Nov-2016	781,287	PEP_N	02-Jan-1996	09-Apr-2020	2,228,345
MYL_O	29-Dec-2008	09-Apr-2020	1,436,118	PETM_O	02-Jan-1996	11-Mar-2015	1,632,708
M_N	01-Jun-2007	09-Apr-2020	1,262,010	PFE_N	02-Jan-1996	09-Apr-2020	2,328,477
NAVI_O	22-Oct-1999	09-Apr-2020	1,193,194	PFG_N	02-Jan-1996	09-Apr-2020	1,789,089
NBL_N	02-Jan-1996	09-Apr-2020	2,044,903	PGR_N	02-Jan-1996	09-Apr-2020	2,121,030
NBR_N	03-Nov-2005	09-Apr-2020	1,414,483	PG_N	02-Jan-1996	09-Apr-2020	2,314,239
NDAQ_O	10-Feb-2005	09-Apr-2020	1,704,069	PHM_N	02-Jan-1996	09-Apr-2020	2,042,253
NEE_N	23-Jun-2010	09-Apr-2020	961,305	PH_N	02-Jan-1996	09-Apr-2020	2,107,013
NEM_N	03-Jan-1996	09-Apr-2020	2,211,055	PKI_N	26-Oct-1999	09-Apr-2020	1,933,481
NE_N	29-Mar-1996	09-Apr-2020	2,165,382	PLD_N	02-Jul-1998	09-Apr-2020	1,875,169
NFLX_O	23-May-2002	09-Apr-2020	2,392,224	PLL_N	02-Jan-1996	11-Nov-2019	1,605,514
NFX_N	02-Jan-1996	13-Feb-2019	1,832,255	PM_N	31-Mar-2008	09-Apr-2020	1,181,089
NI_N	02-Jan-1996	09-Apr-2020	2,020,446	PNC_N	02-Jan-1996	09-Apr-2020	2,197,949
NKE_N	02-Jan-1996	09-Apr-2020	2,224,473	PNR_N	04-Mar-1996	09-Apr-2020	1,949,350
NLSN_N	26-Jan-2011	09-Apr-2020	897,829	PNW_N	02-Jan-1996	09-Apr-2020	2,019,549
NOC_N	02-Jan-1996	09-Apr-2020	2,094,598	POM_N	02-Jan-1996	24-Mar-2016	1,602,116
NOVL_O	02-Jan-1996	27-Apr-2011	1,349,640	PPG_N	02-Jan-1996	09-Apr-2020	2,167,029
NOV_N	02-Jan-1996	09-Apr-2020	1,573,503	PPL_N	02-Jan-1996	09-Apr-2020	2,067,127
NRG_N	31-May-2000	09-Apr-2020	1,672,350	PRGO_N	06-Jun-2013	09-Apr-2020	668,722
NSC_N	02-Jan-1996	09-Apr-2020	2,175,649	PRU_N	13-Dec-2001	09-Apr-2020	1,782,469
NSM_N	02-Jan-1996	31-Jul-2018	1,996,944	PSA_N	02-Jan-1996	09-Apr-2020	1,888,269
NTAP_O	02-Jan-1996	09-Apr-2020	2,487,241	PSX_N	02-Jan-1996	09-Apr-2020	803,084
NTRS_O	02-Jan-1996	09-Apr-2020	2,332,169	PTV_N	05-Nov-1999	16-Nov-2010	976,124
NUE_N	02-Jan-1996	09-Apr-2020	2,135,579	PVH_N	02-Jan-1996	09-Apr-2020	1,807,982
NU_N	02-Jan-1996	18-Feb-2015	1,448,773	PWR_N	12-Feb-1998	09-Apr-2020	1,868,736
NVDA_O	22-Jan-1999	09-Apr-2020	2,684,496	PXD_N	08-Aug-1997	09-Apr-2020	1,950,527
NVLS_O	02-Jan-1996	24-Jul-2017	1,809,544	PX_N	02-Jan-1996	30-Oct-2018	1,990,693
NWL_N	02-Jan-1996	09-Apr-2020	2,131,328	QCOM_O	02-Jan-1996	09-Apr-2020	2,849,469
NWSA_O	29-Dec-2008	09-Apr-2020	1,263,345	QEP_N	01-Jul-2010	09-Apr-2020	957,557
NYT_N	25-Sep-1997	09-Apr-2020	2,041,906	QLGC_O	03-Jan-1996	16-Aug-2016	2,001,183
NYX_N	08-Mar-2006	12-Nov-2013	754,846	Q_N	03-Jan-2000	14-Nov-2017	1,515,869

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
RAI_N	02-Aug-2004	25-Jul-2017	1,272,362	SYMC_O	02-Jan-1996	04-Nov-2019	2,399,193
RCL_N	02-Jan-1996	09-Apr-2020	1,976,250	SYY_N	02-Jan-1996	09-Apr-2020	2,148,710
RDC_N	02-Jan-1996	11-Apr-2019	2,076,304	S_N	02-Jan-1996	01-Apr-2020	2,204,212
REGN_O	02-Jan-1996	09-Apr-2020	2,087,151	TAP_N	23-Apr-1996	09-Apr-2020	1,593,289
RF_N	03-May-2002	09-Apr-2020	1,749,826	TDC_N	01-Oct-2007	09-Apr-2020	1,225,942
RHI_N	02-Jan-1996	09-Apr-2020	2,035,511	TEG_N	24-Feb-1997	29-Jun-2015	831,809
RHT_N	18-Nov-2002	09-Jul-2019	1,312,607	TEL_N	02-Jan-1996	09-Apr-2020	1,260,966
RIG_N	02-Jan-1996	09-Apr-2020	2,180,099	TER_N	02-Jan-1996	09-Apr-2020	2,193,190
RL_N	02-Jan-1996	09-Apr-2020	1,902,044	TE_N	02-Jan-1996	30-Jun-2016	1,673,046
ROK_N	02-Jan-1996	09-Apr-2020	2,150,278	TGT_N	31-Jan-2000	09-Apr-2020	1,956,846
ROP_N	01-Nov-1996	09-Apr-2020	1,862,049	THC_N	02-Jan-1996	09-Apr-2020	2,141,146
ROST_O	02-Jan-1996	09-Apr-2020	2,319,058	TIE_N	16-Jul-1998	07-Jan-2013	971,804
RRC_N	26-Aug-1998	09-Apr-2020	1,748,909	TIF_N	02-Jan-1996	09-Apr-2020	2,070,864
RRD_O	05-Aug-2009	19-Aug-2016	787,510	TJX_N	02-Jan-1996	09-Apr-2020	2,135,203
RSG_N	02-Jul-1998	09-Apr-2020	1,932,858	TLAB_O	02-Jan-1996	03-Dec-2013	1,771,722
RTN_N	02-Jan-1996	03-Apr-2020	1,930,725	TMK_N	02-Jan-1996	08-Aug-2019	2,007,666
R_N	02-Jan-1996	09-Apr-2020	2,019,734	TMO_N	02-Jan-1996	09-Apr-2020	2,106,333
SAI_N	02-Jan-1996	27-Sep-2013	821,032	TRIP_O	02-Jan-1996	09-Apr-2020	1,017,153
SBUX_O	02-Jan-1996	09-Apr-2020	2,680,668	TROW_O	02-Jan-1996	09-Apr-2020	2,278,480
SCG_N	02-Jan-1996	31-Dec-2018	1,857,179	TRV_N	02-Jan-1996	09-Apr-2020	1,492,936
SCHW_N	05-Mar-2010	09-Apr-2020	991,593	TSCO_O	02-Jan-1996	09-Apr-2020	2,079,629
SEE_N	02-Jan-1996	09-Apr-2020	2,050,409	TSN_N	17-Oct-1997	09-Apr-2020	1,950,834
SE_N	08-Mar-1996	09-Apr-2020	1,722,290	TSO_N	02-Jan-1996	31-Jul-2017	1,637,192
SHLD_O	28-Mar-2005	23-Oct-2018	1,552,837	TSS_N	03-Jan-1996	17-Sep-2019	1,772,421
SHW_N	03-Jan-1996	09-Apr-2020	2,099,741	TWC_N	01-Mar-2007	17-May-2016	900,725
SIAL_O	03-Jan-1996	18-Nov-2015	1,744,630	TWX_N	03-Jan-1996	14-Jun-2018	1,821,187
SII_N	02-Jan-1996	27-Aug-2010	1,205,460	TXN_N	02-Jan-1996	09-Apr-2020	1,668,418
SJM_N	29-Aug-2000	09-Apr-2020	1,773,704	TXT_N	02-Jan-1996	09-Apr-2020	2,151,982
SLB_N	02-Jan-1996	09-Apr-2020	2,271,445	TYC_N	02-Jan-1996	02-Sep-2016	1,860,870
SLE_N	02-Jan-1996	28-Jun-2012	1,439,969	T_N	03-Jan-1996	09-Apr-2020	2,255,745
SNA_N	02-Jan-1996	09-Apr-2020	1,994,519	UA_N	02-Jan-1996	09-Apr-2020	1,305,754
SNDK_O	02-Jan-1996	12-May-2016	2,043,724	UHS_N	02-Jan-1996	09-Apr-2020	1,953,488
SNI_N	01-Jul-2008	14-Mar-2016	752,499	UNH_N	02-Jan-1996	09-Apr-2020	2,189,587
SO_N	02-Jan-1996	09-Apr-2020	2,179,579	UNM_N	02-Jan-1996	09-Apr-2020	2,133,439
SPG_N	02-Jan-1996	09-Apr-2020	1,982,000	UNP_N	02-Jan-1996	09-Apr-2020	2,158,248
SPLS_O	02-Jan-1996	12-Sep-2017	2,140,893	UPS_N	10-Nov-1999	09-Apr-2020	1,951,671
SRCL_O	23-Aug-1996	09-Apr-2020	2,074,532	URBN_O	03-Jan-1996	09-Apr-2020	2,109,211
SRE_N	29-Jun-1998	09-Apr-2020	2,012,253	URI_N	18-Dec-1997	09-Apr-2020	1,905,870
STI_N	02-Jan-1996	06-Dec-2019	2,140,876	USB_N	04-Aug-1997	09-Apr-2020	2,124,485
STJ_N	02-Dec-1996	04-Jan-2017	1,749,978	UTX_N	02-Jan-1996	02-Apr-2020	2,232,005
STR_N	02-Jan-1996	16-Sep-2016	1,614,366	VAR_N	02-Jan-1996	09-Apr-2020	1,980,031
STT_N	02-Jan-1996	09-Apr-2020	2,150,837	VFC_N	03-Jan-1996	09-Apr-2020	2,086,784
STX_O	16-Sep-2008	09-Apr-2020	1,412,496	VIAB_O	01-Dec-2011	04-Dec-2019	944,616
STZ_N	20-Sep-2000	09-Apr-2020	1,835,118	VLO_N	02-Jan-1996	09-Apr-2020	2,006,094
SUN_N	02-Jan-1996	09-Apr-2020	1,843,540	VMC_N	02-Jan-1996	09-Apr-2020	1,999,038
SVU_N	02-Jan-1996	22-Oct-2018	1,907,957	VNO_N	02-Jan-1996	09-Apr-2020	1,890,700
SWK_N	02-Jan-1996	09-Apr-2020	2,052,025	VRSN_O	30-Jan-1998	09-Apr-2020	2,447,514
SWN_N	02-Jan-1996	09-Apr-2020	1,798,725	VRTX_O	02-Jan-1996	09-Apr-2020	2,333,086
SWY_N	02-Jan-1996	29-Jan-2015	1,654,047	VTR_N	01-May-1998	09-Apr-2020	1,814,627
SYK_N	25-Jul-1997	09-Apr-2020	2,004,851	VZ_N	03-Jul-2000	09-Apr-2020	1,926,918

Table 5.1: Stock Ticker List

Ticker	Start Date	EndDate	Number of Quotes	Ticker	Start Date	EndDate	Number of Quotes
V_N	02-Jan-1996	09-Apr-2020	1,687,854	WYNN_O	25-Oct-2002	09-Apr-2020	2,048,467
WAG_N	02-Jan-1996	30-Dec-2014	1,687,283	WYN_N	21-May-1996	31-May-2018	1,243,807
WAT_N	02-Jan-1996	09-Apr-2020	2,012,484	WY_N	02-Jan-1996	09-Apr-2020	2,190,983
WBA_O	31-Dec-2014	09-Apr-2020	699,263	XEC_N	01-Oct-2002	09-Apr-2020	1,671,344
WDC_N	02-Jan-1996	09-Apr-2020	1,509,382	XEL_N	02-Jan-1996	09-Apr-2020	1,892,938
WEC_N	02-Jan-1996	09-Apr-2020	2,008,516	XLNX_O	02-Jan-1996	09-Apr-2020	2,671,151
WFC_N	02-Jan-1996	09-Apr-2020	2,257,351	XL_N	02-Jan-1996	12-Sep-2018	1,837,282
WFMI_O	02-Jan-1996	05-May-2011	1,235,929	XOM_N	01-Dec-1999	09-Apr-2020	1,980,937
WFM_O	06-May-2011	28-Aug-2017	789,154	XRAY_O	02-Jan-1996	09-Apr-2020	2,166,171
WFR_N	02-Jan-1996	31-May-2013	1,205,154	XRX_N	02-Jan-1996	09-Apr-2020	2,265,252
WHR_N	02-Jan-1996	09-Apr-2020	2,127,939	XTO_N	02-Jan-1996	25-Jun-2010	997,655
WIN_N	02-Jan-1996	05-Mar-2019	983,338	XYL_N	01-Nov-2011	09-Apr-2020	823,287
WLP_N	02-Jan-1996	02-Dec-2014	1,540,070	X_N	02-Jan-1996	09-Apr-2020	2,129,784
WMB_N	02-Jan-1996	09-Apr-2020	2,183,170	YHOO_O	12-Apr-1996	16-Jun-2017	2,558,951
WMT_N	02-Jan-1996	09-Apr-2020	2,257,067	YUM_N	07-Oct-1997	09-Apr-2020	2,060,980
WM_N	09-Dec-1998	09-Apr-2020	1,966,129	ZION_O	02-Jan-1996	09-Apr-2020	2,232,899
WPI_N	18-Sep-1997	23-Jan-2013	1,332,685	ZMH_N	07-Aug-2001	26-Jun-2015	1,338,825
WPX_N	03-Jan-2012	09-Apr-2020	810,182	ZTS_N	01-Feb-2013	09-Apr-2020	702,710
WU_N	02-Oct-2006	09-Apr-2020	1,325,573				

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
AUDCAD	1D	15-Jun-2017	08-May-2020	1,524,960	AUDSEK	6M	02-Jan-2015	08-May-2020	2,813,760
AUDCAD	1M	02-Jan-2015	08-May-2020	2,813,760	AUDSEK	9M	15-Jun-2017	08-May-2020	1,524,960
AUDCAD	2M	02-Jan-2015	08-May-2020	2,813,760	AUDSEK	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDCAD	3M	02-Jan-2015	08-May-2020	2,813,760	AUDSEK	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDCAD	6M	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	1D	15-Jun-2017	08-May-2020	1,524,960
AUDCAD	9M	15-Jun-2017	08-May-2020	1,524,960	AUDSGD	1M	02-Jan-2015	08-May-2020	2,813,760
AUDCAD	1Y	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	2M	02-Jan-2015	08-May-2020	2,813,760
AUDCAD	2Y	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	3M	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	1D	15-Jun-2017	08-May-2020	1,524,960	AUDSGD	6M	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	1M	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	9M	15-Jun-2017	08-May-2020	1,524,960
AUDCHF	2M	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	3M	02-Jan-2015	08-May-2020	2,813,760	AUDSGD	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	6M	02-Jan-2015	08-May-2020	2,813,760	AUDUSD	1D	15-Jun-2017	08-May-2020	1,524,960
AUDCHF	9M	15-Jun-2017	08-May-2020	1,524,960	AUDUSD	1M	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	1Y	02-Jan-2015	08-May-2020	2,813,760	AUDUSD	2M	02-Jan-2015	08-May-2020	2,813,760
AUDCHF	2Y	02-Jan-2015	08-May-2020	2,813,760	AUDUSD	3M	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	1D	15-Jun-2017	08-May-2020	1,524,960	AUDUSD	6M	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	1M	02-Jan-2015	08-May-2020	2,813,760	AUDUSD	9M	15-Jun-2017	08-May-2020	1,524,960
AUDDKK	2M	02-Jan-2015	08-May-2020	2,813,760	AUDUSD	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	3M	05-Jan-2015	08-May-2020	2,809,440	AUDUSD	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	6M	02-Jan-2015	08-May-2020	2,813,760	CADCHF	1D	15-Jun-2017	08-May-2020	1,524,960
AUDDKK	9M	15-Jun-2017	08-May-2020	1,524,960	CADCHF	1M	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	1Y	02-Jan-2015	08-May-2020	2,813,760	CADCHF	2M	02-Jan-2015	08-May-2020	2,813,760
AUDDKK	2Y	02-Jan-2015	08-May-2020	2,813,760	CADCHF	3M	02-Jan-2015	08-May-2020	2,813,760
AUDHKD	1D	15-Jun-2017	08-May-2020	1,524,960	CADCHF	6M	02-Jan-2015	08-May-2020	2,813,760
AUDHKD	1M	02-Jan-2015	08-May-2020	2,813,760	CADCHF	9M	15-Jun-2017	08-May-2020	1,524,960
AUDHKD	2M	02-Jan-2015	08-May-2020	2,813,760	CADCHF	1Y	05-Jan-2015	08-May-2020	2,809,440
AUDHKD	3M	02-Jan-2015	08-May-2020	2,813,760	CADCHF	2Y	05-Jan-2015	08-May-2020	2,809,440
AUDHKD	6M	02-Jan-2015	08-May-2020	2,813,760	CADDKK	1D	15-Jun-2017	08-May-2020	1,524,960
AUDHKD	9M	15-Jun-2017	08-May-2020	1,524,960	CADDKK	1M	02-Jan-2015	08-May-2020	2,813,760
AUDHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	CADDKK	2M	02-Jan-2015	08-May-2020	2,813,760
AUDHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	CADDKK	3M	05-Jan-2015	08-May-2020	2,809,440
AUDJPY	1D	15-Jun-2017	08-May-2020	1,524,960	CADDKK	6M	02-Jan-2015	08-May-2020	2,813,760
AUDJPY	1M	02-Jan-2015	08-May-2020	2,813,760	CADDKK	9M	15-Jun-2017	08-May-2020	1,524,960
AUDJPY	2M	02-Jan-2015	08-May-2020	2,813,760	CADDKK	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDJPY	3M	02-Jan-2015	08-May-2020	2,813,760	CADDKK	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDJPY	6M	02-Jan-2015	08-May-2020	2,813,760	CADJPY	1D	15-Jun-2017	08-May-2020	1,524,960
AUDJPY	9M	15-Jun-2017	08-May-2020	1,524,960	CADJPY	1M	02-Jan-2015	08-May-2020	2,813,760
AUDJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	CADJPY	2M	02-Jan-2015	08-May-2020	2,813,760
AUDJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	CADJPY	3M	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	1D	15-Jun-2017	08-May-2020	1,524,960	CADJPY	6M	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	1M	02-Jan-2015	08-May-2020	2,813,760	CADJPY	9M	15-Jun-2017	08-May-2020	1,524,960
AUDNOK	2M	02-Jan-2015	08-May-2020	2,813,760	CADJPY	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	3M	02-Jan-2015	08-May-2020	2,813,760	CADJPY	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	6M	02-Jan-2015	08-May-2020	2,813,760	CADNOK	1D	15-Jun-2017	08-May-2020	1,524,960
AUDNOK	9M	15-Jun-2017	08-May-2020	1,524,960	CADNOK	1M	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	1Y	02-Jan-2015	08-May-2020	2,813,760	CADNOK	2M	02-Jan-2015	08-May-2020	2,813,760
AUDNOK	2Y	02-Jan-2015	08-May-2020	2,813,760	CADNOK	3M	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	1D	15-Jun-2017	08-May-2020	1,524,960	CADNOK	6M	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	1M	05-Jan-2015	08-May-2020	2,809,440	CADNOK	9M	15-Jun-2017	08-May-2020	1,524,960
AUDNZD	2M	06-Jan-2015	08-May-2020	2,808,000	CADNOK	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	3M	06-Jan-2015	08-May-2020	2,808,000	CADNOK	2Y	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	6M	02-Jan-2015	08-May-2020	2,813,760	CADSEK	1D	15-Jun-2017	08-May-2020	1,524,960
AUDNZD	9M	15-Jun-2017	06-May-2020	1,522,080	CADSEK	1M	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	1Y	02-Jan-2015	08-May-2020	2,813,760	CADSEK	2M	02-Jan-2015	08-May-2020	2,813,760
AUDNZD	2Y	02-Jan-2015	08-May-2020	2,813,760	CADSEK	3M	02-Jan-2015	08-May-2020	2,813,760
AUDSEK	1D	15-Jun-2017	08-May-2020	1,524,960	CADSEK	6M	02-Jan-2015	08-May-2020	2,813,760
AUDSEK	1M	02-Jan-2015	08-May-2020	2,813,760	CADSEK	9M	15-Jun-2017	08-May-2020	1,524,960
AUDSEK	2M	02-Jan-2015	08-May-2020	2,813,760	CADSEK	1Y	02-Jan-2015	08-May-2020	2,813,760
AUDSEK	3M	02-Jan-2015	08-May-2020	2,813,760	CADSEK	2Y	02-Jan-2015	08-May-2020	2,813,760

Notes: this table presents the ticker list used in the foreign exchange study. The 'Currency' column shows the underlying asset. The 'Maturity' column shows the contract maturity for corresponding currency. The 'Start Date' column shows the first recorded date for trades in database of my study. The 'End Date' column shows the last recorded date for trades in database of my study. The 'Number of Quotes' column shows the number of trades in my database for corresponding contract.

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
CADSGD	1D	15-Jun-2017	08-May-2020	1,524,960	CHFZAR	1D	15-Jun-2017	08-May-2020	1,524,960
CADSGD	1M	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	1M	02-Jan-2015	08-May-2020	2,813,760
CADSGD	2M	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	2M	02-Jan-2015	08-May-2020	2,813,760
CADSGD	3M	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	3M	02-Jan-2015	08-May-2020	2,813,760
CADSGD	6M	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	6M	02-Jan-2015	08-May-2020	2,813,760
CADSGD	9M	15-Jun-2017	08-May-2020	1,524,960	CHFZAR	9M	15-Jun-2017	08-May-2020	1,524,960
CADSGD	1Y	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	1Y	02-Jan-2015	08-May-2020	2,813,760
CADSGD	2Y	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	2Y	15-Jun-2017	08-May-2020	1,524,960
CHFDKK	1D	15-Jun-2017	08-May-2020	1,524,960	CNHSGD	3M	08-Jun-2018	10-Jan-2019	312,480
CHFDKK	1M	02-Jan-2015	08-May-2020	2,813,760	CNHSGD	6M	08-Jun-2018	10-Jan-2019	312,480
CHFDKK	2M	02-Jan-2015	08-May-2020	2,813,760	CNHSGD	9M	08-Jun-2018	10-Jan-2019	312,480
CHFDKK	3M	02-Jan-2015	08-May-2020	2,813,760	CNHSGD	1Y	08-Jun-2018	10-Jan-2019	312,480
CHFDKK	6M	02-Jan-2015	08-May-2020	2,813,760	CNHSGD	2Y	08-Jun-2018	10-Jan-2019	312,480
CHFDKK	9M	15-Jun-2017	08-May-2020	1,524,960	CHFZAR	9M	15-Jun-2017	08-May-2020	1,524,960
CHFDKK	1Y	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	1Y	02-Jan-2015	08-May-2020	2,813,760
CHFDKK	2Y	02-Jan-2015	08-May-2020	2,813,760	CHFZAR	2Y	15-Jun-2017	08-May-2020	1,524,960
CHFHKD	1D	15-Jun-2017	10-Jan-2019	828,000	DKKJPY	1D	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	1M	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	1M	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	2M	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	2M	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	3M	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	3M	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	6M	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	6M	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	9M	15-Jun-2017	28-Apr-2020	1,510,560	DKKJPY	9M	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	1Y	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	1Y	07-Jul-2017	08-May-2020	1,493,280
CHFHKD	2Y	02-Jan-2015	28-Apr-2020	2,799,360	DKKJPY	2Y	07-Jul-2017	08-May-2020	1,493,280
CHFHUF	1D	15-Jun-2017	08-May-2020	1,524,960	DKKNOK	1M	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	1M	02-Jan-2015	08-May-2020	2,813,760	DKKNOK	2M	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	2M	02-Jan-2015	08-May-2020	2,813,760	DKKNOK	3M	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	3M	02-Jan-2015	08-May-2020	2,813,760	DKKNOK	6M	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	6M	05-Jan-2015	08-May-2020	2,809,440	DKKNOK	9M	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	9M	15-Jun-2017	08-May-2020	1,524,960	DKKNOK	1Y	02-Jan-2015	21-Jul-2016	816,480
CHFHUF	1Y	02-Jan-2015	08-May-2020	2,813,760	DKKJPY	1Y	07-Jul-2017	08-May-2020	1,493,280
CHFHUF	2Y	02-Jan-2015	08-May-2020	2,813,760	DKKJPY	2Y	07-Jul-2017	08-May-2020	1,493,280
CHFJPY	1D	15-Jun-2017	08-May-2020	1,524,960	DKKSEK	1M	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	1M	02-Jan-2015	08-May-2020	2,813,760	DKKSEK	2M	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	2M	02-Jan-2015	08-May-2020	2,813,760	DKKSEK	3M	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	3M	02-Jan-2015	08-May-2020	2,813,760	DKKSEK	6M	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	6M	02-Jan-2015	08-May-2020	2,813,760	DKKSEK	9M	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	9M	15-Jun-2017	08-May-2020	1,524,960	DKKSEK	1Y	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	DKKSEK	2Y	02-Jan-2015	22-Jul-2016	817,920
CHFJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	DKKJPY	2Y	07-Jul-2017	08-May-2020	1,493,280
CHFMXN	1D	15-Jun-2017	08-May-2020	1,524,960	EURAUD	1D	15-Jun-2017	08-May-2020	1,524,960
CHFMXN	1M	02-Jan-2015	08-May-2020	2,813,760	EURAUD	1M	02-Jan-2015	08-May-2020	2,813,760
CHFMXN	2M	02-Jan-2015	08-May-2020	2,813,760	EURAUD	2M	02-Jan-2015	08-May-2020	2,813,760
CHFMXN	3M	02-Jan-2015	08-May-2020	2,813,760	EURAUD	3M	02-Jan-2015	08-May-2020	2,813,760
CHFMXN	6M	02-Jan-2015	08-May-2020	2,813,760	EURAUD	6M	02-Jan-2015	08-May-2020	2,813,760
CHFMXN	9M	15-Jun-2017	08-May-2020	1,524,960	EURAUD	9M	15-Jun-2017	08-May-2020	1,524,960
CHFMXN	1Y	02-Jan-2015	08-May-2020	2,813,760	EURAUD	1Y	02-Jan-2015	08-May-2020	2,813,760
CHFMXN	2Y	02-Jan-2015	08-May-2020	2,813,760	EURAUD	2Y	02-Jan-2015	08-May-2020	2,813,760
CHFNOK	1D	15-Jun-2017	08-May-2020	1,524,960	EURBRL	1D	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	1M	02-Jan-2015	08-May-2020	2,813,760	EURBRL	1M	02-Jan-2015	08-May-2020	2,813,760
CHFNOK	2M	02-Jan-2015	08-May-2020	2,813,760	EURBRL	2M	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	3M	02-Jan-2015	08-May-2020	2,813,760	EURBRL	3M	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	6M	02-Jan-2015	08-May-2020	2,813,760	EURBRL	6M	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	9M	15-Jun-2017	08-May-2020	1,524,960	EURBRL	9M	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	1Y	02-Jan-2015	08-May-2020	2,813,760	EURBRL	1Y	15-Jun-2017	08-May-2020	1,524,960
CHFNOK	2Y	02-Jan-2015	08-May-2020	2,813,760	EURBRL	2Y	15-Jun-2017	08-May-2020	1,524,960
CHFPLN	1D	15-Jun-2017	08-May-2020	1,524,960	EURCAD	1D	15-Jun-2017	08-May-2020	1,524,960
CHFPLN	1M	02-Jan-2015	08-May-2020	2,813,760	EURCAD	1M	02-Jan-2015	08-May-2020	2,813,760
CHFPLN	2M	02-Jan-2015	08-May-2020	2,813,760	EURCAD	2M	02-Jan-2015	08-May-2020	2,813,760
CHFPLN	3M	02-Jan-2015	08-May-2020	2,813,760	EURCAD	3M	02-Jan-2015	08-May-2020	2,813,760

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
CHFPLN	6M	02-Jan-2015	08-May-2020	2,813,760	EURCAD	6M	02-Jan-2015	08-May-2020	2,813,760
CHFPLN	9M	15-Jun-2017	08-May-2020	1,524,960	EURCAD	9M	15-Jun-2017	08-May-2020	1,524,960
CHFPLN	1Y	02-Jan-2015	08-May-2020	2,813,760	EURCAD	1Y	02-Jan-2015	08-May-2020	2,813,760
CHFPLN	2Y	02-Jan-2015	08-May-2020	2,813,760	EURCAD	2Y	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	1D	15-Jun-2017	08-May-2020	1,524,960	EURCHF	1D	15-Jun-2017	08-May-2020	1,524,960
CHFSEK	1M	02-Jan-2015	08-May-2020	2,813,760	EURCHF	1M	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	2M	02-Jan-2015	08-May-2020	2,813,760	EURCHF	2M	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	3M	02-Jan-2015	08-May-2020	2,813,760	EURCHF	3M	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	6M	02-Jan-2015	08-May-2020	2,813,760	EURCHF	6M	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	9M	15-Jun-2017	08-May-2020	1,524,960	EURCHF	9M	15-Jun-2017	08-May-2020	1,524,960
CHFSEK	1Y	02-Jan-2015	08-May-2020	2,813,760	EURCHF	1Y	02-Jan-2015	08-May-2020	2,813,760
CHFSEK	2Y	02-Jan-2015	08-May-2020	2,813,760	EURCHF	2Y	02-Jan-2015	08-May-2020	2,813,760
CHFSGD	1D	15-Jun-2017	08-May-2020	1,524,960	EURCNH	1D	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	1M	02-Jan-2015	08-May-2020	2,813,760	EURCNH	1M	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	2M	02-Jan-2015	08-May-2020	2,813,760	EURCNH	2M	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	3M	02-Jan-2015	08-May-2020	2,813,760	EURCNH	3M	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	6M	02-Jan-2015	08-May-2020	2,813,760	EURCNH	6M	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	9M	15-Jun-2017	08-May-2020	1,524,960	EURCNH	9M	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	1Y	05-Jan-2015	08-May-2020	2,809,440	EURCNH	1Y	08-Jun-2018	08-May-2020	1,009,440
CHFSGD	2Y	02-Jan-2015	08-May-2020	2,813,760	EURCNH	2Y	08-Jun-2018	08-May-2020	1,009,440
EURCNY	1D	15-Jun-2017	08-May-2020	1,524,960	EURMXN	1D	15-Jun-2017	08-May-2020	1,524,960
EURCNY	1M	02-Jan-2015	08-May-2020	2,813,760	EURMXN	1M	02-Jan-2015	08-May-2020	2,813,760
EURCNY	2M	02-Jan-2015	08-May-2020	2,813,760	EURMXN	2M	02-Jan-2015	08-May-2020	2,813,760
EURCNY	3M	02-Jan-2015	08-May-2020	2,813,760	EURMXN	3M	02-Jan-2015	08-May-2020	2,813,760
EURCNY	6M	02-Jan-2015	08-May-2020	2,813,760	EURMXN	6M	02-Jan-2015	08-May-2020	2,813,760
EURCNY	9M	15-Jun-2017	08-May-2020	1,524,960	EURMXN	9M	15-Jun-2017	08-May-2020	1,524,960
EURCNY	1Y	02-Jan-2015	07-May-2020	2,812,320	EURMXN	1Y	02-Jan-2015	08-May-2020	2,813,760
EURCNY	2Y	02-Jan-2015	08-May-2020	2,813,760	EURMXN	2Y	02-Jan-2015	08-May-2020	2,813,760
EURCZK	1D	15-Jun-2017	08-May-2020	1,524,960	EURNOK	1D	15-Jun-2017	08-May-2020	1,524,960
EURCZK	1M	02-Jan-2015	08-May-2020	2,813,760	EURNOK	1M	02-Jan-2015	08-May-2020	2,813,760
EURCZK	2M	02-Jan-2015	08-May-2020	2,813,760	EURNOK	2M	02-Jan-2015	08-May-2020	2,813,760
EURCZK	3M	02-Jan-2015	08-May-2020	2,813,760	EURNOK	3M	02-Jan-2015	08-May-2020	2,813,760
EURCZK	6M	02-Jan-2015	08-May-2020	2,813,760	EURNOK	6M	02-Jan-2015	08-May-2020	2,813,760
EURCZK	9M	15-Jun-2017	08-May-2020	1,524,960	EURNOK	9M	15-Jun-2017	08-May-2020	1,524,960
EURCZK	1Y	15-Jun-2017	08-May-2020	1,524,960	EURNOK	1Y	02-Jan-2015	08-May-2020	2,813,760
EURCZK	2Y	02-Jan-2015	08-May-2020	2,813,760	EURNOK	2Y	02-Jan-2015	08-May-2020	2,813,760
EURDKK	1D	07-Jul-2017	08-May-2020	1,493,280	EURNZD	1D	15-Jun-2017	08-May-2020	1,524,960
EURDKK	1M	07-Jul-2017	08-May-2020	1,493,280	EURNZD	1M	02-Jan-2015	08-May-2020	2,813,760
EURDKK	2M	07-Jul-2017	08-May-2020	1,493,280	EURNZD	2M	02-Jan-2015	08-May-2020	2,813,760
EURDKK	3M	07-Jul-2017	08-May-2020	1,493,280	EURNZD	3M	02-Jan-2015	08-May-2020	2,813,760
EURDKK	6M	07-Jul-2017	08-May-2020	1,493,280	EURNZD	6M	02-Jan-2015	08-May-2020	2,813,760
EURDKK	9M	07-Jul-2017	08-May-2020	1,493,280	EURNZD	9M	15-Jun-2017	08-May-2020	1,524,960
EURDKK	1Y	07-Jul-2017	08-May-2020	1,493,280	EURNZD	1Y	02-Jan-2015	08-May-2020	2,813,760
EURDKK	2Y	07-Jul-2017	08-May-2020	1,493,280	EURNZD	2Y	02-Jan-2015	08-May-2020	2,813,760
EURGBP	1D	15-Jun-2017	08-May-2020	1,524,960	EURPLN	1D	15-Jun-2017	08-May-2020	1,524,960
EURGBP	1M	02-Jan-2015	08-May-2020	2,813,760	EURPLN	1M	02-Jan-2015	08-May-2020	2,813,760
EURGBP	2M	02-Jan-2015	08-May-2020	2,813,760	EURPLN	2M	02-Jan-2015	08-May-2020	2,813,760
EURGBP	3M	02-Jan-2015	08-May-2020	2,813,760	EURPLN	3M	02-Jan-2015	08-May-2020	2,813,760
EURGBP	6M	02-Jan-2015	08-May-2020	2,813,760	EURPLN	6M	02-Jan-2015	08-May-2020	2,813,760
EURGBP	9M	15-Jun-2017	08-May-2020	1,524,960	EURPLN	9M	15-Jun-2017	08-May-2020	1,524,960
EURGBP	1Y	02-Jan-2015	08-May-2020	2,813,760	EURPLN	1Y	02-Jan-2015	08-May-2020	2,813,760
EURGBP	2Y	02-Jan-2015	08-May-2020	2,813,760	EURPLN	2Y	02-Jan-2015	08-May-2020	2,813,760
EURHKD	1D	15-Jun-2017	08-May-2020	1,524,960	EURSEK	1D	15-Jun-2017	08-May-2020	1,524,960
EURHKD	1M	02-Jan-2015	08-May-2020	2,813,760	EURSEK	1M	02-Jan-2015	08-May-2020	2,813,760
EURHKD	2M	02-Jan-2015	08-May-2020	2,813,760	EURSEK	2M	02-Jan-2015	08-May-2020	2,813,760
EURHKD	3M	02-Jan-2015	08-May-2020	2,813,760	EURSEK	3M	02-Jan-2015	08-May-2020	2,813,760
EURHKD	6M	02-Jan-2015	08-May-2020	2,813,760	EURSEK	6M	02-Jan-2015	08-May-2020	2,813,760
EURHKD	9M	15-Jun-2017	08-May-2020	1,524,960	EURSEK	9M	15-Jun-2017	08-May-2020	1,524,960
EURHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	EURSEK	1Y	02-Jan-2015	08-May-2020	2,813,760
EURHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	EURSEK	2Y	02-Jan-2015	08-May-2020	2,813,760
EURHUF	1D	15-Jun-2017	08-May-2020	1,524,960	EURSGD	1D	15-Jun-2017	08-May-2020	1,524,960

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
EURHUF	1M	02-Jan-2015	08-May-2020	2,813,760	EURSGD	1M	02-Jan-2015	08-May-2020	2,813,760
EURHUF	2M	02-Jan-2015	08-May-2020	2,813,760	EURSGD	2M	02-Jan-2015	08-May-2020	2,813,760
EURHUF	3M	02-Jan-2015	08-May-2020	2,813,760	EURSGD	3M	02-Jan-2015	08-May-2020	2,813,760
EURHUF	6M	02-Jan-2015	08-May-2020	2,813,760	EURSGD	6M	02-Jan-2015	08-May-2020	2,813,760
EURHUF	9M	15-Jun-2017	08-May-2020	1,524,960	EURSGD	9M	15-Jun-2017	08-May-2020	1,524,960
EURHUF	1Y	02-Jan-2015	08-May-2020	2,813,760	EURSGD	1Y	02-Jan-2015	08-May-2020	2,813,760
EURHUF	2Y	02-Jan-2015	08-May-2020	2,813,760	EURSGD	2Y	02-Jan-2015	08-May-2020	2,813,760
EURILS	1D	15-Jun-2017	08-May-2020	1,524,960	EURTRY	1D	15-Jun-2017	08-May-2020	1,524,960
EURILS	1M	02-Jan-2015	08-May-2020	2,813,760	EURTRY	1M	02-Jan-2015	08-May-2020	2,813,760
EURILS	2M	02-Jan-2015	08-May-2020	2,813,760	EURTRY	2M	02-Jan-2015	08-May-2020	2,813,760
EURILS	3M	02-Jan-2015	08-May-2020	2,813,760	EURTRY	3M	02-Jan-2015	08-May-2020	2,813,760
EURILS	6M	02-Jan-2015	08-May-2020	2,813,760	EURTRY	6M	02-Jan-2015	08-May-2020	2,813,760
EURILS	9M	15-Jun-2017	08-May-2020	1,524,960	EURTRY	9M	15-Jun-2017	08-May-2020	1,524,960
EURILS	1Y	02-Jan-2015	08-May-2020	2,813,760	EURTRY	1Y	02-Jan-2015	08-May-2020	2,813,760
EURILS	2Y	02-Jan-2015	08-May-2020	2,813,760	EURTRY	2Y	02-Jan-2015	08-May-2020	2,813,760
EURINR	1D	11-Mar-2020	08-May-2020	84,960	EURUSD	1D	15-Jun-2017	08-May-2020	1,524,960
EURINR	1M	11-Mar-2020	08-May-2020	84,960	EURUSD	1M	02-Jan-2015	08-May-2020	2,813,760
EURINR	2M	11-Mar-2020	08-May-2020	84,960	EURUSD	2M	02-Jan-2015	08-May-2020	2,813,760
EURINR	3M	11-Mar-2020	08-May-2020	84,960	EURUSD	3M	02-Jan-2015	08-May-2020	2,813,760
EURINR	6M	11-Mar-2020	08-May-2020	84,960	EURUSD	6M	02-Jan-2015	08-May-2020	2,813,760
EURINR	9M	11-Mar-2020	08-May-2020	84,960	EURUSD	9M	15-Jun-2017	08-May-2020	1,524,960
EURINR	1Y	11-Mar-2020	08-May-2020	84,960	EURUSD	1Y	02-Jan-2015	08-May-2020	2,813,760
EURINR	2Y	11-Mar-2020	08-May-2020	84,960	EURUSD	2Y	02-Jan-2015	08-May-2020	2,813,760
EURJPY	1D	15-Jun-2017	08-May-2020	1,524,960	EURZAR	1D	15-Jun-2017	08-May-2020	1,524,960
EURJPY	1M	02-Jan-2015	08-May-2020	2,813,760	EURZAR	1M	02-Jan-2015	08-May-2020	2,813,760
EURJPY	2M	02-Jan-2015	08-May-2020	2,813,760	EURZAR	2M	02-Jan-2015	08-May-2020	2,813,760
EURJPY	3M	02-Jan-2015	08-May-2020	2,813,760	EURZAR	3M	02-Jan-2015	08-May-2020	2,813,760
EURJPY	6M	02-Jan-2015	08-May-2020	2,813,760	EURZAR	6M	02-Jan-2015	08-May-2020	2,813,760
EURJPY	9M	15-Jun-2017	08-May-2020	1,524,960	EURZAR	9M	15-Jun-2017	08-May-2020	1,524,960
EURJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	EURZAR	1Y	02-Jan-2015	08-May-2020	2,813,760
EURJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	EURZAR	2Y	02-Jan-2015	08-May-2020	2,813,760
EURKRW	1D	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	1D	15-Jun-2017	08-May-2020	1,524,960
EURKRW	1M	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	1M	02-Jan-2015	08-May-2020	2,813,760
EURKRW	2M	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	2M	02-Jan-2015	08-May-2020	2,813,760
EURKRW	3M	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	3M	02-Jan-2015	08-May-2020	2,813,760
EURKRW	6M	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	6M	02-Jan-2015	08-May-2020	2,813,760
EURKRW	9M	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	9M	15-Jun-2017	08-May-2020	1,524,960
EURKRW	1Y	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	1Y	02-Jan-2015	08-May-2020	2,813,760
EURKRW	2Y	03-Jul-2018	10-Jan-2019	276,480	GBPAUD	2Y	15-Jun-2017	08-May-2020	1,524,960
GBPCAD	1D	15-Jun-2017	08-May-2020	1,524,960	HKDJPY	1D	15-Jun-2017	08-May-2020	1,524,960
GBPCAD	1M	02-Jan-2015	08-May-2020	2,813,760	HKDJPY	1M	02-Jan-2015	08-May-2020	2,813,760
GBPCAD	2M	02-Jan-2015	08-May-2020	2,813,760	HKDJPY	2M	02-Jan-2015	08-May-2020	2,813,760
GBPCAD	3M	02-Jan-2015	08-May-2020	2,813,760	HKDJPY	3M	02-Jan-2015	08-May-2020	2,813,760
GBPCAD	6M	02-Jan-2015	08-May-2020	2,813,760	HKDJPY	6M	02-Jan-2015	08-May-2020	2,813,760
GBPCAD	9M	15-Jun-2017	08-May-2020	1,524,960	HKDJPY	9M	15-Jun-2017	08-May-2020	1,524,960
GBPCAD	1Y	07-Jan-2015	08-May-2020	2,806,560	HKDJPY	1Y	02-Jan-2015	08-May-2020	2,813,760
GBPCAD	2Y	02-Jan-2015	08-May-2020	2,813,760	HKDJPY	2Y	02-Jan-2015	08-May-2020	2,813,760
GBPCHF	1D	15-Jun-2017	08-May-2020	1,524,960	HKDSGD	1D	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	1M	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	1M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	2M	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	2M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	3M	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	3M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	6M	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	6M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	9M	15-Jun-2017	08-May-2020	1,524,960	HKDSGD	9M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	1Y	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	1Y	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	2Y	02-Jan-2015	08-May-2020	2,813,760	HKDSGD	2Y	08-Jun-2018	08-May-2020	1,009,440
GBPDKK	1M	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	1D	08-Jun-2018	08-May-2020	1,009,440
GBPDKK	2M	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	1M	08-Jun-2018	08-May-2020	1,009,440
GBPDKK	3M	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	2M	08-Jun-2018	08-May-2020	1,009,440
GBPDKK	6M	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	3M	08-Jun-2018	08-May-2020	1,009,440
GBPDKK	1Y	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	6M	08-Jun-2018	08-May-2020	1,009,440

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
GBPDKK	2Y	02-Jan-2015	21-Jul-2016	816,480	IDRJPY	9M	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	1Y	02-Jan-2015	08-May-2020	2,813,760	IDRJPY	1Y	08-Jun-2018	08-May-2020	1,009,440
GBPCHF	2Y	02-Jan-2015	08-May-2020	2,813,760	IDRJPY	2Y	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	1D	15-Jun-2017	08-May-2020	1,524,960	JPYKRW	1D	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	1M	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	1M	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	2M	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	2M	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	3M	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	3M	20-Jun-2018	08-May-2020	992,160
GBPHKD	6M	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	6M	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	9M	15-Jun-2017	08-May-2020	1,524,960	JPYKRW	9M	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	1Y	08-Jun-2018	08-May-2020	1,009,440
GBPHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	JPYKRW	2Y	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	1D	15-Jun-2017	08-May-2020	1,524,960	JPYTWD	1D	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	1M	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	1M	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	2M	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	2M	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	3M	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	3M	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	6M	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	6M	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	9M	15-Jun-2017	08-May-2020	1,524,960	JPYTWD	9M	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	1Y	08-Jun-2018	08-May-2020	1,009,440
GBPJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	JPYTWD	2Y	08-Jun-2018	06-Apr-2020	963,360
GBPNOK	1D	15-Jun-2017	08-May-2020	1,524,960	KRWJPY	1D	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	1M	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	1M	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	2M	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	2M	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	3M	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	3M	20-Jun-2018	08-May-2020	992,160
GBPNOK	6M	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	6M	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	9M	15-Jun-2017	08-May-2020	1,524,960	KRWJPY	9M	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	1Y	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	1Y	08-Jun-2018	08-May-2020	1,009,440
GBPNOK	2Y	02-Jan-2015	08-May-2020	2,813,760	KRWJPY	2Y	08-Jun-2018	08-May-2020	1,009,440
GBPNZD	1D	15-Jun-2017	08-May-2020	1,524,960	MXNJPY	1D	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	1M	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	1M	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	2M	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	2M	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	3M	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	3M	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	6M	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	6M	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	9M	15-Jun-2017	08-May-2020	1,524,960	MXNJPY	9M	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	1Y	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	1Y	08-Jun-2018	10-Jan-2019	312,480
GBPNZD	2Y	02-Jan-2015	08-May-2020	2,813,760	MXNJPY	2Y	08-Jun-2018	10-Jan-2019	312,480
GBPSEK	1D	15-Jun-2017	08-May-2020	1,524,960	NOKJPY	1D	15-Jun-2017	08-May-2020	1,524,960
GBPSEK	1M	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	1M	02-Jan-2015	08-May-2020	2,813,760
GBPSEK	2M	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	2M	02-Jan-2015	08-May-2020	2,813,760
GBPSEK	3M	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	3M	02-Jan-2015	08-May-2020	2,813,760
GBPSEK	6M	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	6M	02-Jan-2015	08-May-2020	2,813,760
GBPSEK	9M	15-Jun-2017	08-May-2020	1,524,960	NOKJPY	9M	15-Jun-2017	08-May-2020	1,524,960
GBPSEK	1Y	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	1Y	02-Jan-2015	08-May-2020	2,813,760
GBPSEK	2Y	02-Jan-2015	08-May-2020	2,813,760	NOKJPY	2Y	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	1D	15-Jun-2017	08-May-2020	1,524,960	NOKSEK	1D	15-Jun-2017	08-May-2020	1,524,960
GBPSGD	1M	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	1M	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	2M	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	2M	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	3M	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	3M	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	6M	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	6M	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	9M	15-Jun-2017	08-May-2020	1,524,960	NOKSEK	9M	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	1Y	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	1Y	02-Jan-2015	08-May-2020	2,813,760
GBPSGD	2Y	02-Jan-2015	08-May-2020	2,813,760	NOKSEK	2Y	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	1D	15-Jun-2017	08-May-2020	1,524,960	NZDCAD	1D	15-Jun-2017	08-May-2020	1,524,960
GBPUSD	1M	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	1M	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	2M	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	2M	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	3M	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	3M	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	6M	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	6M	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	9M	15-Jun-2017	08-May-2020	1,524,960	NZDCAD	9M	15-Jun-2017	08-May-2020	1,524,960
GBPUSD	1Y	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	1Y	02-Jan-2015	08-May-2020	2,813,760
GBPUSD	2Y	02-Jan-2015	08-May-2020	2,813,760	NZDCAD	2Y	02-Jan-2015	08-May-2020	2,813,760

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
NZDCHF	1D	15-Jun-2017	08-May-2020	1,524,960	SGDJPY	1D	15-Jun-2017	08-May-2020	1,524,960
NZDCHF	1M	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	1M	02-Jan-2015	08-May-2020	2,813,760
NZDCHF	2M	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	2M	02-Jan-2015	08-May-2020	2,813,760
NZDCHF	3M	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	3M	02-Jan-2015	08-May-2020	2,813,760
NZDCHF	6M	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	6M	02-Jan-2015	08-May-2020	2,813,760
NZDCHF	9M	15-Jun-2017	08-May-2020	1,524,960	SGDJPY	9M	15-Jun-2017	08-May-2020	1,524,960
NZDCHF	1Y	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	1Y	02-Jan-2015	08-May-2020	2,813,760
NZDCHF	2Y	02-Jan-2015	08-May-2020	2,813,760	SGDJPY	2Y	02-Jan-2015	08-May-2020	2,813,760
NZDHKD	1D	15-Jun-2017	08-May-2020	1,524,960	SGDMYR	1D	19-Jun-2018	08-May-2020	993,600
NZDHKD	1M	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	1M	19-Jun-2018	08-May-2020	993,600
NZDHKD	2M	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	2M	19-Jun-2018	08-May-2020	993,600
NZDHKD	3M	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	3M	19-Jun-2018	08-May-2020	993,600
NZDHKD	6M	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	6M	19-Jun-2018	08-May-2020	993,600
NZDHKD	9M	15-Jun-2017	08-May-2020	1,524,960	SGDMYR	9M	19-Jun-2018	08-May-2020	993,600
NZDHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	1Y	19-Jun-2018	08-May-2020	993,600
NZDHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	SGDMYR	2Y	19-Jun-2018	08-May-2020	993,600
NZDJPY	1D	15-Jun-2017	08-May-2020	1,524,960	USDBRL	1D	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	1M	02-Jan-2015	08-May-2020	2,813,760	USDBRL	1M	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	2M	02-Jan-2015	08-May-2020	2,813,760	USDBRL	2M	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	3M	02-Jan-2015	08-May-2020	2,813,760	USDBRL	3M	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	6M	02-Jan-2015	08-May-2020	2,813,760	USDBRL	6M	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	9M	15-Jun-2017	08-May-2020	1,524,960	USDBRL	9M	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	USDBRL	1Y	15-Jun-2017	08-May-2020	1,524,960
NZDJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	USDBRL	2Y	15-Jun-2017	08-May-2020	1,524,960
NZDNOK	1D	15-Jun-2017	08-May-2020	1,524,960	USDCAD	1D	15-Jun-2017	08-May-2020	1,524,960
NZDNOK	1M	02-Jan-2015	08-May-2020	2,813,760	USDCAD	1M	02-Jan-2015	08-May-2020	2,813,760
NZDNOK	2M	02-Jan-2015	08-May-2020	2,813,760	USDCAD	2M	02-Jan-2015	08-May-2020	2,813,760
NZDNOK	3M	02-Jan-2015	08-May-2020	2,813,760	USDCAD	3M	02-Jan-2015	08-May-2020	2,813,760
NZDNOK	6M	02-Jan-2015	08-May-2020	2,813,760	USDCAD	6M	02-Jan-2015	08-May-2020	2,813,760
NZDNOK	9M	15-Jun-2017	08-May-2020	1,524,960	USDCAD	9M	15-Jun-2017	08-May-2020	1,524,960
NZDNOK	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCAD	1Y	02-Jan-2015	08-May-2020	2,813,760
NZDNOK	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCAD	2Y	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	1D	15-Jun-2017	08-May-2020	1,524,960	USDCHF	1D	15-Jun-2017	08-May-2020	1,524,960
NZDSEK	1M	02-Jan-2015	08-May-2020	2,813,760	USDCHF	1M	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	2M	02-Jan-2015	08-May-2020	2,813,760	USDCHF	2M	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	3M	02-Jan-2015	08-May-2020	2,813,760	USDCHF	3M	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	6M	02-Jan-2015	08-May-2020	2,813,760	USDCHF	6M	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	9M	15-Jun-2017	08-May-2020	1,524,960	USDCHF	9M	15-Jun-2017	08-May-2020	1,524,960
NZDSEK	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCHF	1Y	02-Jan-2015	08-May-2020	2,813,760
NZDSEK	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCHF	2Y	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	1D	15-Jun-2017	08-May-2020	1,524,960	USDCLP	1D	15-Jun-2017	08-May-2020	1,524,960
NZDSGD	1M	02-Jan-2015	08-May-2020	2,813,760	USDCLP	1M	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	2M	02-Jan-2015	08-May-2020	2,813,760	USDCLP	2M	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	3M	02-Jan-2015	07-May-2020	2,812,320	USDCLP	3M	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	6M	02-Jan-2015	08-May-2020	2,813,760	USDCLP	6M	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	9M	15-Jun-2017	08-May-2020	1,524,960	USDCLP	9M	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCLP	1Y	02-Jan-2015	08-May-2020	2,813,760
NZDSGD	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCLP	2Y	15-Jun-2017	08-May-2020	1,524,960
NZDUSD	1D	15-Jun-2017	08-May-2020	1,524,960	USDCNH	1D	15-Jun-2017	08-May-2020	1,524,960
NZDUSD	1M	02-Jan-2015	08-May-2020	2,813,760	USDCNH	1M	02-Jan-2015	08-May-2020	2,813,760
NZDUSD	2M	02-Jan-2015	08-May-2020	2,813,760	USDCNH	2M	02-Jan-2015	08-May-2020	2,813,760
NZDUSD	3M	02-Jan-2015	08-May-2020	2,813,760	USDCNH	3M	02-Jan-2015	08-May-2020	2,813,760
NZDUSD	6M	02-Jan-2015	08-May-2020	2,813,760	USDCNH	6M	02-Jan-2015	08-May-2020	2,813,760
NZDUSD	9M	15-Jun-2017	08-May-2020	1,524,960	USDCNH	9M	15-Jun-2017	08-May-2020	1,524,960
NZDUSD	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCNH	1Y	02-Jan-2015	08-May-2020	2,813,760
NZDUSD	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCNH	2Y	02-Jan-2015	08-May-2020	2,813,760
SEKJPY	1D	15-Jun-2017	08-May-2020	1,524,960	USDCNY	1D	15-Jun-2017	08-May-2020	1,524,960
SEKJPY	1M	02-Jan-2015	08-May-2020	2,813,760	USDCNY	1M	02-Jan-2015	08-May-2020	2,813,760
SEKJPY	2M	02-Jan-2015	08-May-2020	2,813,760	USDCNY	2M	02-Jan-2015	08-May-2020	2,813,760

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
SEKJPY	3M	02-Jan-2015	08-May-2020	2,813,760	USDCNY	3M	02-Jan-2015	08-May-2020	2,813,760
SEKJPY	6M	02-Jan-2015	08-May-2020	2,813,760	USDCNY	6M	02-Jan-2015	08-May-2020	2,813,760
SEKJPY	9M	15-Jun-2017	08-May-2020	1,524,960	USDCNY	9M	15-Jun-2017	08-May-2020	1,524,960
SEKJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCNY	1Y	02-Jan-2015	08-May-2020	2,813,760
SEKJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCNY	2Y	02-Jan-2015	08-May-2020	2,813,760
SGDCNH	3M	08-Jun-2018	10-Jan-2019	312,480	USDCOP	1D	15-Jun-2017	08-May-2020	1,524,960
SGDCNH	6M	08-Jun-2018	10-Jan-2019	312,480	USDCOP	1M	15-Jun-2017	08-May-2020	1,524,960
SGDCNH	9M	08-Jun-2018	10-Jan-2019	312,480	USDCOP	2M	15-Jun-2017	08-May-2020	1,524,960
SGDCNH	1Y	08-Jun-2018	10-Jan-2019	312,480	USDCOP	3M	15-Jun-2017	08-May-2020	1,524,960
SGDCNH	2Y	08-Jun-2018	10-Jan-2019	312,480	USDCOP	6M	15-Jun-2017	08-May-2020	1,524,960
SEKJPY	9M	15-Jun-2017	08-May-2020	1,524,960	USDCOP	9M	15-Jun-2017	08-May-2020	1,524,960
SEKJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCOP	1Y	15-Jun-2017	08-May-2020	1,524,960
SEKJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCOP	2Y	15-Jun-2017	08-May-2020	1,524,960
SGDHKD	1D	15-Jun-2017	08-May-2020	1,524,960	USDCZK	1D	15-Jun-2017	08-May-2020	1,524,960
SGDHKD	1M	02-Jan-2015	08-May-2020	2,813,760	USDCZK	1M	02-Jan-2015	08-May-2020	2,813,760
SGDHKD	2M	02-Jan-2015	08-May-2020	2,813,760	USDCZK	2M	02-Jan-2015	07-May-2020	2,812,320
SGDHKD	3M	02-Jan-2015	08-May-2020	2,813,760	USDCZK	3M	02-Jan-2015	08-May-2020	2,813,760
SGDHKD	6M	02-Jan-2015	08-May-2020	2,813,760	USDCZK	6M	02-Jan-2015	08-May-2020	2,813,760
SGDHKD	9M	15-Jun-2017	08-May-2020	1,524,960	USDCZK	9M	15-Jun-2017	08-May-2020	1,524,960
SGDHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	USDCZK	1Y	15-Jun-2017	08-May-2020	1,524,960
SGDHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	USDCZK	2Y	02-Jan-2015	08-May-2020	2,813,760
USDDKK	1D	15-Jun-2017	08-May-2020	1,524,960	USDNOK	1D	15-Jun-2017	08-May-2020	1,524,960
USDDKK	1M	02-Jan-2015	08-May-2020	2,813,760	USDNOK	1M	02-Jan-2015	08-May-2020	2,813,760
USDDKK	2M	02-Jan-2015	08-May-2020	2,813,760	USDNOK	2M	02-Jan-2015	08-May-2020	2,813,760
USDDKK	3M	02-Jan-2015	08-May-2020	2,813,760	USDNOK	3M	02-Jan-2015	08-May-2020	2,813,760
USDDKK	6M	02-Jan-2015	08-May-2020	2,813,760	USDNOK	6M	02-Jan-2015	08-May-2020	2,813,760
USDDKK	9M	02-Jan-2015	08-May-2020	2,813,760	USDNOK	9M	02-Jan-2015	08-May-2020	2,813,760
USDDKK	1Y	02-Jan-2015	08-May-2020	2,813,760	USDNOK	1Y	02-Jan-2015	08-May-2020	2,813,760
USDDKK	2Y	02-Jan-2015	08-May-2020	2,813,760	USDNOK	2Y	02-Jan-2015	08-May-2020	2,813,760
USDHKD	1D	15-Jun-2017	08-May-2020	1,524,960	USDPHP	1D	29-Jun-2018	08-May-2020	979,200
USDHKD	1M	05-Jan-2015	08-May-2020	2,809,440	USDPHP	1M	08-Jun-2018	08-May-2020	1,009,440
USDHKD	2M	02-Jan-2015	08-May-2020	2,813,760	USDPHP	2M	29-Jun-2018	08-May-2020	979,200
USDHKD	3M	02-Jan-2015	08-May-2020	2,813,760	USDPHP	3M	29-Jun-2018	08-May-2020	979,200
USDHKD	6M	02-Jan-2015	08-May-2020	2,813,760	USDPHP	6M	29-Jun-2018	08-May-2020	979,200
USDHKD	9M	15-Jun-2017	08-May-2020	1,524,960	USDPHP	9M	29-Jun-2018	08-May-2020	979,200
USDHKD	1Y	02-Jan-2015	08-May-2020	2,813,760	USDPHP	1Y	08-Jun-2018	08-May-2020	1,009,440
USDHKD	2Y	02-Jan-2015	08-May-2020	2,813,760	USDPHP	2Y	08-Jun-2018	08-May-2020	1,009,440
USDHUF	1D	15-Jun-2017	08-May-2020	1,524,960	USDPLN	1D	15-Jun-2017	08-May-2020	1,524,960
USDHUF	1M	02-Jan-2015	08-May-2020	2,813,760	USDPLN	1M	02-Jan-2015	08-May-2020	2,813,760
USDHUF	2M	02-Jan-2015	08-May-2020	2,813,760	USDPLN	2M	02-Jan-2015	08-May-2020	2,813,760
USDHUF	3M	02-Jan-2015	08-May-2020	2,813,760	USDPLN	3M	02-Jan-2015	08-May-2020	2,813,760
USDHUF	6M	02-Jan-2015	08-May-2020	2,813,760	USDPLN	6M	02-Jan-2015	08-May-2020	2,813,760
USDHUF	9M	15-Jun-2017	08-May-2020	1,524,960	USDPLN	9M	15-Jun-2017	08-May-2020	1,524,960
USDHUF	1Y	02-Jan-2015	08-May-2020	2,813,760	USDPLN	1Y	02-Jan-2015	08-May-2020	2,813,760
USDHUF	2Y	02-Jan-2015	08-May-2020	2,813,760	USDPLN	2Y	02-Jan-2015	08-May-2020	2,813,760
USDIDR	1D	15-Jun-2017	08-May-2020	1,524,960	USDSEK	1D	15-Jun-2017	08-May-2020	1,524,960
USDIDR	1M	02-Jan-2015	08-May-2020	2,813,760	USDSEK	1M	02-Jan-2015	08-May-2020	2,813,760
USDIDR	2M	02-Jan-2015	08-May-2020	2,813,760	USDSEK	2M	02-Jan-2015	08-May-2020	2,813,760
USDIDR	3M	02-Jan-2015	08-May-2020	2,813,760	USDSEK	3M	02-Jan-2015	08-May-2020	2,813,760
USDIDR	6M	02-Jan-2015	08-May-2020	2,813,760	USDSEK	6M	02-Jan-2015	08-May-2020	2,813,760
USDIDR	9M	15-Jun-2017	08-May-2020	1,524,960	USDSEK	9M	02-Jan-2015	08-May-2020	2,813,760
USDIDR	1Y	02-Jan-2015	08-May-2020	2,813,760	USDSEK	1Y	02-Jan-2015	08-May-2020	2,813,760
USDIDR	2Y	15-Jun-2017	08-May-2020	1,524,960	USDSEK	2Y	15-Jun-2017	08-May-2020	1,524,960
USDILS	1D	15-Jun-2017	08-May-2020	1,524,960	USDSGD	1D	15-Jun-2017	08-May-2020	1,524,960
USDILS	1M	02-Jan-2015	08-May-2020	2,813,760	USDSGD	1M	02-Jan-2015	08-May-2020	2,813,760
USDILS	2M	02-Jan-2015	08-May-2020	2,813,760	USDSGD	2M	02-Jan-2015	08-May-2020	2,813,760
USDILS	3M	02-Jan-2015	08-May-2020	2,813,760	USDSGD	3M	02-Jan-2015	08-May-2020	2,813,760
USDILS	6M	02-Jan-2015	08-May-2020	2,813,760	USDSGD	6M	02-Jan-2015	08-May-2020	2,813,760
USDILS	9M	15-Jun-2017	08-May-2020	1,524,960	USDSGD	9M	15-Jun-2017	08-May-2020	1,524,960
USDILS	1Y	02-Jan-2015	08-May-2020	2,813,760	USDSGD	1Y	02-Jan-2015	08-May-2020	2,813,760
USDILS	2Y	02-Jan-2015	08-May-2020	2,813,760	USDSGD	2Y	02-Jan-2015	08-May-2020	2,813,760
USDINR	1D	19-Jun-2017	08-May-2020	1,519,200	USDTHB	1D	15-Jun-2017	08-May-2020	1,524,960

Table 5.2: Option Ticker List

Currency	Maturity	StartDate	EndDate	Number of Quotes	Currency	Maturity	StartDate	EndDate	Number of Quotes
USDINR	1M	02-Jan-2015	08-May-2020	2,813,760	USDTHB	1M	02-Jan-2015	08-May-2020	2,813,760
USDINR	2M	02-Jan-2015	08-May-2020	2,813,760	USDTHB	2M	02-Jan-2015	08-May-2020	2,813,760
USDINR	3M	02-Jan-2015	08-May-2020	2,813,760	USDTHB	3M	02-Jan-2015	08-May-2020	2,813,760
USDINR	6M	02-Jan-2015	08-May-2020	2,813,760	USDTHB	6M	02-Jan-2015	08-May-2020	2,813,760
USDINR	9M	15-Jun-2017	08-May-2020	1,524,960	USDTHB	9M	15-Jun-2017	08-May-2020	1,524,960
USDINR	1Y	15-Jun-2017	08-May-2020	1,524,960	USDTHB	1Y	02-Jan-2015	08-May-2020	2,813,760
USDINR	2Y	02-Jan-2015	08-May-2020	2,813,760	USDTHB	2Y	05-Jan-2015	08-May-2020	2,809,440
USDJPY	1D	15-Jun-2017	08-May-2020	1,524,960	USDTRY	1D	15-Jun-2017	08-May-2020	1,524,960
USDJPY	1M	02-Jan-2015	08-May-2020	2,813,760	USDTRY	1M	02-Jan-2015	08-May-2020	2,813,760
USDJPY	2M	02-Jan-2015	08-May-2020	2,813,760	USDTRY	2M	02-Jan-2015	08-May-2020	2,813,760
USDJPY	3M	02-Jan-2015	08-May-2020	2,813,760	USDTRY	3M	02-Jan-2015	08-May-2020	2,813,760
USDJPY	6M	02-Jan-2015	08-May-2020	2,813,760	USDTRY	6M	02-Jan-2015	08-May-2020	2,813,760
USDJPY	9M	15-Jun-2017	08-May-2020	1,524,960	USDTRY	9M	15-Jun-2017	08-May-2020	1,524,960
USDJPY	1Y	02-Jan-2015	08-May-2020	2,813,760	USDTRY	1Y	02-Jan-2015	08-May-2020	2,813,760
USDJPY	2Y	02-Jan-2015	08-May-2020	2,813,760	USDTRY	2Y	02-Jan-2015	08-May-2020	2,813,760
USDKRW	1D	15-Jun-2017	08-May-2020	1,524,960	USDTWD	1D	15-Jun-2017	08-May-2020	1,524,960
USDKRW	1M	02-Jan-2015	08-May-2020	2,813,760	USDTWD	1M	02-Jan-2015	08-May-2020	2,813,760
USDKRW	2M	02-Jan-2015	08-May-2020	2,813,760	USDTWD	2M	02-Jan-2015	08-May-2020	2,813,760
USDKRW	3M	02-Jan-2015	08-May-2020	2,813,760	USDTWD	3M	02-Jan-2015	08-May-2020	2,813,760
USDKRW	6M	02-Jan-2015	08-May-2020	2,813,760	USDTWD	6M	02-Jan-2015	08-May-2020	2,813,760
USDKRW	9M	15-Jun-2017	08-May-2020	1,524,960	USDTWD	9M	15-Jun-2017	08-May-2020	1,524,960
USDKRW	1Y	02-Jan-2015	08-May-2020	2,813,760	USDTWD	1Y	02-Jan-2015	08-May-2020	2,813,760
USDKRW	2Y	15-Jun-2017	08-May-2020	1,524,960	USDTWD	2Y	02-Jan-2015	08-May-2020	2,813,760
USDMXN	1D	15-Jun-2017	08-May-2020	1,524,960	USDZAR	1D	15-Jun-2017	08-May-2020	1,524,960
USDMXN	1M	02-Jan-2015	08-May-2020	2,813,760	USDZAR	1M	02-Jan-2015	08-May-2020	2,813,760
USDMXN	2M	02-Jan-2015	08-May-2020	2,813,760	USDZAR	2M	02-Jan-2015	08-May-2020	2,813,760
USDMXN	3M	02-Jan-2015	08-May-2020	2,813,760	USDZAR	3M	02-Jan-2015	08-May-2020	2,813,760
USDMXN	6M	02-Jan-2015	08-May-2020	2,813,760	USDZAR	6M	02-Jan-2015	08-May-2020	2,813,760
USDMXN	9M	15-Jun-2017	08-May-2020	1,524,960	USDZAR	9M	15-Jun-2017	08-May-2020	1,524,960
USDMXN	1Y	02-Jan-2015	08-May-2020	2,813,760	USDZAR	1Y	02-Jan-2015	08-May-2020	2,813,760
USDMXN	2Y	02-Jan-2015	08-May-2020	2,813,760	USDZAR	2Y	02-Jan-2015	08-May-2020	2,813,760
USDMYR	1D	15-Jun-2017	08-May-2020	1,524,960	XAGUSD	1D	15-Jun-2017	08-May-2020	1,524,960
USDMYR	1M	02-Jan-2015	08-May-2020	2,813,760	XAGUSD	1M	02-Jan-2015	08-May-2020	2,813,760
USDMYR	2M	02-Jan-2015	08-May-2020	2,813,760	XAGUSD	2M	02-Jan-2015	08-May-2020	2,813,760
USDMYR	3M	02-Jan-2015	08-May-2020	2,813,760	XAGUSD	3M	02-Jan-2015	08-May-2020	2,813,760
USDMYR	6M	02-Jan-2015	08-May-2020	2,813,760	XAGUSD	6M	02-Jan-2015	08-May-2020	2,813,760
USDMYR	9M	15-Jun-2017	08-May-2020	1,524,960	XAGUSD	9M	02-Jan-2015	08-May-2020	2,813,760
USDMYR	1Y	02-Jan-2015	08-May-2020	2,813,760	XAGUSD	1Y	02-Jan-2015	08-May-2020	2,813,760
USDMYR	2Y	15-Jun-2017	08-May-2020	1,524,960	XAGUSD	2Y	02-Jan-2015	08-May-2020	2,813,760
XAUUSD	1D	15-Jun-2017	08-May-2020	1,524,960	XAUUSD	6M	02-Jan-2015	08-May-2020	2,813,760
XAUUSD	1M	02-Jan-2015	08-May-2020	2,813,760	XAUUSD	9M	02-Jan-2015	08-May-2020	2,813,760
XAUUSD	2M	02-Jan-2015	08-May-2020	2,813,760	XAUUSD	1Y	02-Jan-2015	08-May-2020	2,813,760
XAUUSD	3M	02-Jan-2015	08-May-2020	2,813,760	XAUUSD	2Y	02-Jan-2015	08-May-2020	2,813,760

Appendix B

Figure 5.1: Minimum eigenvalue versus market data: AUDCAD put10

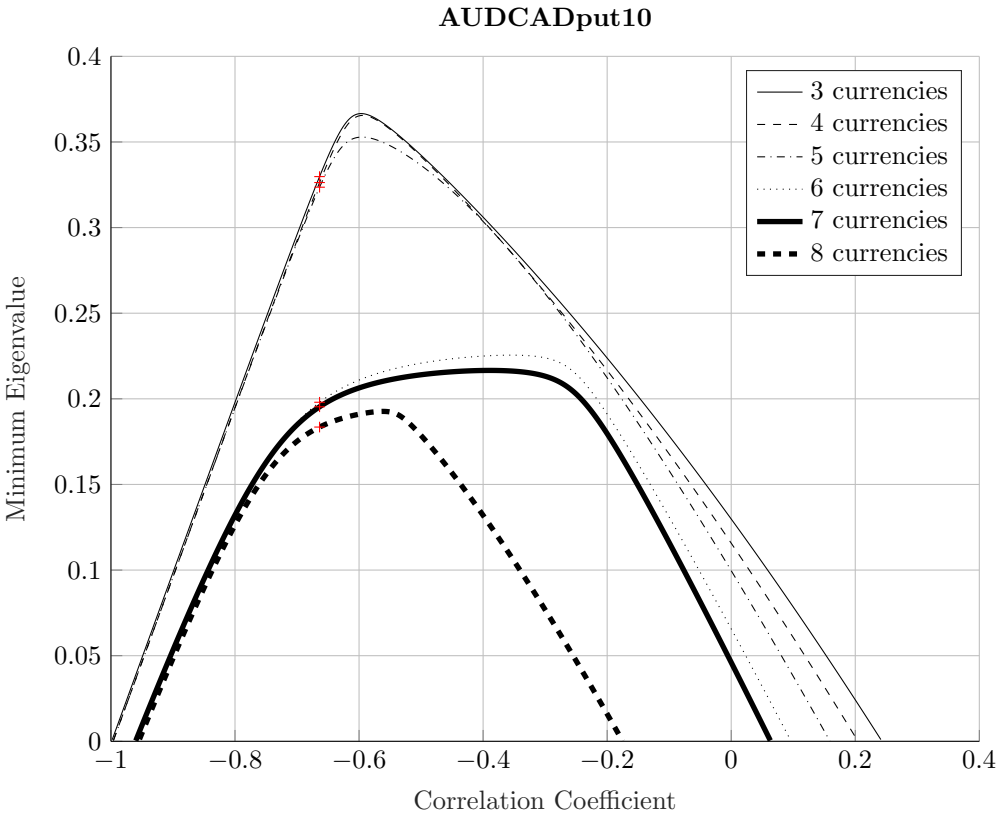


Figure 5.2: Minimum eigenvalue versus market data: AUDCAD put25

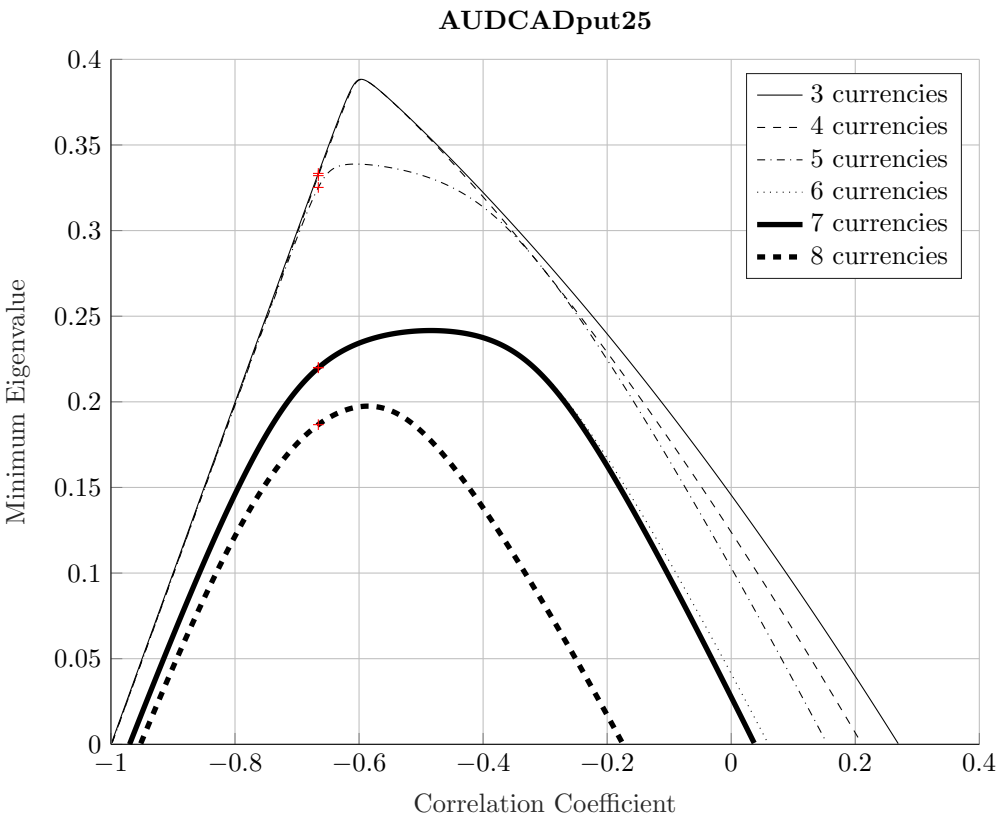


Figure 5.3: Minimum eigenvalue versus market data: AUDCAD call10

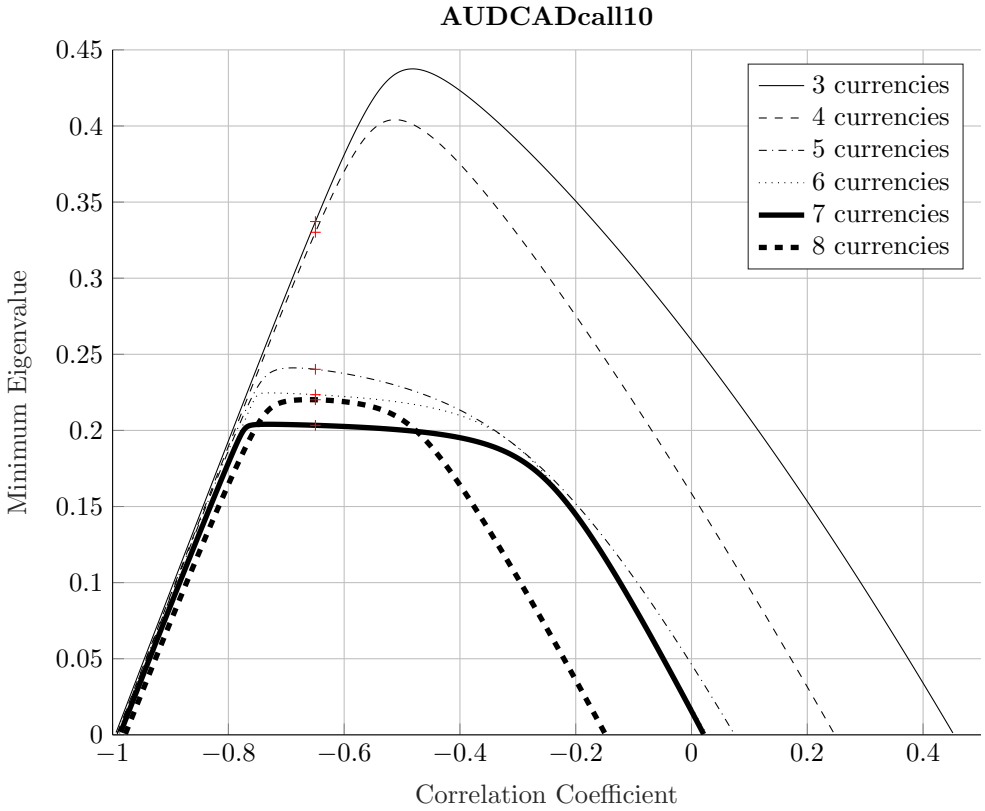


Figure 5.4: Minimum eigenvalue versus market data: AUDCAD call25

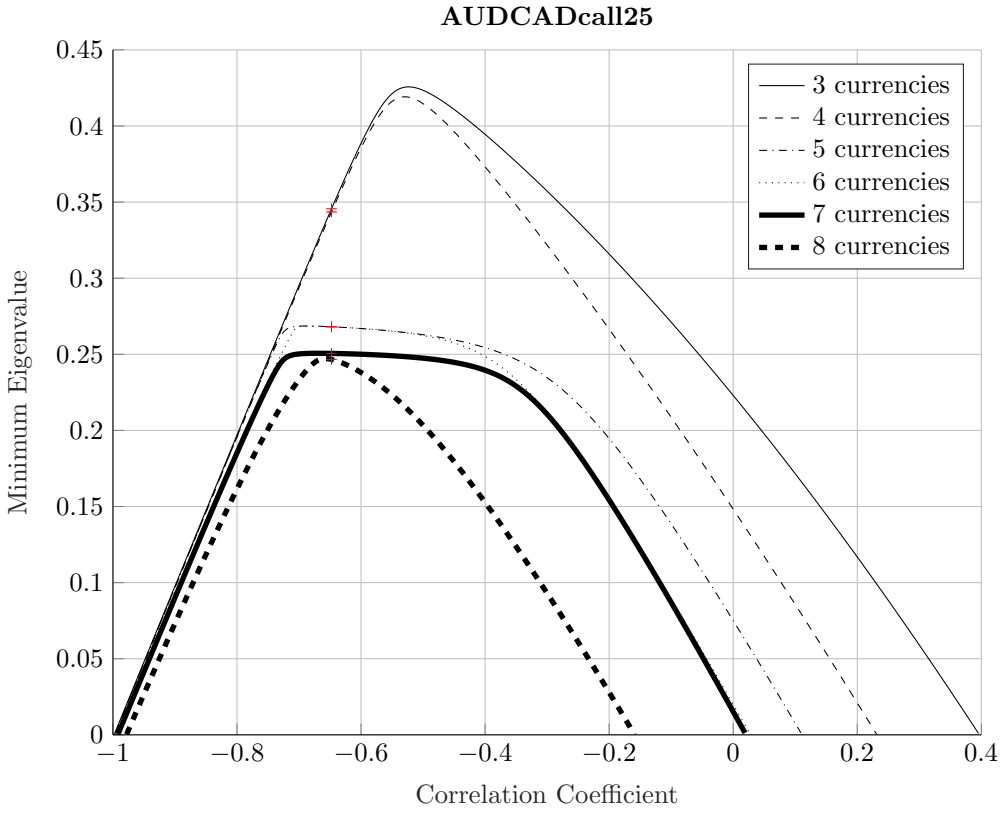


Figure 5.5: Minimum eigenvalue versus market data: CHFJPY atm

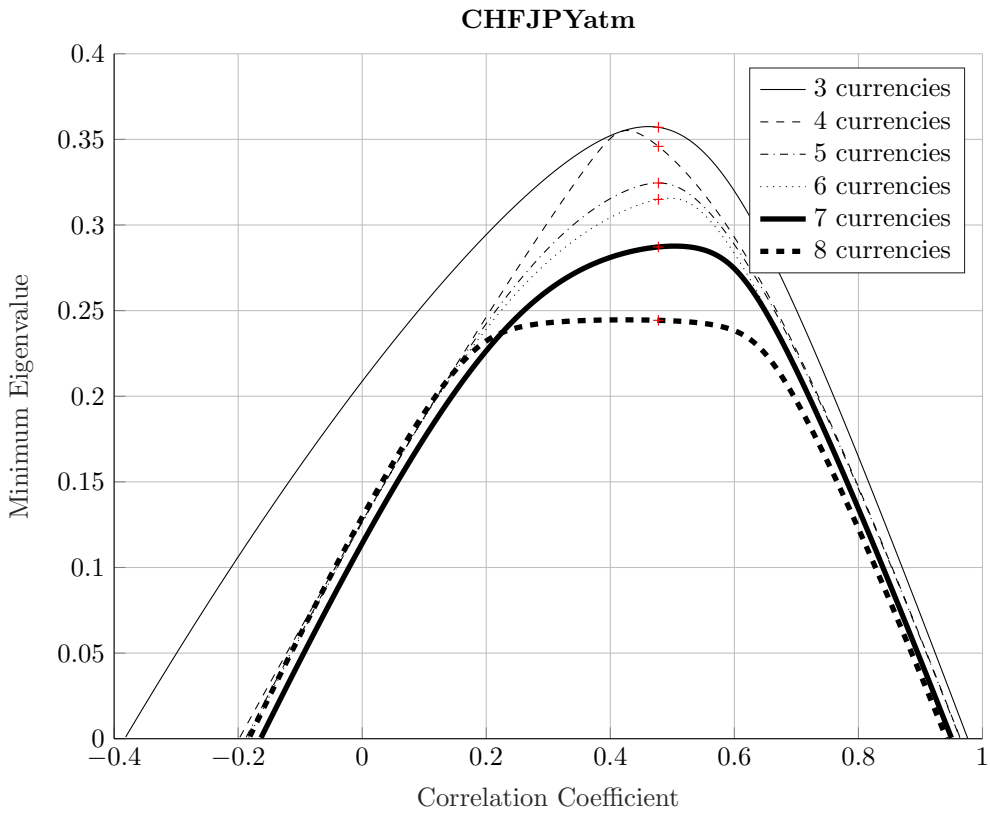


Figure 5.6: Minimum eigenvalue versus market data: CHFJPY put10

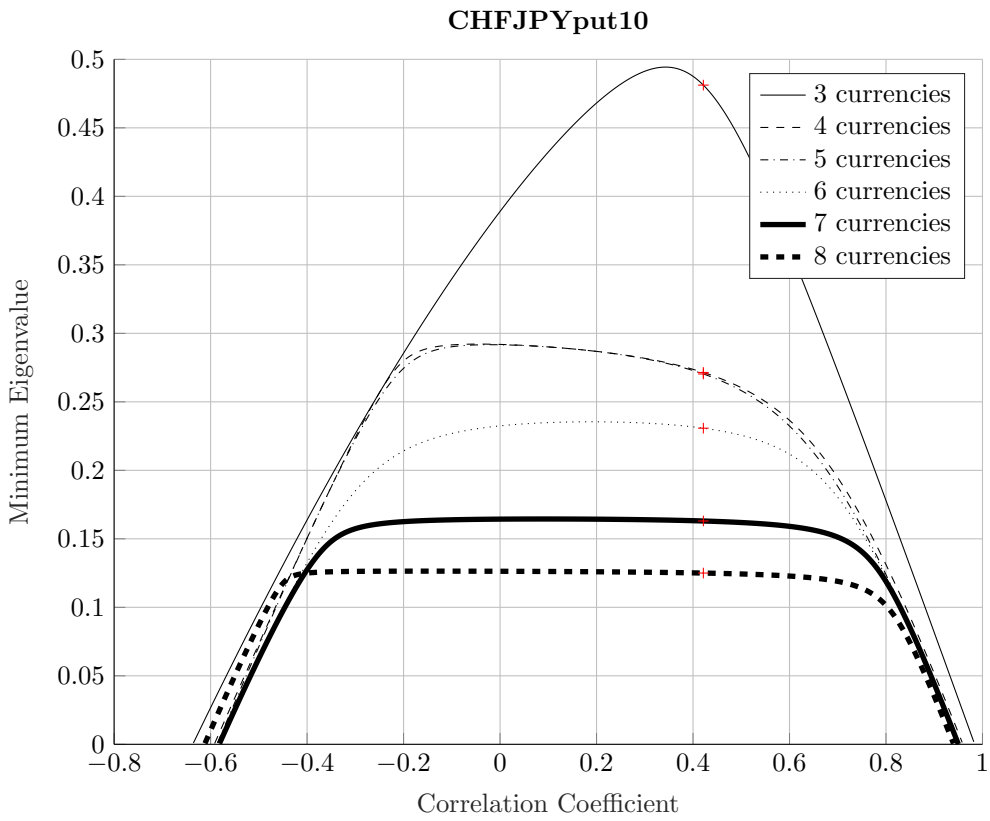


Figure 5.7: Minimum eigenvalue versus market data: CHFJPY put25

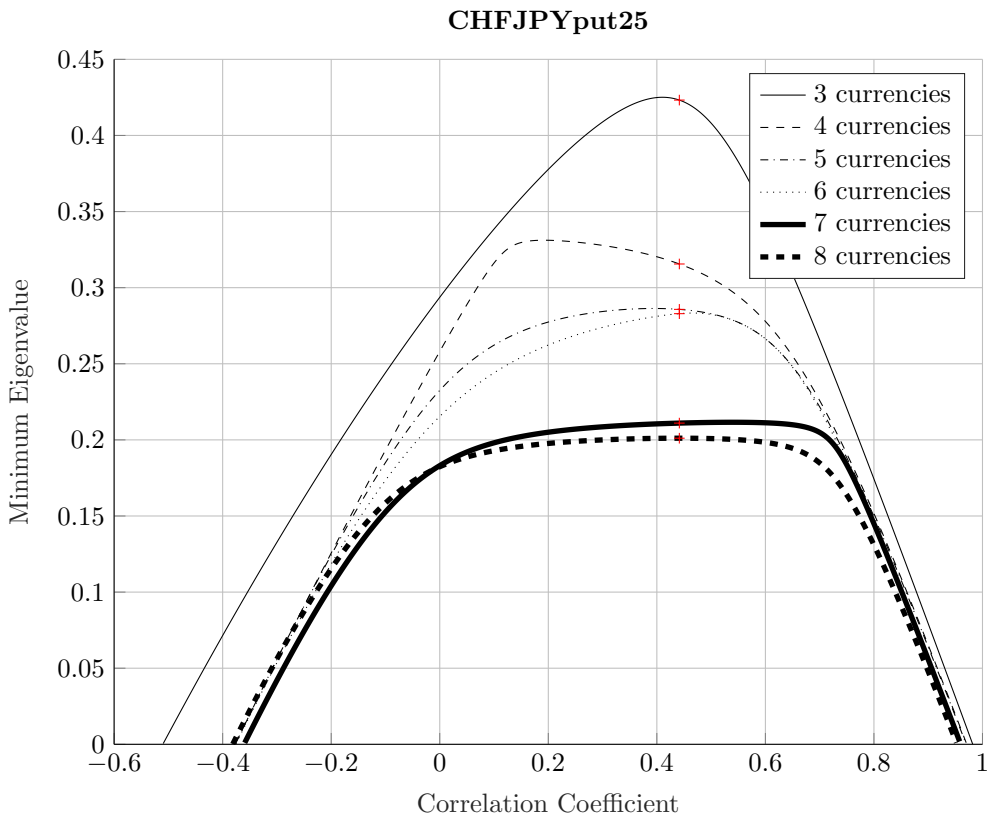


Figure 5.8: Minimum eigenvalue versus market data:CHFJPY call10

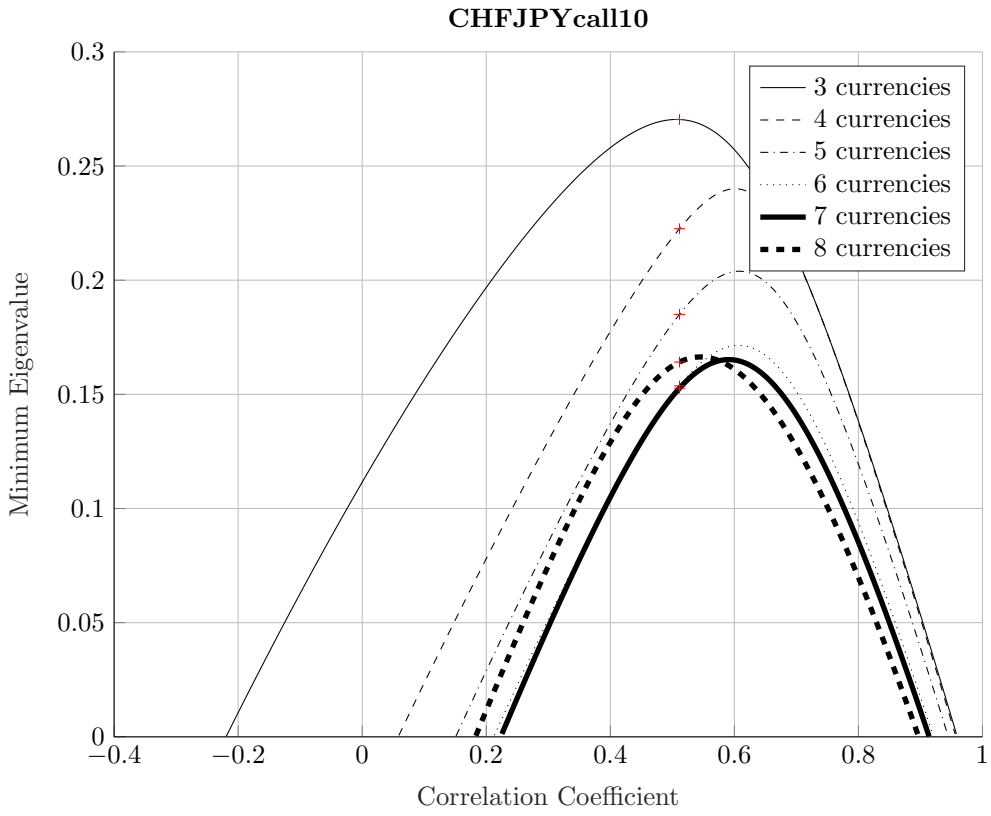


Figure 5.9: Minimum eigenvalue versus market data: CHFJPY call25

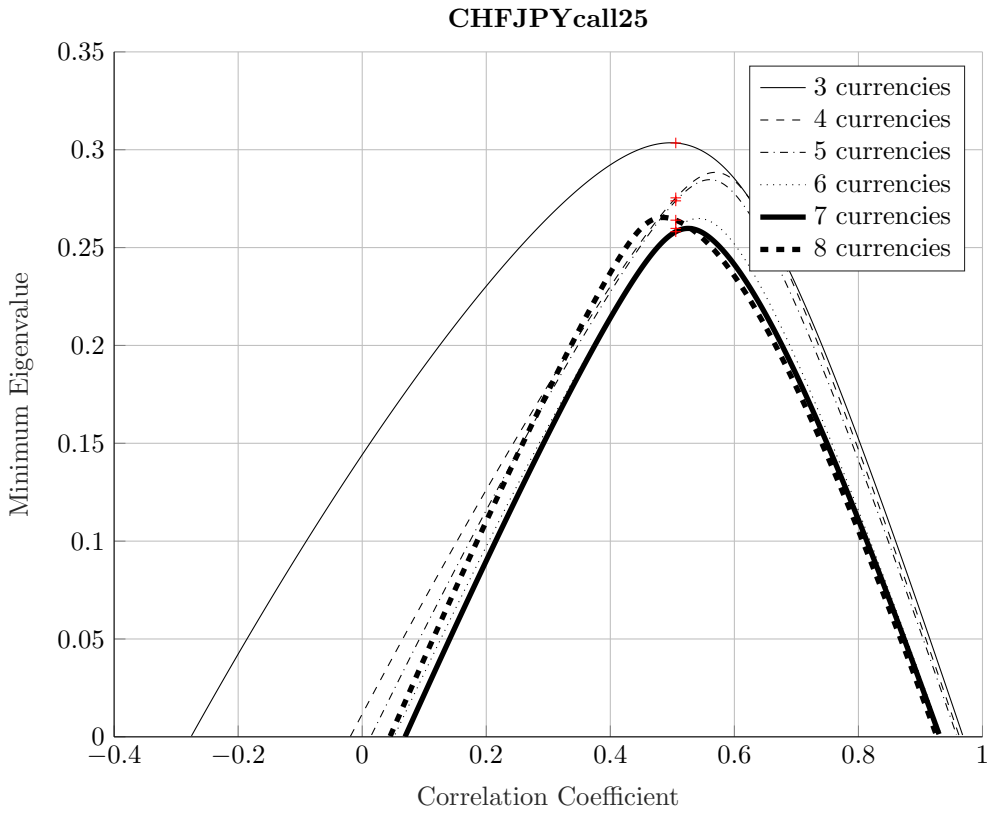


Figure 5.10: Minimum eigenvalue versus market data: EURGBP atm

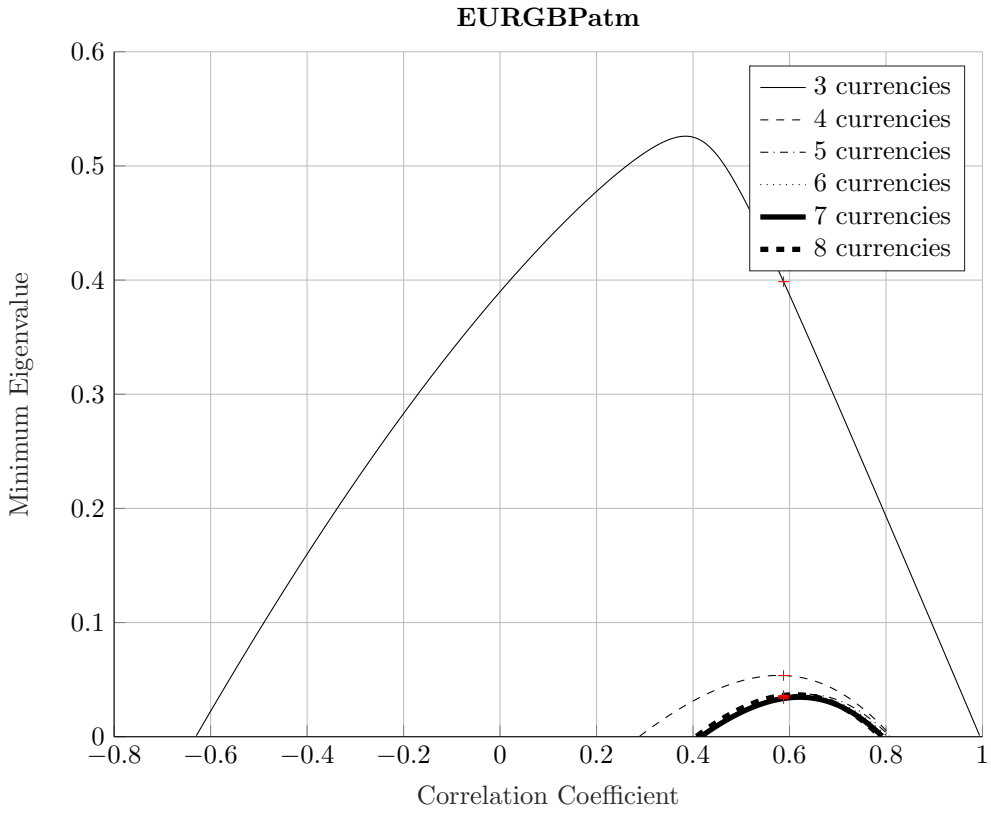


Figure 5.11: Minimum eigenvalue versus market data: EURGBP put10

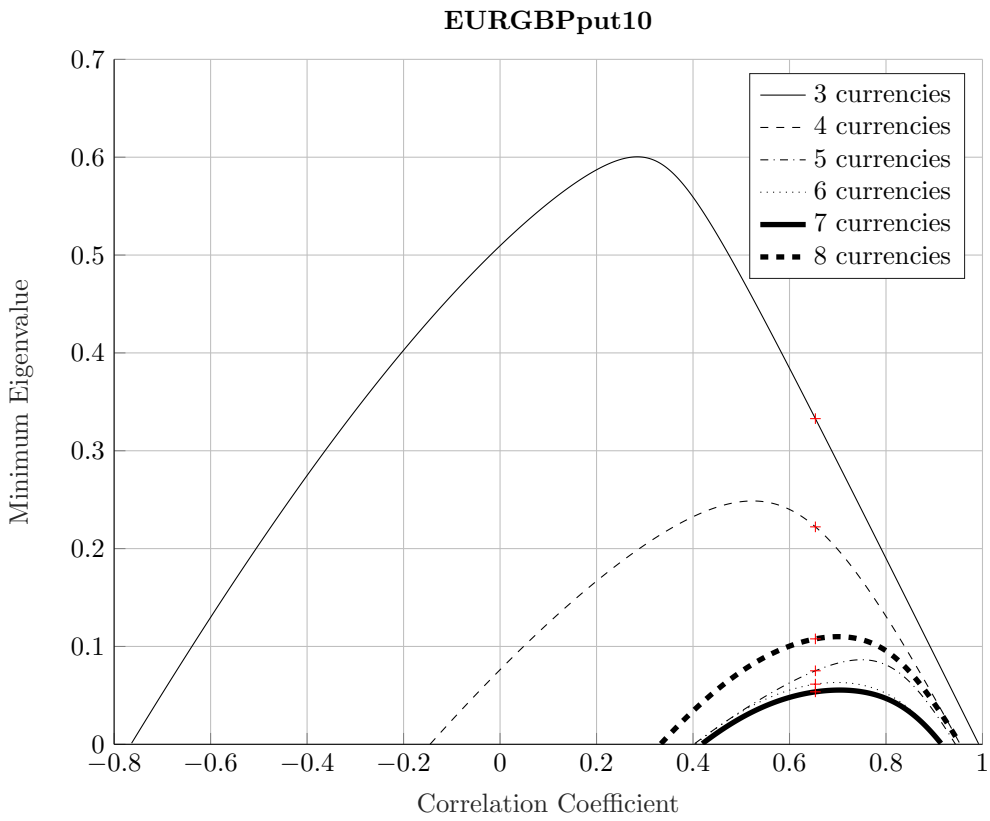


Figure 5.12: Minimum eigenvalue versus market data: EURGBP put25

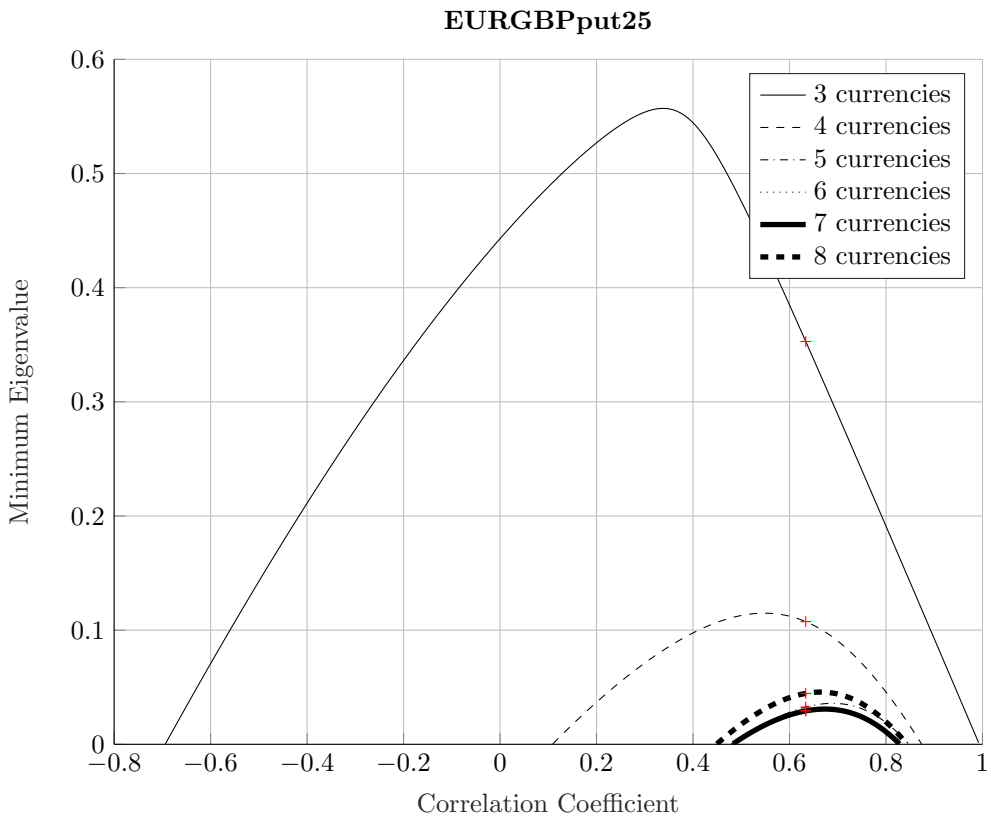


Figure 5.13: Minimum eigenvalue versus market data: EURGBP put10

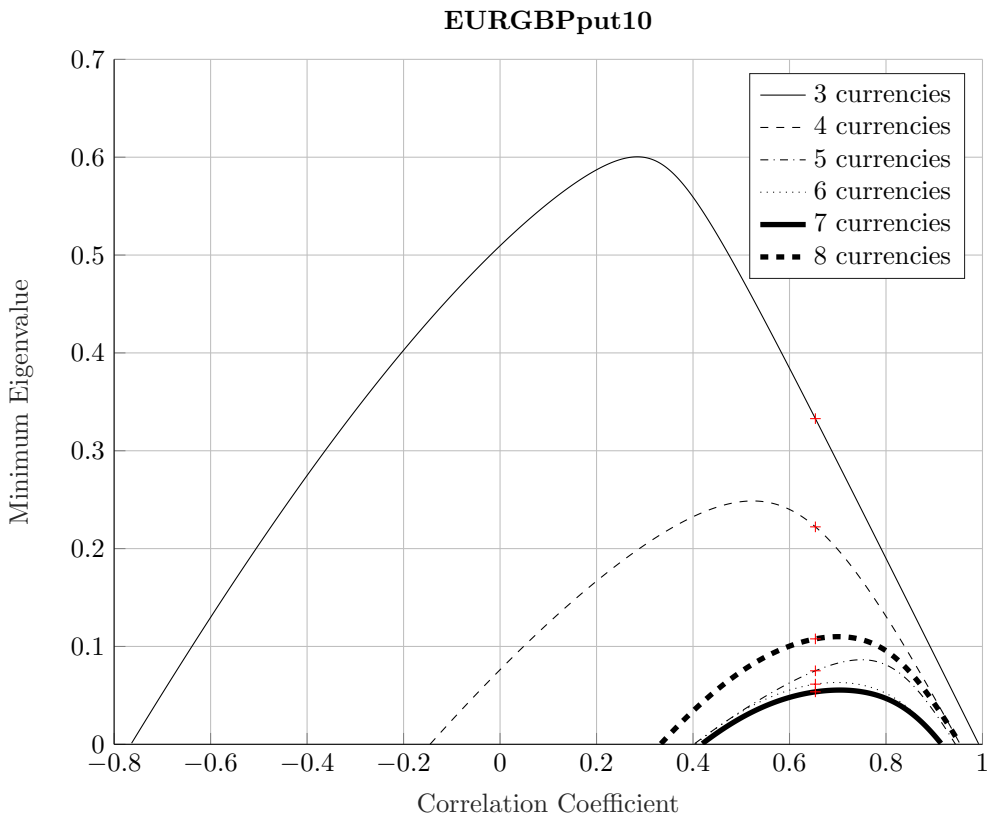


Figure 5.14: Minimum eigenvalue versus market data: EURGBP put25

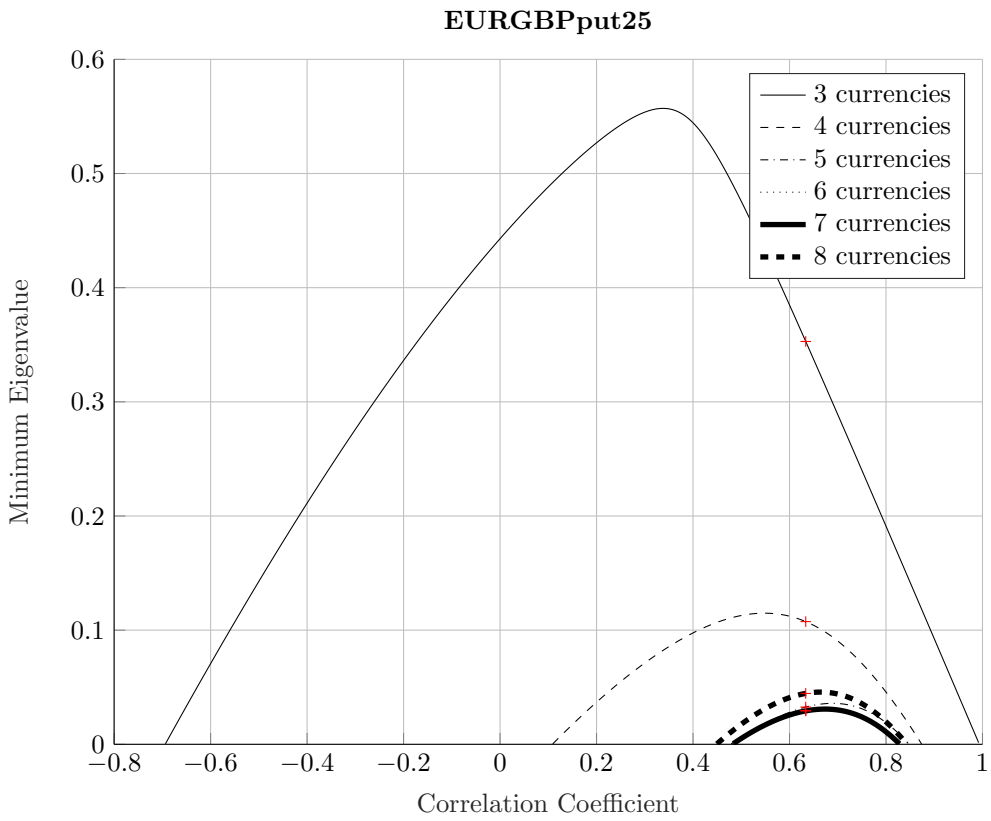


Table 5.3: DM statistic of AUD

			AUD			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	11	0.8182	0.0909	0.0909	18.2294	-30.6205
p25_1M	11	0.9091	0.0909	0.0000	18.6378	-28.5866
atm_1M	11	0.9091	0.0909	0.0000	3.9040	-18.9252
c25_1M	11	0.9091	0.0909	0.0000	23.3820	-53.4192
c10_1M	11	0.9091	0.0909	0.0000	24.4042	-56.4495
p10_1Y	10	0.8000	0.2000	0.0000	33.0361	-19.0071
p25_1Y	10	0.8000	0.2000	0.0000	29.6469	-16.3134
atm_1Y	10	0.9000	0.1000	0.0000	22.1387	-16.9485
c25_1Y	10	0.4000	0.6000	0.0000	21.6595	-9.6022
c10_1Y	10	0.4000	0.6000	0.0000	23.1520	-9.5685
p10_2M	11	0.8182	0.1818	0.0000	12.3093	-34.6361
p25_2M	11	0.9091	0.0000	0.0909	1.7227	-31.5431
atm_2M	11	0.9091	0.0909	0.0000	6.0789	-20.9103
c25_2M	11	1.0000	0.0000	0.0000	-22.0452	-60.1846
c10_2M	11	1.0000	0.0000	0.0000	-21.9920	-63.3258
p10_2Y	10	0.8000	0.2000	0.0000	50.8070	-23.5446
p25_2Y	10	0.8000	0.2000	0.0000	48.4055	-33.4990
atm_2Y	10	0.9000	0.1000	0.0000	38.5949	-23.9586
c25_2Y	10	0.9000	0.1000	0.0000	10.2190	-64.2200
c10_2Y	10	0.9000	0.1000	0.0000	23.5116	-45.6603
p10_3M	10	0.8000	0.2000	0.0000	42.3141	-35.4867
p25_3M	10	0.8000	0.2000	0.0000	32.7108	-36.4696
atm_3M	10	0.9000	0.1000	0.0000	21.3948	-21.6456
c25_3M	10	0.9000	0.1000	0.0000	3.7533	-60.1739
c10_3M	10	0.9000	0.1000	0.0000	10.3187	-66.9004
p10_6M	10	0.8000	0.1000	0.1000	62.2352	-34.4791
p25_6M	10	0.8000	0.2000	0.0000	50.3425	-44.5909
atm_6M	10	0.9000	0.1000	0.0000	23.6354	-22.4302
c25_6M	10	0.9000	0.1000	0.0000	9.6663	-65.1257
c10_6M	10	0.9000	0.1000	0.0000	17.0660	-70.4651

Notes: This table presents the Diebold-Mariano test summary for eigen model and GARCH model performance in European countries. Deltas and Maturities are separated. 'Count' column report amount of corresponding delta/maturity. The 'Eigen' column reports the ratio of eigen model outperforming GARCH model in DM test. The 'GARCH' column reports the ratio of GARCH model outperforming eigen model in DM tests. The 'Similar' column reports the ratio of two models do not present significant difference. The 'Max' column provides max Diebold-Mariano statistic for corresponding strategy. The 'Min' column provides min Diebold-Mariano statistic for corresponding strategy.

Table 5.4: DM statistic of CAD

			CAD			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	10	0.9000	0.1000	0.0000	18.2294	-29.5053
p25_1M	10	0.9000	0.1000	0.0000	18.6378	-28.8176
atm_1M	10	1.0000	0.0000	0.0000	-2.8731	-19.3781
c25_1M	10	0.9000	0.1000	0.0000	23.3820	-69.4086
c10_1M	10	0.9000	0.1000	0.0000	24.4042	-75.2249
p10_1Y	9	0.7778	0.2222	0.0000	9.1251	-34.2080
p25_1Y	9	0.7778	0.1111	0.1111	3.8784	-32.6344
atm_1Y	9	1.0000	0.0000	0.0000	-3.1629	-24.0592
c25_1Y	9	0.0000	0.8889	0.1111	22.7859	-1.1500
c10_1Y	9	0.0000	0.8889	0.1111	31.9683	-0.5686
p10_2M	10	0.9000	0.1000	0.0000	3.1780	-33.3516
p25_2M	10	0.9000	0.0000	0.1000	1.7227	-30.3624
atm_2M	10	1.0000	0.0000	0.0000	-7.1119	-22.2537
c25_2M	10	1.0000	0.0000	0.0000	-29.8788	-71.0102
c10_2M	10	1.0000	0.0000	0.0000	-28.2558	-80.1015
p10_2Y	9	0.8889	0.1111	0.0000	9.9381	-53.7598
p25_2Y	9	0.8889	0.1111	0.0000	6.5557	-42.9838
atm_2Y	9	1.0000	0.0000	0.0000	-4.3726	-22.6098
c25_2Y	9	1.0000	0.0000	0.0000	-23.3123	-107.7967
c10_2Y	9	1.0000	0.0000	0.0000	-14.1693	-135.3757
p10_3M	9	0.8889	0.1111	0.0000	3.6505	-33.0891
p25_3M	9	0.8889	0.1111	0.0000	3.0891	-34.1116
atm_3M	9	1.0000	0.0000	0.0000	-6.2252	-23.2795
c25_3M	9	1.0000	0.0000	0.0000	-30.1983	-73.1755
c10_3M	9	1.0000	0.0000	0.0000	-24.6828	-83.7153
p10_6M	9	0.8889	0.0000	0.1111	0.2449	-36.1528
p25_6M	9	0.8889	0.1111	0.0000	3.0525	-37.9996
atm_6M	9	1.0000	0.0000	0.0000	-4.6976	-24.0055
c25_6M	9	1.0000	0.0000	0.0000	-31.0399	-76.1640
c10_6M	9	1.0000	0.0000	0.0000	-24.1586	-88.4647

Table 5.5: DM statistic of CHF

			CHF			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	15	0.9333	0.0667	0.0000	6.7314	-29.1511
p25_1M	15	0.9333	0.0667	0.0000	4.4995	-21.7177
atm_1M	15	1.0000	0.0000	0.0000	-11.7295	-33.9460
c25_1M	15	1.0000	0.0000	0.0000	-4.1788	-52.4873
c10_1M	15	1.0000	0.0000	0.0000	-4.2040	-44.7699
p10_1Y	14	0.6429	0.2857	0.0714	13.0086	-296.1665
p25_1Y	14	0.8571	0.0714	0.0714	3.9915	-224.7012
atm_1Y	14	1.0000	0.0000	0.0000	-11.1723	-6884.5212
c25_1Y	14	0.5714	0.4286	0.0000	204.6191	-229.7451
c10_1Y	14	0.3571	0.5714	0.0714	190.4506	-318.7011
p10_2M	15	0.9333	0.0667	0.0000	4.7315	-27.7732
p25_2M	15	0.9333	0.0667	0.0000	3.7557	-21.5918
atm_2M	15	1.0000	0.0000	0.0000	-11.2113	-34.4652
c25_2M	15	1.0000	0.0000	0.0000	-14.6171	-53.1467
c10_2M	15	1.0000	0.0000	0.0000	-14.5599	-55.7931
p10_2Y	13	0.5385	0.2308	0.2308	7.0177	-21.5834
p25_2Y	13	0.9231	0.0769	0.0000	2.8713	-25.1596
atm_2Y	13	1.0000	0.0000	0.0000	-9.6281	-24.4899
c25_2Y	13	1.0000	0.0000	0.0000	-14.9733	-83.5110
c10_2Y	13	1.0000	0.0000	0.0000	-7.4305	-85.5667
p10_3M	13	0.9231	0.0769	0.0000	5.0725	-28.4463
p25_3M	13	0.9231	0.0769	0.0000	3.9950	-21.9804
atm_3M	13	1.0000	0.0000	0.0000	-10.7450	-35.2838
c25_3M	13	1.0000	0.0000	0.0000	-15.2078	-46.7379
c10_3M	13	1.0000	0.0000	0.0000	-15.1203	-61.5365
p10_6M	13	0.9231	0.0769	0.0000	4.5352	-274.8602
p25_6M	13	0.9231	0.0769	0.0000	5.1170	-211.5120
atm_6M	13	1.0000	0.0000	0.0000	-11.6008	-3848.0014
c25_6M	13	1.0000	0.0000	0.0000	-15.6066	-202.3363
c10_6M	13	1.0000	0.0000	0.0000	-15.6579	-293.6030

Table 5.6: DM statistic of DKK

			DKK			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	6	0.6667	0.3333	0.0000	6.7314	-29.5053
p25_1M	6	0.6667	0.3333	0.0000	4.4995	-28.8176
atm_1M	6	0.8333	0.1667	0.0000	2.0347	-14.7638
c25_1M	6	1.0000	0.0000	0.0000	-4.1788	-69.4086
c10_1M	6	1.0000	0.0000	0.0000	-4.2040	-75.2249
p10_1Y	5	0.8000	0.2000	0.0000	4.3500	-34.2080
p25_1Y	5	0.6000	0.2000	0.2000	3.9915	-29.2354
atm_1Y	5	0.8000	0.0000	0.2000	-1.7943	-27.1125
c25_1Y	5	0.0000	1.0000	0.0000	226.3842	15.0806
c10_1Y	5	0.0000	1.0000	0.0000	228.6671	18.2777
p10_2M	6	0.8333	0.1667	0.0000	4.7315	-36.4356
p25_2M	6	0.8333	0.1667	0.0000	3.7557	-32.7695
atm_2M	6	0.8333	0.0000	0.1667	1.4991	-16.8310
c25_2M	6	1.0000	0.0000	0.0000	-16.8131	-104.9686
c10_2M	6	1.0000	0.0000	0.0000	-21.6443	-103.0454
p10_2Y	4	0.7500	0.2500	0.0000	6.0771	-53.7598
p25_2Y	4	0.7500	0.2500	0.0000	2.8713	-42.3919
atm_2Y	4	0.7500	0.2500	0.0000	17.6115	-10.7207
c25_2Y	4	1.0000	0.0000	0.0000	-43.7092	-118.6174
c10_2Y	4	1.0000	0.0000	0.0000	-45.6603	-135.3757
p10_3M	5	0.8000	0.2000	0.0000	5.0725	-33.8814
p25_3M	5	0.8000	0.2000	0.0000	3.9950	-30.7564
atm_3M	5	0.8000	0.0000	0.2000	0.3895	-18.9971
c25_3M	5	1.0000	0.0000	0.0000	-24.5892	-108.0858
c10_3M	5	1.0000	0.0000	0.0000	-32.4429	-99.8500
p10_6M	5	0.8000	0.2000	0.0000	4.5352	-32.8712
p25_6M	5	0.8000	0.2000	0.0000	5.1170	-32.1848
atm_6M	5	0.8000	0.0000	0.2000	-1.3130	-23.0889
c25_6M	5	1.0000	0.0000	0.0000	-37.2711	-117.5965
c10_6M	5	1.0000	0.0000	0.0000	-47.1952	-119.7842

Table 5.7: DM statistic of EUR

			EUR			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	14	0.7857	0.2143	0.0000	16.5219	-26.8282
p25_1M	14	0.7857	0.2143	0.0000	12.3349	-31.6937
atm_1M	14	0.9286	0.0714	0.0000	8.9472	-39.3868
c25_1M	14	0.9286	0.0714	0.0000	7.7845	-48.4671
c10_1M	14	0.8571	0.0714	0.0714	10.3045	-43.9540
p10_1Y	13	0.6923	0.3077	0.0000	25.4486	-227.6371
p25_1Y	13	0.8462	0.0769	0.0769	5.0622	-236.1357
atm_1Y	13	1.0000	0.0000	0.0000	-3.1629	-8390.2853
c25_1Y	13	0.5385	0.4615	0.0000	27.6876	-239.4677
c10_1Y	13	0.5385	0.4615	0.0000	21.7169	-285.0797
p10_2M	14	0.7857	0.2143	0.0000	16.5498	-29.5885
p25_2M	14	0.8571	0.1429	0.0000	10.0830	-38.0599
atm_2M	14	0.9286	0.0714	0.0000	3.9961	-43.4816
c25_2M	14	0.9286	0.0714	0.0000	2.9341	-59.7588
c10_2M	14	0.8571	0.0714	0.0714	7.5845	-54.8172
p10_2Y	13	0.5385	0.3846	0.0769	21.4690	-238.3610
p25_2Y	13	0.8462	0.1538	0.0000	10.0274	-227.9619
atm_2Y	13	0.8462	0.0769	0.0769	5.6602	-6680.2448
c25_2Y	13	0.9231	0.0000	0.0769	-1.9151	-137.3422
c10_2Y	13	1.0000	0.0000	0.0000	-4.3108	-191.2037
p10_3M	13	0.6923	0.3077	0.0000	17.3839	-190.7252
p25_3M	13	0.8462	0.0769	0.0769	8.2441	-211.4651
atm_3M	13	0.9231	0.0000	0.0769	1.1868	-2552.0749
c25_3M	13	0.9231	0.0000	0.0769	-0.5854	-210.2047
c10_3M	13	0.9231	0.0769	0.0000	5.2464	-245.7514
p10_6M	13	0.6923	0.3077	0.0000	17.6768	-192.6914
p25_6M	13	0.8462	0.0769	0.0769	8.4480	-223.1722
atm_6M	13	0.9231	0.0769	0.0000	1.9788	-4502.3086
c25_6M	13	0.9231	0.0000	0.0769	-1.4065	-221.9277
c10_6M	13	0.9231	0.0769	0.0000	4.4637	-266.8076

Table 5.8: DM statistic of HKD

			HKD			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	7	0.8571	0.0000	0.1429	-0.3379	-28.9432
p25_1M	7	0.8571	0.1429	0.0000	2.9423	-31.6937
atm_1M	7	0.8571	0.1429	0.0000	6.2366	-18.5061
c25_1M	7	1.0000	0.0000	0.0000	-19.7004	-67.3236
c10_1M	7	1.0000	0.0000	0.0000	-20.5042	-51.6861
p10_1Y	5	1.0000	0.0000	0.0000	-5.1039	-17.6337
p25_1Y	5	1.0000	0.0000	0.0000	-4.2648	-18.4343
atm_1Y	5	0.8000	0.2000	0.0000	27.6925	-26.4096
c25_1Y	5	1.0000	0.0000	0.0000	-9.3385	-49.6491
c10_1Y	5	0.8000	0.0000	0.2000	-1.2495	-29.5102
p10_2M	7	1.0000	0.0000	0.0000	-2.8197	-29.5885
p25_2M	7	0.8571	0.0000	0.1429	0.9947	-38.0599
atm_2M	7	0.8571	0.1429	0.0000	6.2392	-19.8409
c25_2M	7	1.0000	0.0000	0.0000	-18.7184	-55.0104
c10_2M	7	1.0000	0.0000	0.0000	-18.6320	-63.3258
p10_2Y	5	0.8000	0.2000	0.0000	7.0177	-27.2766
p25_2Y	5	1.0000	0.0000	0.0000	-10.4007	-18.5741
atm_2Y	5	0.8000	0.2000	0.0000	24.3935	-21.8503
c25_2Y	5	1.0000	0.0000	0.0000	-12.5702	-83.5110
c10_2Y	5	1.0000	0.0000	0.0000	-13.2483	-38.2942
p10_3M	4	1.0000	0.0000	0.0000	-13.1922	-31.0737
p25_3M	4	1.0000	0.0000	0.0000	-12.5004	-35.1699
atm_3M	4	1.0000	0.0000	0.0000	-6.5738	-25.8603
c25_3M	4	1.0000	0.0000	0.0000	-11.6490	-51.8846
c10_3M	4	1.0000	0.0000	0.0000	-11.1476	-66.9004
p10_6M	4	1.0000	0.0000	0.0000	-16.0991	-34.4791
p25_6M	4	1.0000	0.0000	0.0000	-14.0986	-38.5545
atm_6M	4	1.0000	0.0000	0.0000	-7.5398	-40.2538
c25_6M	4	1.0000	0.0000	0.0000	-14.3382	-65.1257
c10_6M	4	1.0000	0.0000	0.0000	-12.7448	-67.6831

Table 5.9: DM statistic of HUF

	Count	Eigen	HUF GARCH	Similar	Max	Min
p10_1M	2	0.5000	0.5000	0.0000	16.5219	-13.0171
p25_1M	2	0.5000	0.5000	0.0000	5.0221	-12.7034
atm_1M	2	1.0000	0.0000	0.0000	-4.6390	-14.9916
c25_1M	2	1.0000	0.0000	0.0000	-3.6609	-15.1998
c10_1M	2	0.5000	0.0000	0.5000	-0.2666	-14.8773
p10_1Y	2	0.0000	1.0000	0.0000	25.4486	13.0086
p25_1Y	2	1.0000	0.0000	0.0000	-6.6376	-16.1851
atm_1Y	2	1.0000	0.0000	0.0000	-10.0836	-19.1509
c25_1Y	2	1.0000	0.0000	0.0000	-6.3959	-34.8987
c10_1Y	2	0.0000	1.0000	0.0000	19.7571	18.4245
p10_2M	2	0.5000	0.5000	0.0000	15.2330	-13.5269
p25_2M	2	0.5000	0.5000	0.0000	2.6734	-13.0620
atm_2M	2	1.0000	0.0000	0.0000	-7.8526	-15.6161
c25_2M	2	1.0000	0.0000	0.0000	-4.9789	-14.6442
c10_2M	2	0.5000	0.0000	0.5000	-1.5705	-14.5599
p10_2Y	2	0.5000	0.5000	0.0000	21.4690	-12.8732
p25_2Y	2	1.0000	0.0000	0.0000	-2.8934	-15.5541
atm_2Y	2	1.0000	0.0000	0.0000	-12.9173	-24.4899
c25_2Y	2	1.0000	0.0000	0.0000	-16.5330	-31.5384
c10_2Y	2	1.0000	0.0000	0.0000	-7.4305	-21.7613
p10_3M	2	0.5000	0.5000	0.0000	14.9605	-14.9184
p25_3M	2	0.5000	0.0000	0.5000	1.0016	-13.9023
atm_3M	2	1.0000	0.0000	0.0000	-9.9193	-17.1983
c25_3M	2	1.0000	0.0000	0.0000	-7.1963	-15.2078
c10_3M	2	1.0000	0.0000	0.0000	-4.5516	-15.1203
p10_6M	2	0.5000	0.5000	0.0000	15.2650	-15.7124
p25_6M	2	0.5000	0.0000	0.5000	-1.5073	-14.6123
atm_6M	2	1.0000	0.0000	0.0000	-12.4708	-19.8917
c25_6M	2	1.0000	0.0000	0.0000	-13.8001	-15.6892
c10_6M	2	1.0000	0.0000	0.0000	-13.4655	-15.6579

Table 5.10: DM statistic of JPY

			JPY			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	10	1.0000	0.0000	0.0000	-9.4526	-28.9432
p25_1M	10	1.0000	0.0000	0.0000	-9.8135	-27.7085
atm_1M	10	1.0000	0.0000	0.0000	-10.4868	-16.2791
c25_1M	10	1.0000	0.0000	0.0000	-23.4524	-56.2945
c10_1M	10	1.0000	0.0000	0.0000	-32.0813	-54.3505
p10_1Y	8	1.0000	0.0000	0.0000	-5.5215	-33.1486
p25_1Y	8	1.0000	0.0000	0.0000	-8.0022	-39.4705
atm_1Y	8	1.0000	0.0000	0.0000	-11.3983	-24.2970
c25_1Y	8	0.2500	0.7500	0.0000	17.2953	-43.2870
c10_1Y	8	0.2500	0.7500	0.0000	19.4179	-65.4266
p10_2M	10	1.0000	0.0000	0.0000	-11.3351	-28.2756
p25_2M	10	1.0000	0.0000	0.0000	-13.0823	-33.1929
atm_2M	10	1.0000	0.0000	0.0000	-13.4549	-20.0938
c25_2M	10	1.0000	0.0000	0.0000	-29.3425	-59.7588
c10_2M	10	1.0000	0.0000	0.0000	-37.7449	-76.1977
p10_2Y	8	0.8750	0.0000	0.1250	-0.0010	-30.7686
p25_2Y	8	1.0000	0.0000	0.0000	-9.9269	-45.7838
atm_2Y	8	1.0000	0.0000	0.0000	-14.3755	-22.2126
c25_2Y	8	1.0000	0.0000	0.0000	-38.7292	-103.4898
c10_2Y	8	1.0000	0.0000	0.0000	-38.8979	-127.9406
p10_3M	8	1.0000	0.0000	0.0000	-10.9305	-28.3817
p25_3M	8	1.0000	0.0000	0.0000	-13.6435	-22.4003
atm_3M	8	1.0000	0.0000	0.0000	-15.3642	-22.5637
c25_3M	8	1.0000	0.0000	0.0000	-31.0754	-72.6189
c10_3M	8	1.0000	0.0000	0.0000	-45.8785	-82.6625
p10_6M	8	1.0000	0.0000	0.0000	-7.6812	-36.0108
p25_6M	8	1.0000	0.0000	0.0000	-13.8275	-29.2672
atm_6M	8	1.0000	0.0000	0.0000	-17.0100	-24.2879
c25_6M	8	1.0000	0.0000	0.0000	-33.9206	-91.5836
c10_6M	8	1.0000	0.0000	0.0000	-53.6227	-107.4420

Table 5.11: DM statistic of MXN

	Count	Eigen	MXN GARCH	Similar	Max	Min
p10_1M	2	1.0000	0.0000	0.0000	-13.6784	-17.5267
p25_1M	2	1.0000	0.0000	0.0000	-14.3721	-17.5698
atm_1M	2	1.0000	0.0000	0.0000	-9.5718	-17.8434
c25_1M	2	1.0000	0.0000	0.0000	-27.1090	-52.4873
c10_1M	2	1.0000	0.0000	0.0000	-26.0623	-35.7629
p10_1Y	2	1.0000	0.0000	0.0000	-14.5573	-22.4152
p25_1Y	2	1.0000	0.0000	0.0000	-16.4942	-26.0675
atm_1Y	2	1.0000	0.0000	0.0000	-10.1562	-21.6513
c25_1Y	2	1.0000	0.0000	0.0000	-29.8439	-31.4369
c10_1Y	2	1.0000	0.0000	0.0000	-29.3032	-33.1386
p10_2M	2	1.0000	0.0000	0.0000	-13.3032	-18.0575
p25_2M	2	1.0000	0.0000	0.0000	-13.9831	-18.8394
atm_2M	2	1.0000	0.0000	0.0000	-12.0346	-19.7571
c25_2M	2	1.0000	0.0000	0.0000	-27.2047	-53.1467
c10_2M	2	1.0000	0.0000	0.0000	-26.1917	-33.6412
p10_2Y	2	1.0000	0.0000	0.0000	-16.3072	-26.6507
p25_2Y	2	1.0000	0.0000	0.0000	-19.6141	-28.4553
atm_2Y	2	1.0000	0.0000	0.0000	-13.5995	-17.8957
c25_2Y	2	1.0000	0.0000	0.0000	-29.4403	-30.0264
c10_2Y	2	1.0000	0.0000	0.0000	-28.9665	-32.7945
p10_3M	2	1.0000	0.0000	0.0000	-13.8239	-17.5336
p25_3M	2	1.0000	0.0000	0.0000	-14.4460	-18.2663
atm_3M	2	1.0000	0.0000	0.0000	-10.7877	-20.9295
c25_3M	2	1.0000	0.0000	0.0000	-26.9537	-46.7379
c10_3M	2	1.0000	0.0000	0.0000	-26.9691	-31.4229
p10_6M	2	1.0000	0.0000	0.0000	-14.6735	-18.1488
p25_6M	2	1.0000	0.0000	0.0000	-15.7407	-19.9520
atm_6M	2	1.0000	0.0000	0.0000	-10.2745	-25.7358
c25_6M	2	1.0000	0.0000	0.0000	-28.5884	-32.5174
c10_6M	2	1.0000	0.0000	0.0000	-28.3694	-31.0709

Table 5.12: DM statistic of NOK

			NOK			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	9	1.0000	0.0000	0.0000	-3.4071	-26.9926
p25_1M	9	1.0000	0.0000	0.0000	-12.0444	-26.3993
atm_1M	9	1.0000	0.0000	0.0000	-10.1365	-20.0418
c25_1M	9	1.0000	0.0000	0.0000	-23.8210	-50.9365
c10_1M	9	1.0000	0.0000	0.0000	-23.2784	-53.7121
p10_1Y	8	0.5000	0.3750	0.1250	5.0604	-18.8637
p25_1Y	8	0.6250	0.0000	0.3750	0.7239	-20.2117
atm_1Y	8	1.0000	0.0000	0.0000	-4.5367	-27.1125
c25_1Y	8	0.0000	1.0000	0.0000	129.1794	12.9341
c10_1Y	8	0.0000	1.0000	0.0000	127.9944	9.6072
p10_2M	9	1.0000	0.0000	0.0000	-8.9402	-36.4356
p25_2M	9	1.0000	0.0000	0.0000	-14.4256	-32.7695
atm_2M	9	1.0000	0.0000	0.0000	-11.5879	-25.2440
c25_2M	9	1.0000	0.0000	0.0000	-31.7832	-71.6719
c10_2M	9	1.0000	0.0000	0.0000	-29.6320	-72.6100
p10_2Y	7	0.5714	0.1429	0.2857	4.8741	-30.7686
p25_2Y	7	1.0000	0.0000	0.0000	-12.7793	-40.6065
atm_2Y	7	1.0000	0.0000	0.0000	-11.6764	-37.9599
c25_2Y	7	1.0000	0.0000	0.0000	-29.2050	-91.5547
c10_2Y	7	1.0000	0.0000	0.0000	-26.6438	-99.6952
p10_3M	8	1.0000	0.0000	0.0000	-11.7346	-33.8814
p25_3M	8	1.0000	0.0000	0.0000	-14.8797	-32.7944
atm_3M	8	1.0000	0.0000	0.0000	-11.9942	-27.9034
c25_3M	8	1.0000	0.0000	0.0000	-31.2752	-69.2927
c10_3M	8	1.0000	0.0000	0.0000	-28.9112	-81.4955
p10_6M	8	1.0000	0.0000	0.0000	-7.1495	-32.1492
p25_6M	8	1.0000	0.0000	0.0000	-14.9140	-35.4839
atm_6M	8	1.0000	0.0000	0.0000	-11.7536	-31.6372
c25_6M	8	1.0000	0.0000	0.0000	-30.9521	-70.3608
c10_6M	8	1.0000	0.0000	0.0000	-28.2652	-102.3565

Table 5.13: DM statistic of NZD

			NZD			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	10	0.8000	0.0000	0.2000	1.1893	-25.4062
p25_1M	10	0.9000	0.1000	0.0000	2.9423	-24.4555
atm_1M	10	0.8000	0.2000	0.0000	6.2366	-14.0262
c25_1M	10	1.0000	0.0000	0.0000	-19.6580	-50.9812
c10_1M	10	1.0000	0.0000	0.0000	-20.5042	-48.4528
p10_1Y	9	0.7778	0.2222	0.0000	33.0361	-15.9456
p25_1Y	9	0.7778	0.1111	0.1111	29.6469	-21.8507
atm_1Y	9	0.7778	0.2222	0.0000	27.6925	-18.8978
c25_1Y	9	0.2222	0.7778	0.0000	22.7859	-24.5429
c10_1Y	9	0.3333	0.6667	0.0000	25.2141	-34.0439
p10_2M	10	0.8000	0.1000	0.1000	12.3093	-33.3516
p25_2M	10	0.9000	0.0000	0.1000	0.9947	-30.3624
atm_2M	10	0.8000	0.2000	0.0000	6.2392	-17.2506
c25_2M	10	1.0000	0.0000	0.0000	-23.2996	-59.0178
c10_2M	10	1.0000	0.0000	0.0000	-3.3392	-51.5189
p10_2Y	9	0.7778	0.2222	0.0000	50.8070	-29.6234
p25_2Y	9	0.8889	0.1111	0.0000	48.4055	-42.9838
atm_2Y	9	0.7778	0.2222	0.0000	38.5949	-19.0021
c25_2Y	9	0.8889	0.1111	0.0000	10.2190	-67.7749
c10_2Y	9	0.8889	0.1111	0.0000	23.5116	-101.8494
p10_3M	9	0.7778	0.2222	0.0000	42.3141	-33.0891
p25_3M	9	0.8889	0.1111	0.0000	32.7108	-31.6674
atm_3M	9	0.8889	0.1111	0.0000	21.3948	-19.3330
c25_3M	9	0.8889	0.1111	0.0000	3.7533	-58.4228
c10_3M	9	0.8889	0.1111	0.0000	10.3187	-60.5544
p10_6M	9	0.7778	0.2222	0.0000	62.2352	-36.2897
p25_6M	9	0.8889	0.1111	0.0000	50.3425	-37.9452
atm_6M	9	0.8889	0.1111	0.0000	23.6354	-19.7670
c25_6M	9	0.8889	0.1111	0.0000	9.6663	-69.0351
c10_6M	9	0.8889	0.1111	0.0000	17.0660	-68.5679

Table 5.14: DM statistic of PLN

			PLN			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	2	0.5000	0.5000	0.0000	15.7237	-13.4000
p25_1M	2	0.5000	0.5000	0.0000	12.3349	-13.2524
atm_1M	2	0.5000	0.5000	0.0000	8.9472	-15.3271
c25_1M	2	0.5000	0.5000	0.0000	7.7845	-14.0573
c10_1M	2	0.5000	0.5000	0.0000	10.3045	-15.1918
p10_1Y	2	0.5000	0.5000	0.0000	13.6135	-19.4848
p25_1Y	2	1.0000	0.0000	0.0000	-5.5587	-5.5949
atm_1Y	2	1.0000	0.0000	0.0000	-10.5875	-16.3335
c25_1Y	2	1.0000	0.0000	0.0000	-3.9474	-31.6957
c10_1Y	2	0.5000	0.0000	0.5000	-0.4845	-19.1670
p10_2M	2	0.5000	0.5000	0.0000	16.5498	-13.7867
p25_2M	2	0.5000	0.5000	0.0000	10.0830	-13.6357
atm_2M	2	0.5000	0.5000	0.0000	3.9961	-16.0216
c25_2M	2	0.5000	0.5000	0.0000	2.9341	-14.6171
c10_2M	2	0.5000	0.5000	0.0000	7.5845	-16.6015
p10_2Y	2	0.5000	0.5000	0.0000	15.7199	-15.8174
p25_2Y	2	0.5000	0.5000	0.0000	10.0274	-15.6955
atm_2Y	2	0.5000	0.5000	0.0000	5.6602	-19.8080
c25_2Y	2	0.5000	0.0000	0.5000	-1.9151	-14.9733
c10_2Y	2	1.0000	0.0000	0.0000	-4.3108	-11.9215
p10_3M	2	0.5000	0.5000	0.0000	17.3839	-14.4341
p25_3M	2	0.5000	0.5000	0.0000	8.2441	-14.0630
atm_3M	2	0.5000	0.0000	0.5000	1.1868	-17.0884
c25_3M	2	0.5000	0.0000	0.5000	-0.5854	-15.4240
c10_3M	2	0.5000	0.5000	0.0000	5.2464	-18.6088
p10_6M	2	0.5000	0.5000	0.0000	17.6768	-15.3223
p25_6M	2	0.5000	0.5000	0.0000	8.4480	-14.7586
atm_6M	2	0.5000	0.5000	0.0000	1.9788	-18.4838
c25_6M	2	0.5000	0.0000	0.5000	-1.4065	-15.6066
c10_6M	2	0.5000	0.5000	0.0000	4.4637	-17.7700

Table 5.15: DM statistic of SEK

			SEK			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	9	0.7778	0.2222	0.0000	8.3220	-30.6205
p25_1M	9	0.7778	0.2222	0.0000	3.7249	-28.5866
atm_1M	9	0.8889	0.1111	0.0000	2.0347	-18.2102
c25_1M	9	1.0000	0.0000	0.0000	-8.1051	-53.4192
c10_1M	9	1.0000	0.0000	0.0000	-7.9267	-54.3505
p10_1Y	8	0.6250	0.3750	0.0000	13.2385	-19.0071
p25_1Y	8	0.6250	0.1250	0.2500	5.0622	-21.8497
atm_1Y	8	0.8750	0.0000	0.1250	-1.7943	-27.7167
c25_1Y	8	0.0000	1.0000	0.0000	226.3842	8.4763
c10_1Y	8	0.0000	1.0000	0.0000	228.6671	9.3257
p10_2M	9	0.8889	0.1111	0.0000	3.6946	-34.6361
p25_2M	9	1.0000	0.0000	0.0000	-12.6023	-32.5446
atm_2M	9	0.8889	0.0000	0.1111	1.4991	-21.2803
c25_2M	9	1.0000	0.0000	0.0000	-22.4779	-104.9686
c10_2M	9	1.0000	0.0000	0.0000	-22.0311	-103.0454
p10_2Y	8	0.6250	0.2500	0.1250	13.4423	-25.1645
p25_2Y	8	0.8750	0.1250	0.0000	4.3776	-40.6065
atm_2Y	8	0.7500	0.1250	0.1250	17.6115	-24.3830
c25_2Y	8	1.0000	0.0000	0.0000	-27.6203	-118.6174
c10_2Y	8	1.0000	0.0000	0.0000	-27.3842	-127.9406
p10_3M	8	0.8750	0.1250	0.0000	8.0535	-35.4867
p25_3M	8	1.0000	0.0000	0.0000	-12.7961	-36.4696
atm_3M	8	0.8750	0.0000	0.1250	0.3895	-25.0295
c25_3M	8	1.0000	0.0000	0.0000	-22.2276	-108.0858
c10_3M	8	1.0000	0.0000	0.0000	-22.8286	-99.8500
p10_6M	8	0.8750	0.1250	0.0000	10.5291	-32.2403
p25_6M	8	1.0000	0.0000	0.0000	-12.7382	-44.5909
atm_6M	8	0.8750	0.0000	0.1250	-1.3130	-28.1438
c25_6M	8	1.0000	0.0000	0.0000	-23.0572	-117.5965
c10_6M	8	1.0000	0.0000	0.0000	-26.2164	-119.7842

Table 5.16: DM statistic of SGD

			SGD			
	Count	Eigen	GARCH	Similar	Max	Min
p10_1M	8	1.0000	0.0000	0.0000	-10.2778	-21.3421
p25_1M	8	1.0000	0.0000	0.0000	-12.8229	-31.5667
atm_1M	8	1.0000	0.0000	0.0000	-9.6956	-18.5061
c25_1M	8	1.0000	0.0000	0.0000	-13.4663	-67.3236
c10_1M	8	1.0000	0.0000	0.0000	-13.5583	-49.9229
p10_1Y	7	1.0000	0.0000	0.0000	-5.1039	-33.2334
p25_1Y	7	1.0000	0.0000	0.0000	-10.3568	-39.4705
atm_1Y	7	1.0000	0.0000	0.0000	-13.2807	-26.4096
c25_1Y	7	0.5714	0.2857	0.1429	8.4109	-43.2870
c10_1Y	7	0.2857	0.4286	0.2857	9.4285	-65.4266
p10_2M	8	0.8750	0.0000	0.1250	-1.3621	-29.1019
p25_2M	8	1.0000	0.0000	0.0000	-11.9645	-31.4412
atm_2M	8	1.0000	0.0000	0.0000	-12.9880	-21.2558
c25_2M	8	1.0000	0.0000	0.0000	-15.1804	-55.8784
c10_2M	8	1.0000	0.0000	0.0000	-3.3392	-61.3744
p10_2Y	7	1.0000	0.0000	0.0000	-4.2243	-32.6278
p25_2Y	7	1.0000	0.0000	0.0000	-13.7023	-42.4682
atm_2Y	7	1.0000	0.0000	0.0000	-10.4288	-21.9973
c25_2Y	7	1.0000	0.0000	0.0000	-12.5702	-103.4898
c10_2Y	7	1.0000	0.0000	0.0000	-13.2483	-74.8874
p10_3M	7	1.0000	0.0000	0.0000	-11.7814	-30.9951
p25_3M	7	1.0000	0.0000	0.0000	-11.7919	-32.1336
atm_3M	7	1.0000	0.0000	0.0000	-14.4269	-25.8603
c25_3M	7	1.0000	0.0000	0.0000	-11.6490	-63.6259
c10_3M	7	1.0000	0.0000	0.0000	-11.1476	-64.4302
p10_6M	7	1.0000	0.0000	0.0000	-11.4564	-36.2897
p25_6M	7	1.0000	0.0000	0.0000	-13.1273	-37.9996
atm_6M	7	1.0000	0.0000	0.0000	-16.4002	-40.2538
c25_6M	7	1.0000	0.0000	0.0000	-14.3382	-72.4900
c10_6M	7	1.0000	0.0000	0.0000	-12.7448	-68.3568

Table 5.17: DM statistic of ZAR

	Count	Eigen	ZAR GARCH	Similar	Max	Min
p10_1M	2	1.0000	0.0000	0.0000	-23.0150	-29.1511
p25_1M	2	1.0000	0.0000	0.0000	-21.7177	-22.5944
atm_1M	2	1.0000	0.0000	0.0000	-33.9460	-39.3868
c25_1M	2	1.0000	0.0000	0.0000	-15.8693	-16.3059
c10_1M	2	1.0000	0.0000	0.0000	-16.7566	-18.0977
p10_1Y	2	1.0000	0.0000	0.0000	-227.6371	-296.1665
p25_1Y	2	1.0000	0.0000	0.0000	-224.7012	-236.1357
atm_1Y	2	1.0000	0.0000	0.0000	-6884.5212	-8390.2853
c25_1Y	2	1.0000	0.0000	0.0000	-229.7451	-239.4677
c10_1Y	2	1.0000	0.0000	0.0000	-285.0797	-318.7011
p10_2M	2	1.0000	0.0000	0.0000	-24.1169	-27.7732
p25_2M	2	1.0000	0.0000	0.0000	-21.5918	-23.5674
atm_2M	2	1.0000	0.0000	0.0000	-34.4652	-43.4816
c25_2M	2	1.0000	0.0000	0.0000	-16.0948	-16.9935
c10_2M	2	1.0000	0.0000	0.0000	-17.6566	-18.0330
p10_2Y	1	1.0000	0.0000	0.0000	-238.3610	-238.3610
p25_2Y	1	1.0000	0.0000	0.0000	-227.9619	-227.9619
atm_2Y	1	1.0000	0.0000	0.0000	-6680.2448	-6680.2448
c25_2Y	1	1.0000	0.0000	0.0000	-137.3422	-137.3422
c10_2Y	1	1.0000	0.0000	0.0000	-191.2037	-191.2037
p10_3M	2	1.0000	0.0000	0.0000	-28.4463	-190.7252
p25_3M	2	1.0000	0.0000	0.0000	-21.9804	-211.4651
atm_3M	2	1.0000	0.0000	0.0000	-35.2838	-2552.0749
c25_3M	2	1.0000	0.0000	0.0000	-16.3438	-210.2047
c10_3M	2	1.0000	0.0000	0.0000	-18.2423	-245.7514
p10_6M	2	1.0000	0.0000	0.0000	-192.6914	-274.8602
p25_6M	2	1.0000	0.0000	0.0000	-211.5120	-223.1722
atm_6M	2	1.0000	0.0000	0.0000	-3848.0014	-4502.3086
c25_6M	2	1.0000	0.0000	0.0000	-202.3363	-221.9277
c10_6M	2	1.0000	0.0000	0.0000	-266.8076	-293.6030

Appendix C

Listing 5.1: Delta Calculator

```
function Delta =
    DeltaCalculator(Volatility,SpotPrice,StrikePrice,RateD,RateF,...
TsTd,ToTe,DeltaType,Type,pipsPer)


if pipsPer
    d1 = (log(SpotPrice./StrikePrice)+(RateD-RateF).*TsTd+...
        0.5.*(Volatility.^2).*ToTe)./(Volatility.*sqrt(ToTe));
    if DeltaType
        Delta = Type.*normcdf(Type.*d1,0,1).*exp(-RateF.*TsTd);
    else
        Delta = Type.*normcdf(Type.*d1,0,1);
    end
else
    ForwardPrice = SpotPrice.*exp((RateD-RateF).*TsTd);
    d2 = (log(SpotPrice./StrikePrice)+(RateD-RateF).*TsTd-...
        0.5.*(Volatility.^2).*ToTe)./(Volatility.*sqrt(ToTe));
    if DeltaType
        Delta =
            Type.*normcdf(Type.*d2,0,1).*exp(-RateD.*TsTd).*(StrikePrice./SpotPrice);
```

```
    else
        Delta = Type.*normcdf(Type.*d2,0,1).*(StrikePrice./ForwardPrice);
    end
end
```

Where Volatility is implied volatility, SpotPrice/Strike Price is the market quote and corresponding strike. RateD/RateF indicate annual domestic/foreign interest rate. DeltaType should be set to 1 for spot or 0 for forward respectively. PipsPer should beset to 1 for pips or 0 for percentage. Type is 1 for call option and -1 for put option. With given input argument, the output provide corresponding Delta for certain strike price.

Listing 5.2: Black Scholes Option Price Calculator

```
function V=FX_Option_Price_BS(vol,S,K,r,q,t,te,ts,td,type)
T1 = (1 + te - ts)./365;
T2 = (1 + td - t)./365;
T3 = (1 + td - t)./365;
r = r.*365;
q = q.*365;
d1 = (log(S./K) + (r - q).*(T1) + 0.5.*vol.^2.*(T2))./(vol.*sqrt(T3));
d2 = (log(S./K) + (r - q).*(T1) - 0.5.*vol.^2.*(T2))./(vol.*sqrt(T3));

N1 = normcdf(type.*d1);
N2 = normcdf(type.*d2);

PVS = S.*exp(-q.*T1);
PVK = K.*exp(-r.*T1);
```

```
P1 = type.*PVS;
P2 = type.*PVK;
V = P1.*N1 - P2.*N2;
```

Where S is spot rate, K is strike price, r is domestic deposit rate meanwhile q is overseas deposit rate. ts is the time to spot, te is the time to expiry. td is the time to deposit and V is the value of the options. In this piece of code, I provide the formula of typical Black Scholes model for FX option.

Listing 5.3: Affine Jump Diffusion Option Price

```
function [C,P,vol_C,vol_P]=AffineJumpDiffusionOptionPrice(theta,rT,qT,S,K,T)
K = reshape(K,[],1);
M = length(K);
C = zeros(M,1);
%X = 1i.*((exp(linspace(0,6,10000))-1)+realmin)';
EP = 3000;% end point
NP = 2000;% Number of points this should normally be higher than EP so we
    have the fractionals
X = linspace(sqrt(eps),EP,NP);
c = X.*1i;
%F1 = AJ_CF1(theta,S,r,q,T,t);
F1 = JumpCGF1(S,theta,T);
MGF_N1 = Jump_CGF(c+1,theta,S,T);%AJ_CF(theta,X+1,S,r,q,T,t);
MGF_N2 = Jump_CGF(c,theta,S,T);%AJ_CF(theta,X,S,r,q,T,t);
Dr = exp(-rT.*T);

for i=1:M
```

```
    %Y1 = AJ_N1(F1,MGF_N1,X,K(i));
    Y1 = real(((K(i).^(-c)).*MGF_N1)./(c.*F1));
    Y2 = real(((K(i).^(-c)).*MGF_N2)./(c));
    N1 = cumtrapz(X,Y1);
    N2 = cumtrapz(X,Y2);
    P1=(0.5 + (1./pi).*N1);
    P2=(0.5 + (1./pi).*N2);
    %average the end points
    P1 = nanmedian(P1(NP-30:NP));
    P2 = nanmedian(P2(NP-30:NP));
    C(i,1) = Dr.*(F1.*P1 - K(i).*P2); % call price
end
%C = smooth(C,20);% add a five point smoother to the price.
P = C - S.*exp(-qT.*T) + K.*Dr; % put price

r = rT./365;
q = qT./365;
T = T.*365;
t = 1;

C(C<0) = eps;
P(P<0) = eps;
[vol_C]=optionImpliedVolatilityFX(repmat(S,M,1),K,C,repmat(t,M,1),...
    repmat(t,M,1),repmat(T,M,1),repmat(T,M,1),repmat(r,M,1),...
    repmat(q,M,1),ones(M,1));
[vol_P]=optionImpliedVolatilityFX(repmat(S,M,1),K,P,repmat(t,M,1),...
    repmat(t,M,1),repmat(T,M,1),repmat(T,M,1),repmat(r,M,1),...
    repmat(q,M,1),-1.*ones(M,1));
```

```
end
```

Theta is a vector of parameters of the form theta = [thetaV thetaJ thetar thetaq]; rT and qT are the discount rate from t to T; should be in terms of log(1 + Yield), where Yield is a discrete APR. for domestic and overseas S is the spot exchange rate K is a vector of strike prices T is a tenor in fractions of a year. This function calculates the option price for the affine-jump model

Listing 5.4: Get 1-vol Butterfly

```
function
    [sigmaBF]=solveOneVolButterfly(Volatility_C,Volatility_P,pivotDelta,SpotPrice,...
    StrikePrice,RateD,RateF,TsTd,ToTe,DeltaType,pipsPer,r,q,t,T)
flag = 1;
k = 1;
lb = 0.02;
ub = 1;
sigmaBF = NaN;
while flag == 1
    sigmaIn = linspace(lb,ub,10);
    err =
        oneVolButterflyObjective(sigmaIn,Volatility_C,Volatility_P,pivotDelta,...
    SpotPrice,StrikePrice,RateD,RateF,TsTd,ToTe,DeltaType,pipsPer,r,q,t,T);
    [Il,Iu,exact]=zeroCrossing(err);
    gapErr = abs(sum(err([Il Iu]))./2);
    k = k+1;
    if exact
        sigmaBF = sigmaIn(Il);
```

```
        flag = 0;
    end
    if gapErr(1) >= 1e-6
        lb = sigmaIn(Il);
        ub = sigmaIn(Iu);
    end
    if gapErr(1) < 1e-6
        sigmaBF = 0.5.*(sigmaIn(Il) + sigmaIn(Iu));
        flag = 0;
    end
    if k>20
        sigmaBF = 0.5.*(sigmaIn(Il) + sigmaIn(Iu));
        flag = 0;
    end
end
```

This function aim to transfer 2-vol-butterfly to 1-vol-butterfly. Input arguments are listed as follows: TsTd/ToTe- spot to delivery/today to expiry as fraction of year. t/T- number of days to maturity; RateD/RateF- annual domestic/foreign interest rate; r/q- daily domestic/foreign interest rate; pipsPer- 'pips' or 'per'; DeltaType- 'spot' or 'forward'; pivotDelta- 0.1 or 0.25 for BF10 or BF25.

Listing 5.5: Estimate Implied correlation Matrix

```
function
    SinglePair(pair,ccyList,maturity,matPath,scList,alist,opath,deli,method)
leg = 'USD';
%remove aim pair and add them into head and tail of list
```

```
ccy1 = pair(1:3);
ccy2 = pair(4:6);
ind1 = find(ismember(ccyList,ccy1)==1);
ind2 = find(ismember(ccyList,ccy2)==1);
ccyList([ind1,ind2])=[];
ccyList = [ccy1;ccyList;ccy2];
%get the correct name of currency pair by list
pairList = cell(length(ccyList),1);
for k = 1:length(ccyList)
    ccyU = ccyList{k};
    if sum(ismember(scList,ccyU))
        pairList{k,1} = [ccyU,leg];
    else
        pairList{k,1} = [leg,ccyU];
    end
end
%get the str for aim correlation matrix;
strMatrix = list2matrix(pairList,maturity,alist);
%corMatrix could be empty
[corMatrix,A,data1,data2,data3] = str2data(strMatrix,maturity,matPath,deli);
% load('corMatrix.mat')
%get calculation result for the eigenvalue
if ~isempty(corMatrix)
    [rlb,rmax,rub,rmkt,eigenResult,eigCorData] = cor2vol(corMatrix,method);
    geteigCorDataFile(eigenResult,eigCorData,pair,length(pairList));
%     aimpair = strMatrix{1,length(strMatrix)}(end);
%
    cor2iv(rlb,rmax,rub,rmkt,data1,data2,data3,A,aimpair,maturity,opath,eigenResult);
```

```
end
```

This function aim to organize implied correlation matrix from implied volatility. Input argurments are listed as follows: 'ccyList' is a matrix in str format, which contains the currency pairs' name in the leg journey. 'matPath' and 'opath' are folder path for market data and output data respectively. 'method' is 'end of day' or 'intraday' for the calculation of end of day and high frequency intraday respectively.

Listing 5.6: Get Missing Correlation Bound

```
function [Rlb,Rub,rlb,rub,rmax,Emax,eigenCor]=findCorrelationBounds(R,tin)
ind = find(isnan(R));
R(isinf(R)) = 1;
rvec = linspace(-1,1,ceil(1./tin));
% first identify feasibility, can we find a correlation that returns a PD
% matrix, find candidate bounds
Ri = R;
PD = NaN(length(rvec),1);
for i=1:length(rvec)
    Ri(ind) = rvec(i);
    x = eig(Ri);
    ii = find(x<0);
    if isempty(ii)
        PD(i) = 1;
    end
    lbnd(i) = min(x);
    condition(i,1) = cond(Ri);
    lambda(i,1) = max(eig(Ri));
```

```
end
II = find(lbnd>0);
[Emax,mind] = max(lbnd);
eigenCor = [rvec',lbnd'];
%lbnd is the vector record minimum eigen value, find the largest of it.
rmax = rvec(mind);
if ~isempty(II)
    rlb = rvec(II(1));
    rub = rvec(II(end));

    Rlb(ind) = rlb;
    Rub(ind) = rub;
    [~,ind_rho] = min(condition(II));
    try
        rlb = rvec(1,ind_rho-50);
    catch
        rlb = rvec(1,1);
    end
    try
        rub = rvec(1,ind_rho+50);
    catch
        rub = rvec(1,end);
    end
    rmax = (rlb+rub)/2;
    [~,ind_rho] = max(lbnd);
    rho_max = rvec(1,ind_rho);
else
    [~,ind_rho] = max(lbnd);
```

```
    rho_max = rvec(1,ind_rho);
end
Rlb = R;
Rub = R;
Rmax = R;
Rmax(ind) = rmax;
```

This function provide the algorithm to check the eigenvalue structure. With known correlation matrix R, the aim missing correlation coefficient is in the top right corner (the last column of first row and last row of first column). Keep fill the missing element in matrix with values range from -1 to 1 and check the matrix's minimum eigenvalue, two kinds of the data are recorded: 1)the implied correlation coefficient make minimum eigenvalue equal to zero, which provide bounds for missing element. 2) the coefficient make minimum eigenvalue maximized, which indicate the most stable market and is assumed to be the point forecast of our model.

Listing 5.7: Regression Tree

```
function [trainedModel, validationRMSE] = trainRegressionModel(trainingData)
inputTable = array2table(trainingData, 'VariableNames', {'column_1',
    'column_2', 'column_3'});
predictorNames = {'column_2', 'column_3'};
predictors = inputTable(:, predictorNames);
response = inputTable.column_1;
isCategoricalPredictor = [false, false];
% Train a regression model
% This code specifies all the model options and trains the model.
regressionTree = fitrtree(...
```

```
    predictors, ...
    response, ...
    'MinLeafSize', 4, ...
    'Surrogate', 'off');
% Create the result struct with predict function
predictorExtractionFcn = @(x) array2table(x, 'VariableNames', predictorNames);
treePredictFcn = @(x) predict(regressionTree, x);
trainedModel.predictFcn = @(x) treePredictFcn(predictorExtractionFcn(x));
% Add additional fields to the result struct
trainedModel.RegressionTree = regressionTree;
% Extract predictors and response
% This code processes the data into the right shape for training the
% model.
% Convert input to table
inputTable = array2table(trainingData, 'VariableNames', {'column_1',
    'column_2', 'column_3'});
predictorNames = {'column_2', 'column_3'};
predictors = inputTable(:, predictorNames);
response = inputTable.column_1;
isCategoricalPredictor = [false, false];
% Perform cross-validation
partitionedModel = crossval(trainedModel.RegressionTree, 'KFold', 5);
% Compute validation predictions
validationPredictions = kfoldPredict(partitionedModel);
% Compute validation RMSE
validationRMSE = sqrt(kfoldLoss(partitionedModel, 'LossFun', 'mse'));
```

This script returns a trained regression model and its' RMSE. The model provide a

well-trained regression tree to review and predict with new data. For input argument, trainingData is the training data set. For output, trainedModel is a struct containing the trained regression model. The struct contains multiple fields with information about the model; trainedModel.predictFcn is the function to make predictions on new data; validationRMSE is a double variable containing RMSE.

Listing 5.8: R_{oos} Calculator

```
function data = analysisSingleMaturity(data,nameFile)
lengthInd = size(data,1);
for k =1:lengthInd
    dataU = data(k,:);
    market = dataU{1,3};
    cal = dataU{1,2};
    for kk =2:6
        marketU = market(:,kk);
        calU = cal(:,kk);
        windowsInd = floor(length(calU)./3);
        mktWindows = marketU(end-windowsInd:end,1);
        calWindows = calU(end-windowsInd:end,1);

        rUp = sum((calWindows-mktWindows).^2);
        rBot = sum(mktWindows.^2);
        rOos(1,kk-1) = 1-rUp./rBot;
    end
    data{k,4} = rOos;
end
for k =2:lengthInd
```

```
    oosAll = data{1,4};
    oosU = data{k,4};
    ooUdiff = abs(oosU-oosAll);
    data{k,5} = ooUdiff;
    data{k,6} = sum(ooUdiff);
end
```

This script provides the methodology to value R^2_{oos} measurement(For a single maturity from 1 month to 2 year). For input argument, data contains the out-of-sample for both market quote and calculation result, all five key deltas are recorded(put delta10, put delta25, at-the-money, call delta 25, call delta10). For output, data contains the out-of-sample R^2_{oos} for all deltas and its change by making a selected predictors equal to zero.

Listing 5.9: Equity Eigen Structure Main

```
load('dateRange.mat')
pth = 'D:\ChapterThree\AllMat\';
cd(pth);
stockList = dir();
stockList(1:2) = [];
stockName = extractfield(stockList,'name');
DataMatrix = [];
n=1;
for k =1:length(stockName)
    load([pth,stockName{k}]);
    dateU = floor(data(:,1));
    [after,ind1,ind2]= intersect(dateU,dateTested);
    if ~isempty(after)
```

```
        DataMatrix(:,n) = interpStock(dateTested,data);
        NameMatrix(:,n) = cellstr(stockName{k});
        n=n+1;
    end
 end
% %
 Data.stockData = DataMatrix;
 Data.name = NameMatrix;
 Data.Date = dateTested;
 save('stockInfo.mat','Data');
load('stockInfo.mat');
DataMatrix = Data.stockData;
dateTested = Data.Date;
weekendInfo = sum(Data.stockData,2);
ind = find(weekendInfo==0);
DataMatrix(ind,:) =[];
dateTested(ind) = [];

singleInfo = zeros(size(DataMatrix,2),1);
for k =1:size(DataMatrix,2)
    singleInfo(k) = length(find(DataMatrix(:,k)~=0));
end

[~,ind] = sort(singleInfo,'descend');
DataMatrixSorted = DataMatrix(:,ind);
stockSorted = Data.name(ind);
lambda = [];
n=1;
```

```
for k =20:size(DataMatrixSorted,2)
    matrixU = DataMatrixSorted(end-100:end,1:k);
    zeroInd = sum(matrixU);
    if isempty(find(zeroInd==0))
        lambda(n) = singleMatrixTest(matrixU);
        n=n+1;
    end
end
```

This script provides the main function for covariance matrix of equities. There is a folder contains tick data for all equities' in S&P500 index from 1996 to 2020. In this function, the final outcome is a structure with interpreted matrix with equity names and date cover. All equities are sorted by trading volume, the data clean process is also presented in this script.

Listing 5.10: Equity Eigen Structure Simulation

```
load('TimeSeriesEigen.mat');
data = Data.weekly{1,3};
R = diff(data);
S = cov(R);
Shat = nearestSPD(S);
Q = chol(Shat);
for i = 1:10000
    tic
    U = randn(size(R));
    Rstar = U*Q;
    n=1;
```

```
    for k =20:size(Rstar,2)
        matrixU = Rstar(:,1:k);
        zeroInd = sum(matrixU);
        if isempty(find(zeroInd==0))
            [lambda{n,1},lambda{n,3},lambda{n,2}] = singleMatrixTest(matrixU);
            n=n+1;
        end
    end
    test{1,1} = 735972;
    test{1,2} = lambda;
    test{1,3} = matrixU;
    simC(:,i)= simulationTurn(test);
    toc
end
```

This script provides the simulation testbench for covariance matrix of equities. The original market data is used as seed for the simulation, the number of iteration is fixed to 10,000 for this testbench. The outcome of simulation is recorded in variable 'lambda', first element is the normalized eigenvector, second element is the matrix size and the third element is the covariance matrix.