

DEPARTMENT OF MATHEMATICS, FINANCIAL MATHEMATICS
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON

Stochastic Correlation Models in Foreign Exchange Markets

Markus P. Fritz

London, March 2006

E-mail: markus.fritz@imperial.ac.uk

A thesis submitted in partial fulfilment of the requirements for the
degree of Doctor of Philosophy in the Department of Mathematics
of the University of London, and for the Diploma of the
Imperial College of Science, Technology and Medicine

Abstract

We in this thesis study the dynamics of volatility skew in the foreign exchange market. Real market data from the foreign exchange market is used to isolate and describe the dynamic patterns of the volatility surface. We also present a class of models that reproduces the same type of dynamics that we observe in the market. The class is based on the concept of stochastic volatility where skew dynamics is introduced by making the correlation between the spot price and the volatility stochastic. This way the market is described by three stochastic diffusion processes.

A specific model choice is considered and a numerical scheme to price options under this model is presented. Further the impact if making correlation between the spot level and the volatility stochastic is investigated for vanilla options, American barrier options (also known as one-touch options), forward starting options and partial time knock-out options. We find that the new model reproduces phenomena observed in the real market not explained by previously published models.

To my Father

Contents

1	Foreign exchange terminology in short	12
1.1	Fundamental market language	12
1.2	Interpretation of strangles and risk-reversals	16
2	Problem description	19
2.1	The problem we would like to solve	19
2.2	Transition probabilities	21
2.3	What the solution will be used for	23
3	Previous work on similar problems	25
3.1	Local volatility	25
3.1.1	Example of skew dynamics under the smile	26
3.2	Universal volatility model	30
3.3	Stochastic skew model by Jäckel	30
3.4	Time changed Lévy processes	31
4	Historical Market Data	33
4.1	Used set of market data	33
4.2	Fundamental empirical observation	33
4.3	Empirical study of risk-reversal dynamics	36
4.3.1	Risk-reversals and spot-volatility correlation	36
4.3.2	Risk-reversal mean reversion	38
4.4	Independent component analysis (ICA)	41
5	Modelling the market	46
5.1	Modelling the spot price	47
5.2	Modelling stochastic volatility	47
5.2.1	Different model processes	48
5.2.2	The freedom to choose a simple model	50

CONTENTS

6	Stochastic correlation from a model perspective	54
6.1	Heston model fitted to market data	55
7	Introducing the model	59
7.1	What we try to explain and quantify with the model	59
7.2	The model	59
7.3	Drift and diffusion of correlation process	60
7.4	Existence and uniqueness	62
7.5	Admissible values of correlation process	65
7.6	Transformation of correlation process	68
7.7	Expected value and long time run distribution of the correlation process	70
8	The market model and its parameters	74
8.1	Partial differential equation	75
8.2	Mutual correlation	76
9	Markovian properties and market completeness	81
9.1	Assumptions of Markovian properties	81
9.2	Market completeness	81
9.3	Market price of risk	82
10	Qualitative analytic approximations	85
10.1	Asymptotic pricing and series expansions	85
10.2	Hull-White approach to correlation value in a stochastic volatility model	86
10.3	Fourier transform techniques	86
11	Numerical approximations	88
11.1	Monte-Carlo simulations	88
11.2	Tree or lattice models ⁵⁵	88
11.3	Finite differences	89
12	Finite differences	90
12.1	Alternating direction implicit (ADI)	91
12.2	Boundary conditions	92
12.3	Stability, convergence and spurious oscillations	93
12.4	Recycling finite-differences results	95
12.4.1	Call and put options	96
12.4.2	One-touch options	96
12.4.3	Forward starting call and put options	96

CONTENTS

13 Results	99
13.1 Vanilla options	99
13.2 One-touch options and first generation exotic products	106
13.3 Forward setting options	110
13.4 First generation products with a twist	114
14 Model calibration to observed market prices	116
15 Practical use of the model	119
15.1 Products used for hedging new exposure	119
15.2 Trader will still only hedge delta and vega	121
15.3 Bid/ask spreads and transaction costs	123
16 Further areas of research	126
16.1 Term structure of variables	126
16.2 Local parameters	127
16.3 Jumps and events in all three dimensions	128
16.4 Cross currencies	128
A Differentials against option delta	130
B ADI engine (C++)	133

List of Figures

3.1	Stylised implied volatility surface	28
3.2	Local volatility surface	28
3.3	Impact of changed spot under local volatility	29
3.4	Impact on 1M risk-reversal per unit change in spot level under local volatility	29
4.1	Higher frequency of change in risk-reversal over strangle	34
4.2	Daily change in risk-reversal against daily change in spot level . .	35
4.3	Daily change in strangle against daily change in spot level	35
4.4	Evidence of risk-reversals being connected to correlation	37
4.5	Risk-reversal regression showing stability in parameter magnitude	39
4.6	Further regression results for risk-reversals	40
4.7	ICA components describing EURUSD dynamics	43
4.8	ICA components describing USDJY dynamics	43
4.9	ICA components describing GBPUSD dynamics	44
4.10	ICA components describing USDCHF dynamics	44
5.1	Simulation of three models for stochastic volatility	53
6.1	Parameter time series for market calibrated Heston model	56
6.2	Risk-reversal and calibrated Heston correlation	56
6.3	Strangle and calibrated Heston volatility of variance	57
6.4	Implied volatility and square root of calibrated Heston initial variance	57
7.1	Diffusion term and its modification outside $[-1 - \epsilon, +1 + \epsilon]$. . .	63
7.2	Initial correlation value impact on effective correlation	71
7.3	Stationary probability density using a minimisation technique . .	73
8.1	Relations and correlations for market factors	79
12.1	Stylised scheme of node subset holding correct prices	97

LIST OF FIGURES

13.1 Volatility smile under stochastic correlation 101

13.2 Parameter impact on vanilla option premium against strike level 102

13.3 Parameters' impact on vanilla implied volatility plotted against
strike level. 103

13.4 Change in volatility smile due to change in initial correlation value105

13.5 One-touch option premium against spot level 106

13.6 Model premium correction plotted against Black-Scholes premium 107

13.7 Parameter impact on model one-touch premium plotted against
Black-Scholes premium 108

13.8 Model comparison for one-touch premium correction 110

13.9 Impact of parameter '*drw13corr*' on implied forward volatility . 113

List of Tables

13.1 Parameter values used for stochastic correlation model 100

13.2 Strangle impact on forward volatility under pure stochastic volatility 112

13.3 Model comparison for partial knock-out options 115

Acknowledgments

First I would like to thank my supervisor Dr Chris Barnett for many good ideas and patience during my studies at Imperial College. Also I want to thank the Foreign Exchange Options Trading team at Citigroup for sponsoring my PhD studies. The time I spent working in this team gave me knowledge and understanding no academic education could ever replace. Especially I would like to thank my friends and colleagues Dr William McGhee, Dr David Foster and Dr Katia Babbar. Last, but not least, I want to thank my family for all the shown support.

Introduction

The main aim of this research is not to advocate a particular market model but rather to investigate stochastic correlation models based on a stochastic volatility model. We look at real foreign exchange market data to find patterns and ideas on what a generalisation of existing market models might look like.

The work presented in this thesis is of an applied nature. Maybe it can all be summarised by what was said once when I was attending a public academic presentation at a London university on the topic of option pricing. Once the presenter had finished and answered questions from the audience a man on the first row, who worked as a trader, raises his hand and asked - "Ok, but how do I make money out of all this?" No answer was given to this frank but quite innocent question. I do not claim that money can be made on the results in this thesis, but I do claim that making money has been the point of view when looking at the problem in question.

Chapter 1

Foreign exchange terminology in short

1.1 Fundamental market language

We start by explaining some of the essential market terminology. The foreign exchange (FX) market uses a way of quoting the market slightly different from the text book explanation. The different terminology has evolved because it has proved to be more convenient as it makes the numbers involved more consistent. This in turn helps the practitioners to move between different currencies and times to maturity and still use a very similar trading framework and market intuition. For a more comprehensive and thorough description of the foreign exchange market in general we refer the interested reader to other literature (see e.g. Luca[30]).

In the foreign exchange market the underlying assets are different currencies. Market prices are quoted as currency pair exchange rates e.g. EURUSD, USDJPY, GBPUSD or USDCHF. The spot price quotation expresses the unit number of the second currency, called the *term* currency, one must pay (receive) in the market to buy (sell) one unit of the first currency, called the *base* currency.

When describing a regular call or put option the strike level is often not quoted as an absolute level. Instead the option's *delta* is quoted. This corresponds to the absolute value of the regular Black-Scholes value of the options delta expressed in per cent ranging from 0% to 100%. A low delta corresponds to an option out-of-the-money and a high delta corresponds to an option in-the-money. An at-the-money option has approximately a delta of 50%¹. When the

¹Drift and interest rates make small, but non-negligible changes to the delta of an ATM-

strike of a call option passes the at-the-money level the convention is to instead trade put options. This is due to the traders' preference for out-of-the-money options over in-the-money options. An out-of-the-money option needs a smaller delta hedge and put-call-parity ensures put and call options with the same strike and maturity to have the same essential properties. All this results in that one rarely hears quotes of strike levels with a delta value higher than "50". As long as the involved interest rates are known every strike level can uniquely be represented as either a call or a put delta. In short, writing such strikes could be written as "25C" which means the strike level of a 25-delta call option or "10P" which means the strike level of a 10-delta put option.

This way of describing the strike level causes some slight numerical issues in the presence of a volatility smile. An option's delta is dependent on both its absolute strike level and volatility, but if the volatility itself is dependent on the strike level, finding strike levels is not straight forward. Finding an absolute strike level from a delta level demands the use of numerical search techniques.

The delta of an option is approximately equal to the risk-neutral probability of it ending up in-the-money at expiry². Traders often use the two quantities interchangeably, some because of convenience and some because of mild ignorance.

To make things even more complicated there are many different definitions of delta in the FX world. We will not go further into these complications other than mentioning that there are up to eight different definitions of delta. It all has to do with what accounting currency the market participant uses, the way the currency pair is quoted and whether hedging is done with spot or forward contracts.

In textbooks the *at-the-money* (ATM) strike is defined as the strike level equal to the current spot price. In the FX market one instead often chooses the strike level that will give the call and put option the same absolute value of the delta. This definition is naturally called "delta neutral".

Again to complicate things this definition is not unique but there are in total three different types of the at-the-money strike. Apart from the two definitions mentioned above one can also use the forward value as the strike level. This is often the case for emerging markets where the interest rate differential can give the currency pair a significant (risk-neutral) drift.

When trading currencies and derivatives of currencies there are many issues regarding dates to keep in mind. All involved parties need to agree on when the actual exchange of currencies is going to take place. Even though money

option. It is also dependant on what definition of ATM is chosen.

²In the Black-Scholes framework this is a result of

$\Delta = e^{-qT} \cdot N(d_1) \approx N(d_2) = P[S \text{ is in-the-money at expiry}]$

can be transferred electronically by the press of a button, transactions must first be booked, verified, double checked and confirmed with the counter party. This job is mainly the job of the banks' "back office" staff. To give the involved counterparties enough time to perform these tasks all transactions need two, or just one for a few currencies, business days for each party involved in a deal. As not all countries share the same holidays, one needs to keep track of all business days for all of the currencies you are dealing in. The date of delivery is called *value date* or *settlement date*. Buying currencies on the spot market consequently means that money will be exchanged on today's value date. This date is also simply called the *spot date*.

The date on which a contract is agreed on is called the *deal date*. The expiry date of an option is quoted either as a certain date or as a time period, also called *tenor*, from today. Tenors could e.g. be periods of two weeks ("2W"), six months ("6M") or five years ("5Y"). When a specific date is quoted the meaning of the agreement is straight forward. If the specified date does not coincide with a regular tenor the date is called a *broken date*.

If the expiry, however, is quoted as a tenor this has to be translated into a specified date. This translation can be very complicated and again we need to keep track of the different business days in different parts of the world. As the FX option market originally was an offspring of the FX forward market, the definition of tenors are adopted from this market with the difference that an options expiry date is chosen to have its corresponding settlement date equal to the tenor's *forward date*. This way any transactions originating from the option will coincide with the forward date and hence also with transactions originating from trading the forward contract. We shall not further specify how forward dates are calculated rather than noting that this is sometimes not a straightforward procedure.

Once the expiry and the strike level are agreed on, the market participants need to settle on a price or *premium* for the contract. This premium is most often not denoted in actual monetary units but in terms of implied volatility. As the premium, calculated using the Black-Scholes formula, is monotonically increasing with volatility, every premium has a unique corresponding implied volatility. In the same way a market-maker normally quotes both a bid and an ask price one for options quote a bid and ask volatility.

One of the benefits of quoting the premium this way is that the values involved are of the same magnitude for all expiry times and strike levels. The premium of a short dated out-of-the-money option will be tiny in comparison to a long dated at-the-money option. The implied volatility of the two will however be relatively similar.

It is a well known fact that the real world market behaves in a way differ-

ent from the Black-Scholes market model and observed market prices will not be fully explained by this model. Options with the same maturity, but with different strike levels, are priced at different implied volatilities. In the foreign exchange world, low-delta options are often relatively more expensive than at-the-money options. As the Black-Scholes model assumes those options to have the same volatility this immediately reveals to us that the model does not tell the full story. Implied volatility is hence only a convenient way to quote prices and does not have much to do with the “true” or expected volatility from today to expiry. Volatility is because of this, in some sense, even a non-unique concept as we just established that the very model used to define it is in fact faulty.

A market maker must be able to quote prices for arbitrary strike levels and expiry dates. To do this a trader will actually not keep track of, and update, where the market trades every single strike level. Instead, only a few benchmark levels at different strike levels and maturities are updated. All other maturities and strike levels can be calculated using various interpolation techniques. The benchmarks are set for all the liquidly traded tenors and strike levels. As an example a market maker can use the tenors 1W, 2W, 1M, 2M, 3M, 6M, 9M, 1Y and 2Y. At every tenor he or she keeps track of the five levels 10P, 25P, ATM, 25C and 10C. If a particular market is less liquid, fewer tenors could be used, and sometimes the 10-delta levels can be omitted.

Instead of holding a table of volatilities for tenors and strike levels, market makers hold a slightly different set of parameters. The volatility for the ATM option is stored in its pure form. Apart from this, market makers hold the values for what are called *strangles* and *risk-reversals*. A strangle is a structure of a long out-of-the-money call option and a long out-of-the-money put option. A 25-delta strangle (“25STR”) means that the involved call and put options both have a delta of “25”. The volatility number for a strangle is calculated as the mean difference in volatility from the ATM volatility.

A risk-reversal is a structure of a long out-of-the-money call option and a short out-of-the-money put option. As for the strangle a risk-reversal is called a 25-delta risk-reversal (“25RR”) if both the options have a delta of “25”. The volatility number of a risk-reversal is calculated as the difference in volatility between the call and the put. In practice one only quotes the absolute number, and tells whether it favours the call or the put. One can also explicitly write “+” or “-” to clarify which side is relatively more expensive.

As volatility changes frequently, and often in parallel, the advantage of quoting strangles and risk-reversals instead of direct volatilities is that the numbers then need to be updated less often. This practice also gives a much better overview of the shape of the volatility surface and volatility smile curves at explicit tenors. Strangles give information about the curvature and risk-reversals

give information about the skew.

1.2 Interpretation of strangles and risk-reversals

We know that quotation of strangles and risk-reversals is a way for the market to put convenient numbers on the shape of the implied volatility smile curve for different tenors. The curve is, in turn, a way for the market to put more convenient numbers on option premia. But what is the market's reason for the smile curve and hence strangles and risk-reversals? In terms of risk-neutral probability those numbers describe the probabilities of the spot price ending up in different regions and especially the increased probability of ending up in the far-out tails. Many different factors come into play when trying to explain the reasons for this increased tail probability and the asymmetry of its increase. What phenomena is the market accounting for when pricing one direction higher than the opposite direction?

From the market makers' perspective the premium increase for some strike regions is mostly a question about supply and demand or the market makers' own preferences and current position. One must however not forget that supply and demand are results of more fundamental market behaviour based on aggregated views of how the market will behave in different situations.

Positive strangles, equivalent to volatility curve convexity, assign a relatively higher probability of ending up in the distribution tails compared to a log-normal distribution. "Fat tails" of a distribution is in mathematical terms called *kurtosis*. This is more precisely specified as a measure of the distributions fourth moment. In short, fat tails can be explained by the fact that when the underlying spot market actually moves, it often moves by a lot. An out-of-the-money option only pays out if the market makes a large move in the right direction. Conditioned on a large move happening, the probability of a very large move is in the real world higher than what log-normality would predict. One part of the explanation is that conditioned on a large move the probability of a higher volatility rises. In such case the probability of an increased payout is also higher. However this forces us to depart from the basic assumption of constant volatility.

Another part of the explanation is forehand knowledge about possible upcoming events that may (or may not) move the market in a discontinuous, or at least close to discontinuous, way. Events often also make the market more volatile, which again increases the probability of a higher payout. Not all large market moves are predictable and sometimes the market can move discontinuously for apparently no obvious reason. All such movements are examples of

1.2. INTERPRETATION OF STRANGLES AND RISK-REVERSALS

jumps in the market.

So far we have only tried to explain fat tails as a general phenomenon. This only concerns the strangles but not the risk-reversals. In terms of distribution tails the risk-reversals describe the relative difference in tail fatness, which is called the distribution *skew*. This is a measure of the third moment of the distribution. As for kurtosis, the skew can partly be explained by beliefs about future volatility levels conditioned on the future level of the spot price. While convexity assigns a symmetric change in volatility to general changes in spot price, regardless of direction, the skew assigns an asymmetric change in volatility depending on the direction of the spot move. The reasons behind this asymmetry are a subject of debate.

In the equity market one often talks about “leverage effect” or “volatility asymmetry”. This concept too is debated but roughly it means that the impact on the valuation of a company, as a result of changes in business condition for the company, is not directly linked to the current stock price. However the relative impact of such changes on the stock price is linked to the absolute level of the stock price. As a result, the volatility of the stock price will rise as the stock price decreases and vice versa. (Originally the term comes from the financial concept of “leverage” that deals with the difference in return on equity and the return on capital employed by a corporate or institution.)

Reversing the causality one can also argue that an increase in uncertainty, and hence volatility, would make the stock less attractive and hence result in a falling stock price.

The leverage argument is of course not directly translatable into currency areas and the foreign exchange market. Some authors even talk about an “inverted leverage effect”, favouring the upside volatility, due to the asymmetric nature of intervention actions taken by some central banks (or rather, asymmetries in the market’s beliefs about possible interventions). This “invertedness” is dependent on the way the currency pair is quoted and the prevailing monetary policy adopted by the concerned central bank. Rather than having a general direction, volatility skew in the foreign exchange market could be characterized by its quick changes.

One can here mention the well known historical defence of a weak yen by Bank of Japan. In September and October of 2003, and in some sense also in March of 2004, the market was uncertain about the ability of the Bank of Japan to maintain a weak yen. A change in intervention policy could lead to an undefended drop in the USDJPY exchange rate. This risk resulted in a general hike in implied volatility and a large drop in risk-reversals as investors wanted to buy put options as protections against this risk.

1.2. INTERPRETATION OF STRANGLES AND RISK-REVERSALS

In common between the above mentioned explanations is that they mathematically can be represented as a correlation, either positive or negative, between the spot price and volatility. In the above mentioned case of Bank of Japan this correlation would have been negative.

Together the convexity and the skew allow the market to “tweak” the Black-Scholes model, which assumes a flat log-normal distribution, and describe where the market believes volatility will move dependent on future moves in spot prices.

Chapter 2

Problem description

2.1 The problem we would like to solve

When looking at an implied volatility surface of a liquid market it often has a skew or a tilt in some direction. For the equity market the skew typically favours the downside. This means that out-of-the-money put options are relatively more expensive than out-of-the-money call options. Part of this can be explained by the fact that the equity market is always net long assets. Because of this the aggregated risk of the market is on the downside. The result is a net demand for protective put options, which in turn will bid up the prices for downside options. Another explanation for a negative skew specific to the equity market is the negative skew of the historical distribution (see more in Chapter 4 of Schoutens [40]).

In foreign exchange one can also observe this skew. The big difference is though that the FX skew is not static but can sometimes change quite dramatically. The dynamics are often so large that the very direction of the skew changes. Derivative traders are well aware of those dynamics and have developed methods to hedge some of this risk away¹, or at least keep the exposure to these factors under control. This risk is quite well understood for vanilla options and the handling of it is part of the daily routine. However those hedging methods most often apply to only a certain group of exotic derivatives. Included in this group are most of the so called “first generation exotics”. The aim of this research is to find a more general way to handle, and hopefully also to some extent explain, those dynamics.

The methods used by FX derivative traders today are most often based on a theoretical price generated by the Black-Scholes model with an additional cost of hedging various exposures to changes in implied volatility, including the

¹ For a clear and intuitive explanation see various articles by Wystup e.g. [46].

2.1. THE PROBLEM WE WOULD LIKE TO SOLVE

volatility skew. Those hedges often include an initial hedge of the so called *vega*, *vanna*² and *volgamma*³ and could be seen as a sort of initial “static hedge” dealing with effects not covered by the Black-Scholes model. The hedging of those factors is done by, at least in theory, trading regular vanilla options. In the case of path dependent options like knock-out options, which can become worthless before expiry, practitioners also often take the theoretical expected lifetime of the option into account. To fully adjust the cost and value of this type of hedge a trader must estimate the possible unwind value of the hedging options, should the hedge no longer be needed at a point prior to expiry. The concern here is to estimate the future value, or rather the future volatility. Unwinding often takes place at a well known spot level, as is obvious for knock-out options, and an experienced trader can often have a feeling for the market dynamics and assign a rough picture of the volatility smile at the future spot level in question. This is of course a very subjective estimation but it almost always includes the assumption that risk-reversals changes when the spot price changes.

All the just mentioned steps are inconsistent with the very assumptions of the market model used to price them. One of the fundamental assumptions is the assumption of a static level of volatility. Nevertheless this very assumption causes the most complications when pricing options. For further information in this area the interested reader is referred to, among others, Taleb [43].

If we can find a model capable of generating a dynamic skew we can compare the prices and hedges produced by this model to the prices and hedges often used by traders today. Can a new model give similar prices without having to add extra hedging costs to the theoretical price? Will the risky factors in the new model relate to the risk hedged by traders today? We also want to find a mathematically more appealing way to explain the fact that when option traders hedge the previously mentioned risk they in fact hedge “outside the model”. As we mentioned earlier the, traders’ hedging is inconsistent with the assumptions of the market model.

Some of the above questions can be answered by the introduction of stochastic volatility. One can here mention the connection we see between volatility of volatility in a stochastic volatility model (see e.g. Section 3 in Heston [17]) and volgamma in the Black-Scholes model (see e.g. Section 2.7 in Hakala & Wystup [14]). The main part still left unexplained by stochastic volatility is the dynamics of the volatility skew. In a stochastic volatility model, changes in spot level will only change the at-the-money volatility, conditional on a non-

² “Vanna” is the derivative of vega w.r.t. spot or equivalently the derivative of delta w.r.t. volatility.

³ “Volgamma” or “volga” is the second derivative (convexity) of premium w.r.t. to volatility.

zero spot-volatility correlation, but not change the volatility skew. As a result a stochastic volatility model has an essentially static skew. In terms of implied volatility plotted against option strike level the shape of the volatility surface will for such a model “float along” and move in parallel with changes in spot price without a change in smile shape. (Due to spot-volatility correlation the at-the-money volatility level of the smile however also can change if the spot level changes.)

A volatility smile moving along with changes in strike level is commonly said to have *sticky delta* or *sticky moneyness*. After a move in spot price, options with the same option delta will still have the same implied volatility as before the move. Implied volatility for an absolute strike level will however change. If the implied volatility instead is associated with an absolute strike level the behaviour is called *sticky strike*. The two different smile concepts are described in Section 10.4 of Lipton [28].

Even though the sticky delta behaviour describes much of what we empirically can observe in the FX market it does not explain everything. In this market the volatility surface admittedly does move with changes in the spot price in a sticky delta fashion. However such changes also affect the skew of the volatility smile in a way that instead points in the direction of a sticky strike behaviour. The truth seems to be somewhere between the two descriptions. To explain this we need an extended stochastic volatility model and we will show that a stochastic correlation model takes us a step in the right direction.

2.2 Transition probabilities

In more general terms the problem is about finding risk-neutral transition probabilities. Given market prices for vanilla options, which are not path dependent, we are able to obtain the final risk-neutral transition probabilities from today’s spot price to different expiries and strike levels. What we cannot obtain are the intermediate transition probabilities. To do this we need to make assumptions about the involved underlying market processes. Such a model is not necessarily restricted to only the spot price dimension but can extend to other market state dimensions. We have already mentioned one such extra state dimension when we discussed stochastic volatility. The final transition probabilities obtained from vanilla prices only give us information about states of the spot price but gives us no information about transition probabilities of any possible other state dimensions of the market process. One of the simplest assumptions allowing us to uniquely calculate the intermediate transition probabilities is the local volatility model. This model does not make use of the possibility of other

2.2. TRANSITION PROBABILITIES

state dimensions.

Once a market model is chosen and the intermediate transition probabilities are isolated we can price path dependent options and in particular exotic options. In some vague sense we could say that the pricing of exotic options is, or at least has a component that is, “orthogonal” to the pricing of vanilla options. With this we mean that two market models, possibly the same fundamental model but with a different set of parameters, could agree on the prices of vanilla options but disagree on the prices of exotic options.

The problem is not to find a market model replicating the observed vanilla prices. This is the so called “matching problem” and should rather be thought of as a constraint on any model we choose to represent the market. The true problem lies in finding a market model which can also produce sensible prices for exotic options. We are looking for a model for which vanilla prices alone are not enough for model calibration.

If we know all the conditional intermediate transition probabilities we also know all the future conditional implied volatility surfaces or “forward smiles”. The reverse argument is also true – if we know all the future volatility surfaces we also know all the intermediate transition probabilities. Finding a market process that generates plausible future implied volatility surfaces would mean that we might also assume that the model’s transition probabilities are plausible. In turn this would lead to better pricing of all contracts dependent on the intermediate transition probabilities. Those contracts are all path dependent contracts.

To get a more intuitive idea of the problem with intermediate transition probabilities one can imagine a market where both time and spot level take only discrete values. The continuous case could be approached by making the discrete steps finer. The local volatility model can in each time step reach three states in the immediately succeeding time step. This gives three transition probabilities per node. For every state we must match the forward value, the European vanilla option premium and make the three conditional probabilities sum up to the unit probability, one. The system is fully determined and leaves no further degrees of freedom apart from the placement of nodes in time and space. This leaves no possibility to control premia of path dependent derivatives and has hence no “orthogonality” to the pricing of European vanilla options.

If we assume our general market model to be Markovian we need our model to have as many dimensions as our market has degrees of freedom. With degrees of freedom we mean non-excludable market property states – e.g. level of spot price, level of volatility, general shape of the volatility surface and interest rates etc. Those market property parameters can of course be mutually correlated but are not redundant in the sense that one can be uniquely determined as a function of the others.

2.3. WHAT THE SOLUTION WILL BE USED FOR

The immediate danger of increased model dimensionality is the risk of unwanted “data mining”. This is the risk of successful model calibration without giving the model any explanatory or predictive abilities. When calibrating a model with a high-dimensional parameter space to a finite set of observations there also may be many sets of model parameters resulting in a close fit to the observed data. In such case a closer fit does not necessarily mean a better model. As a rule of thumb, fewer parameters are better than more and we need a strong reason to make a model more complicated.

2.3 What the solution will be used for

As previously mentioned, pricing exotic options today often involves many “padding layers” of hedging costs and approximations. A model reproducing more of the observed dynamics will also take more of the risk into account. This would give a more consistent framework for pricing exotic options exposed to skew risk and other dynamics.

Once the option is priced and traded we need to hedge the risk. The practitioners’ method of hedging “outside the model” does not give a very high level of control of the exposure to changes in the skew. The exposure is just hedged locally and the future exposure is not taken into account. A trader can have a good feeling about how the skew moves with the spot price, but finding out what impact those movements will have on the price of an exotic option is another matter. Therefore also the hedging procedure would benefit from a model taking more of the risk into account. We would also get rid of the mathematical inconsistency of hedging factors that are assumed to be constant in our model.

Trading options is always associated with taking risks. One risk taken by trading financial institutions is the risk of the credit worthiness of trading counterparties. This risk is known as *credit risk*. To reduce the impact of this risk, trading market participants impose credit limits and demand security deposits as a form of insurance. All those precautions taken are related to the *value-at-risk* and demand that we actually know something about the counterparty’s risk exposure. The industry has worked out methods and econometric tools to measure most of the risk. Risk that cannot easily be measured can be handled by simply demanding higher deposits. High deposits constraints, however, can result in less business as the counterparty is only allowed, or willing, to have a certain risk exposure. An area where big banks today find it hard to measure risk is the sensitivity to volatility skew of exotic options. Therefore those products often demand a high deposit. For smaller institutions, like small hedge funds, the deposits for trading exotic options can be painfully large. A model

2.3. WHAT THE SOLUTION WILL BE USED FOR

that could take the dynamics of the skew into account could possibly decrease the deposits needed and still give a comfortable control of the risk exposure. This would be beneficial to both parties as it leads to an increase in trading.

Most of the market models used today cannot explain all the dynamics of the volatility empirically observed. If we can find a market model that can explain more of what we observe, we might be able to trade those dynamics. We might even be able to construct a product with a convexity in premium towards a risk factor the market is not fully accounting for. Any change in this risk factor would then lead to an increase in value. This would mean market misspricing and possibility for arbitrage.

A model that explains more of the dynamics of the volatility surface will also help us estimate the future shape of this surface. Even if the qualitative estimations of such a model would not reveal anything that an experienced market participant cannot already make a qualified guess about, it might be able to make more quantitative estimations. As mentioned earlier, making proper use of the estimated future shape of the market when pricing exotic options can be hard even though the estimations themselves are sensible.

As a last area of use of an extended market model we might propose is the interpolation and extrapolation of the implied volatility surface for vanilla options. A model tuned to match the observed market prices can be used to price less liquidly traded options. The presumed contribution in this area is however small as already available methods, like a simple polynomial, can match observed prices very well. Finding a method explaining the observed prices outside the liquidly traded options has though been shown harder.

Chapter 3

Previous work on similar problems

Little work has been done in the area of modelling the dynamics of risk reversals. One explanation is the fact that this phenomena is relatively unique to the foreign exchange market. Another explanation is that this type of dynamics is regarded as less important and that its impact, where it matters, is handled in a more direct manner, like as for the handling of forward volatility. Despite all this there are some researchers who have, explicitly or implicitly, looked at this area.

3.1 Local volatility

The concept of local volatility was introduced and described by Derman & Kani [10], Dupire [12] and Rubinstein [39]. Local volatility is a generalisation of the original Black-Scholes one-dimensional market model process. The constant volatility parameter used by Black-Scholes is here generalized to be a deterministic function of both spot level and time. Despite the deterministic nature of the local volatility, some authors actually call this concept a basic stochastic volatility model due to the fact that volatility is governed by the spot process which itself is stochastic. This model is often said to be the simplest model that can be fully calibrated to any given set of market data. On an SDE form the market model under local volatility is given by

$$dS_t = \mu S_t dt + \sigma(t, S_t) S_t dW_t.$$

The function $\sigma(t, S_t)$ describing the volatility is at the heart of the local volatility model. This function can, up to a few regularity constraints, be chosen

3.1. LOCAL VOLATILITY

freely. Dupire showed that this function easily can be chosen to reproduce an arbitrary function of implied volatility.

Even though none of the authors point it out in the original articles, a local volatility model gives rise to some dynamics of risk-reversals as the spot level changes.¹ Those dynamics are a result of the convexity of the volatility smile surface and are therefore restricted to a positive correlation to changes in spot prices. Even though a market with an implied volatility smile showing negative convexity is mostly academic, such a market would create a negative correlation between changes in risk-reversals and changes in spot price. We will in the next section also show that the skew dynamics created by a local volatility model are of a too large a magnitude.

An obvious shortcoming in this model is that once the curvature of the implied volatility surface is fitted, we cannot change the way the skew or risk-reversal changes with changes in spot. One property locks in the other and we cannot freely choose both of them. Also, the resulting skew dynamics generated by local volatility is of a fully deterministic nature with respect to spot price. This might not be a crucial shortcoming as we are only interested in the expected value of our process under a risk-neutral approach.

A deterministic relation would however not capture any convexity in premium with respect to volatility skew that otherwise would make a continuous contribution to the premium due to quadratic variation. This matter will briefly be handled in a later section.

Further a local volatility model generates very implausible future volatility smiles as time passes. To read more about this and other effects and results of local volatility the interested reader is referred to e.g. Rebonato [36].

3.1.1 Example of skew dynamics under the smile

We will in this section look closer at and investigate the changes in the implied volatility surface generated by a change in spot price under the local volatility model.

The issue of local volatility is closely linked to the symmetry of the two versions of the Black-Scholes pricing PDE. Those versions are the most often used *backward equation* that gives a premium $\Pi(K_0, T_0)$ for a fixed maturity point at any set of $\{S, t\}$, representing different states of the present point, and the less often used *forward equation* or *dual equation*, that gives a premium

¹Especially Bruno Dupire has in later work discussed the local volatility model and the dynamics of the volatility smile arising from changes in spot price.

3.1. LOCAL VOLATILITY

$\Pi(K, T)$ for different maturity points at fixed present point $\{S_0, t_0\}$. Using the pricing version with both dividends and local volatility the two pricing PDEs are given by

$$\begin{aligned} \frac{\partial \Pi}{\partial t} + (r - q)S \frac{\partial \Pi}{\partial S} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 \Pi}{\partial S^2} - r\Pi &= 0 & \text{(Backward equation)} \\ \frac{\partial \Pi}{\partial T} + (r - q)K \frac{\partial \Pi}{\partial K} - \frac{\sigma^2}{2} K^2 \frac{\partial^2 \Pi}{\partial K^2} + q\Pi &= 0 & \text{(Forward equation)} \end{aligned}$$

Rearranging the terms in the forward equation we reach the local volatility and the so called *Dupire formula*

$$\sigma(K, T) = \sqrt{2 \frac{\frac{\partial \Pi}{\partial T} + (r - q)K \frac{\partial \Pi}{\partial K} + q\Pi}{K^2 \frac{\partial^2 \Pi}{\partial K^2}}}. \quad (3.1)$$

To be able to use the Dupire formula it is necessary to know the prices for vanilla options at all strike levels and at all maturities. If we instead know the local volatility we can use the forward equation to calculate today's prices for options expiring in the future. This must most often be done numerically. (A more efficient way to reach local volatility when given implied volatilities is to substitute Π in Dupire's formula (3.1) with the Black-Scholes pricing equation and explicitly write the differentials.)

To investigate the dynamics of the skew due to changes in spot price we must first have today's implied volatility surface. For simplicity we assume this to be a simple stylized surface. We assume this surface to have no term-structure and to have a smile given by a simple quadratic polynomial in the options delta value (calculated using the at-the-money volatility). Looking at historical values for EURUSD it seems plausible to choose the 25-delta strangle to be 0.25% and the 25-delta risk-reversal to be zero

The graph in Figure 3.4 shows that if the local volatility surface is held fixed, a spot price move of 1% changes the 1-month (1M) 25-delta risk-reversal in the same direction as the spot move by approximately 0.7%, measured in implied volatility. For a real currency pair, like EURUSD, a 1% spot move changes the 1-month 25-delta risk-reversal by a magnitude in the region of 0.1%-0.2%. The conclusion is that skew dynamics produced by a pure local volatility model is of a too large magnitude. Or alternatively, in order for a pure local volatility model to produce skew dynamics of the right magnitude, it will have to produce too little smile curvature and hence undervalue strangles.

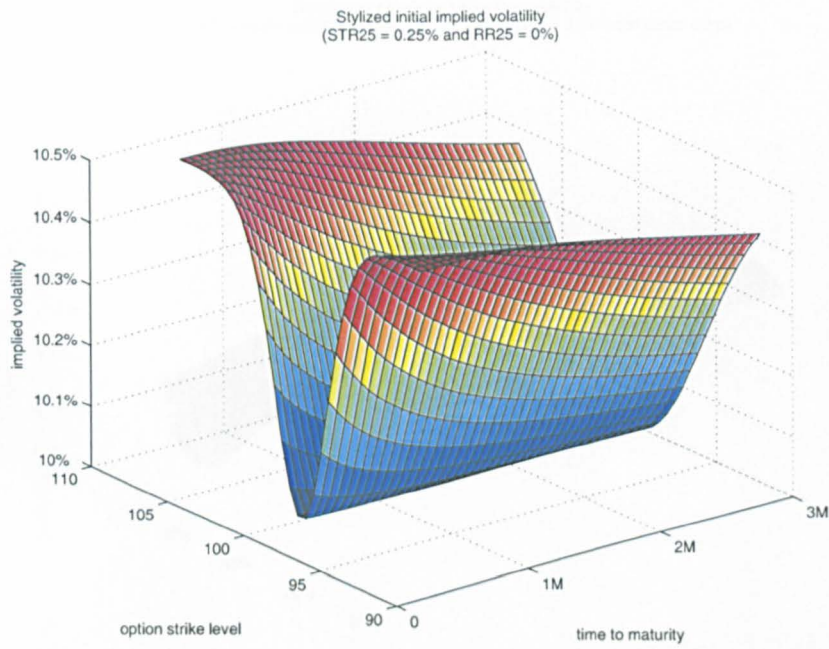


Figure 3.1: Stylised implied volatility surface with $ATM = 10.0\%$, $STR25 = 0.25\%$ and $RR25 = 0.0\%$.

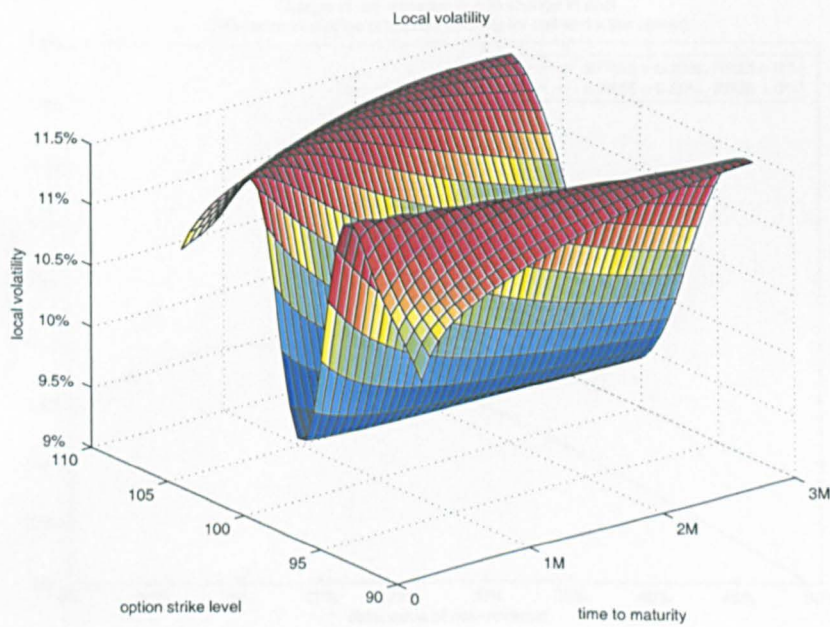


Figure 3.2: Local volatility surface corresponding to stylised implied volatility surface

3.1. LOCAL VOLATILITY

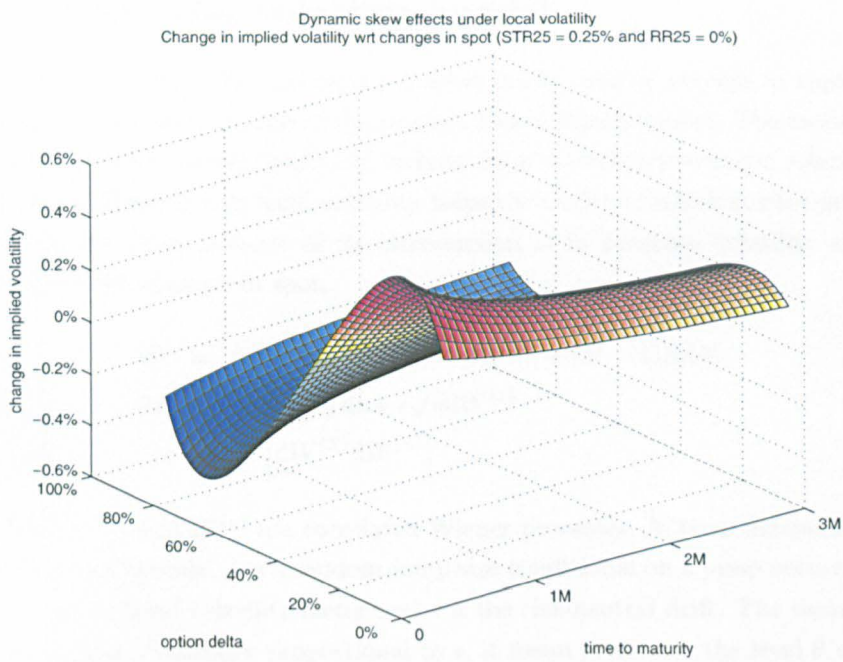


Figure 3.3: Impact on implied volatility at fixed levels of option delta per per unit change in spot level under local volatility

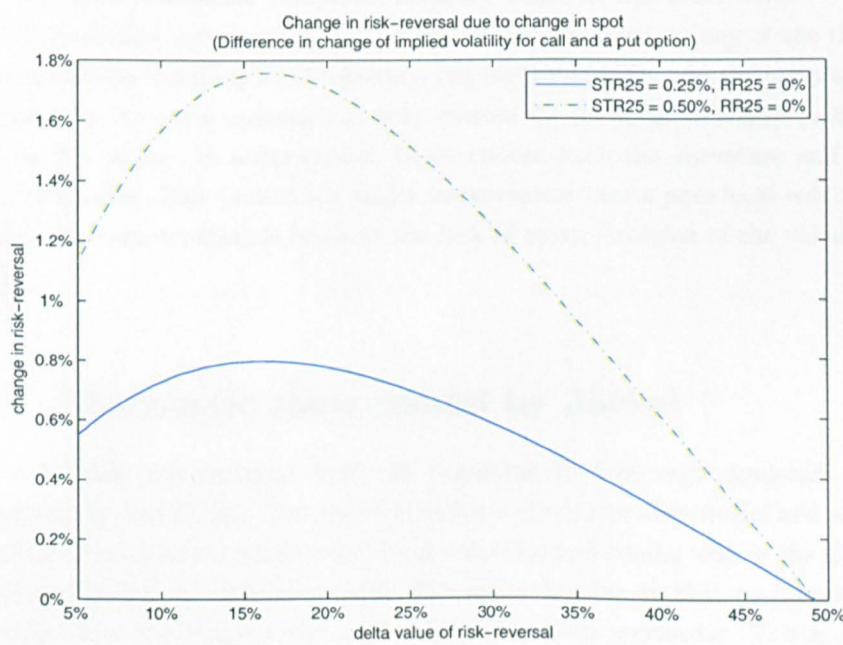


Figure 3.4: Impact on 1M risk-reversal per unit change in spot level under local volatility

3.2 Universal volatility model

Lipton and McGhee [29] proposed a market model general enough to capture all the popular modifications of the original Black-Scholes model. The model is called *universal volatility model* and includes local volatility, stochastic volatility and jumps. Even though local volatility helps the model to match market prices perfectly the main purpose of its introduction is to generate volatility skew dynamics with changes in spot.

$$\begin{aligned} dS &= rSdt + \sqrt{v}\sigma(t, S) SdW^{(S)} + (e^j - 1)SdN \\ dv &= \kappa(\theta - v)dt + \epsilon\sqrt{v}dW^{(v)} \\ \rho &= E\left[dW^{(S)}dW^{(v)}\right] \end{aligned}$$

Here $W^{(S)}$ and $W^{(v)}$ are correlated Wiener processes, N is an independent Poisson jump process, j is a random jump size conditional on a jump occurring, $\sigma(t, S)$ is the local volatility factor and r is the risk-neutral drift. The variance process v has a volatility proportional to ϵ , it mean reverts to the level θ with a reversion speed κ .

Even though the model is in no way dependent on it, this very setup uses a Cox-Ingersoll-Ross (CIR) or Heston-style process for stochastic volatility. In fact any valid process for stochastic volatility would fit this framework.

The volatility curvature in this model can be generated by any of the three generalisations including local volatility, stochastic volatility and the jump component. As the skew dynamics is only created by the local volatility part we can in this model, to some extent, freely choose both the curvature and the skew dynamics. This control is a major improvement over a pure local volatility model. One shortcoming is however the lack of mean reversion of the volatility skew.

3.3 Stochastic skew model by Jäckel

A model that was designed from the beginning to have skew dynamics was presented by Jäckel [24]. The model is called a *stochastic skew model* and has a mixture of stochastic volatility and local volatility in a similar way as the above described universal volatility model. The difference lies in that no jumps are used and that the shape of the local volatility is made stochastic. This is done by describing the local volatility by an exponential function dependent on spot level and a special shape parameter, γ , that in turn is made stochastic. The

model is given by

$$\begin{aligned}
 dS &= \mu S dt + \sigma e^{\gamma(\frac{S}{H}-1)} S dW^{(S)} \\
 d(\log \sigma) &= \kappa_{\sigma} (\log \sigma_{\infty} + \log \sigma_t) dt + \alpha_{\sigma} dW^{(\sigma)} \\
 d\gamma &= \kappa_{\gamma} (\gamma_{\infty} - \gamma_t) dt + \alpha_{\gamma} dW^{(\gamma)} \\
 0 &= E \left[dW^{(S)} dW^{(\sigma)} \right] = E \left[dW^{(S)} dW^{(\gamma)} \right] = E \left[dW^{(\sigma)} dW^{(\gamma)} \right],
 \end{aligned}$$

where H is an arbitrary spot level around which the scaling of the volatility parameter σ is centred and equal to e^{γ} . The level of mean reversion for volatility, σ , and shape parameter, γ , are given by σ_{∞} , and γ_{∞} , respectively.

Even though the model allows for it, Jäckel chooses not to have a correlation between spot level and the shape parameter. As we can see a strong correlation between spot level and risk-reversals in the foreign exchange market, a positive correlation would be a more realistic choice for this particular market.

A problem with this model is the fact that skew dynamics are based on a spot level close to the level H . If the spot price leaves this region the stochastic shape parameter γ will, instead of just only changing the skew, also change the level of volatility and hence effectively increase the volatility of volatility. As mentioned earlier volatility of volatility has a positive relation to the curvature of the volatility smile. As a result this stochastic skew model will have a positive correlation between the spot price and the price of strangles. Such a relation is not observed in real market data.

Jäckel also mentions an interesting phenomenon he calls “jumps without jumps”, referring to the case of a sudden peak in volatility along with a rapid change in spot price. This results in the already large move in spot level being amplified due to the high level of volatility into something that is similar to a jump in the spot level. Such phenomena should be observable in any model where volatility is made stochastic.

3.4 Time changed Lévy processes

Carr and Wu identify the issue of dynamic risk-reversals in [4] and [5] and present a model to generate such dynamics.

Their approach is based on a spot price driven by a combination of two timechanged Lévy processes. Each Lévy processes in turn has an independent diffusion part and a pure jump part. The time change is governed by an activity rate process for each Lévy process. Jumps cause a skew effect and the activity rates make this skew dynamic. By making the drivers of the activity rates be

3.4. TIME CHANGED LÉVY PROCESSES

correlated to the diffusion term of the corresponding Lévy process, a correlation between spot price and skew is created.

The approach is very interesting as it differs much from techniques used in previously published work. While the model, in some sense, is mathematically beautiful, it is less intuitive than a traditional approach. This suggests a resistance threshold for introducing the model to the trading community.

Chapter 4

Historical Market Data

We now turn our eyes at the actual history of the market. Apart from qualitative observations this is essentially a statistical investigation of historical market data. The framework of statistical tools is huge and one could spend much time and effort investigating the form and statistical significance of the mutual dependencies between different market parameters. Even though those aspects are important we will here mostly be looking for quantitative relations.

4.1 Used set of market data

The data set we use is real historical daily market data for spot prices, interest rates and implied volatility for the currency pairs EURUSD, USDJPY, USDGBP and USDCHF. All data is given for the tenors 1W, 2W, 1M, 2M, 3M, 6M, 1Y and 2Y. The volatility surface is, at every tenor, described by values for the strike levels ATM, 25STR, 10STR, 25RR and 10RR. In total our volatility surface data series have a cross section of 40 data points for each business day. The data set starts at the first business day in year 2000 and spans over approximately 4.5 years which is equal to some 1200 readings. In the following text, when we talk about daily changes we implicitly mean changes between business days and hence ignore weekends and holidays.

4.2 Fundamental empirical observation

To gain some qualitative feeling for the data we first do some visual investigations and simple tests for correlation between the most frequently updated market figures. For volatility data the most liquidly traded tenors are around

4.2. FUNDAMENTAL EMPIRICAL OBSERVATION

1M. Those are naturally also the most frequently updated figures. As most liquid currency pairs show a very similar behaviour we will here only present data for EURUSD.

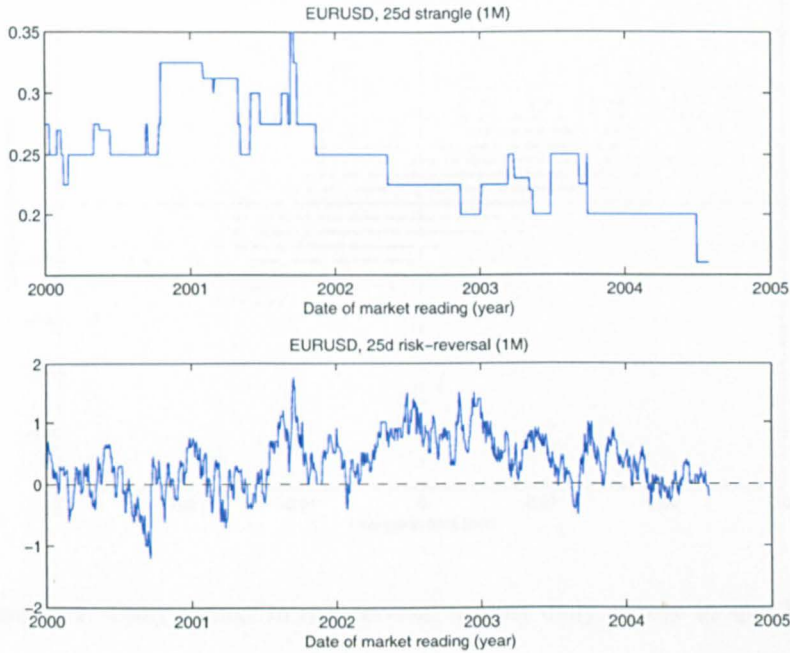


Figure 4.1: EURUSD market data showing a higher frequency of change in risk-reversal over strangle

The first observation is that risk-reversals change very frequently and that its value changes sign on many occasions. Conversely the strangles act very differently, with numbers changing much less often. The same strangle level can prevail for many months and when it eventually changes, it does so in very small relative terms. We must not forget that those numbers are set by traders and the sticky behaviour of the strangles is rather a sign of changes in strangles not being considered as important as changes in risk-reversal. As a result, strangles are updated less often and with a higher relative granularity. The opposite is true for risk reversals to which more attention is paid. Risk reversals are updated much more frequently and with a higher relative precision.

Figure 4.2 shows the daily changes in value of the strangles and risk-reversals plotted against the change in spot price.

Looking at the Figures 4.2 and 4.3 we immediately see a quite strong correlation in one them. The behaviour of risk-reversals show a much higher depen-

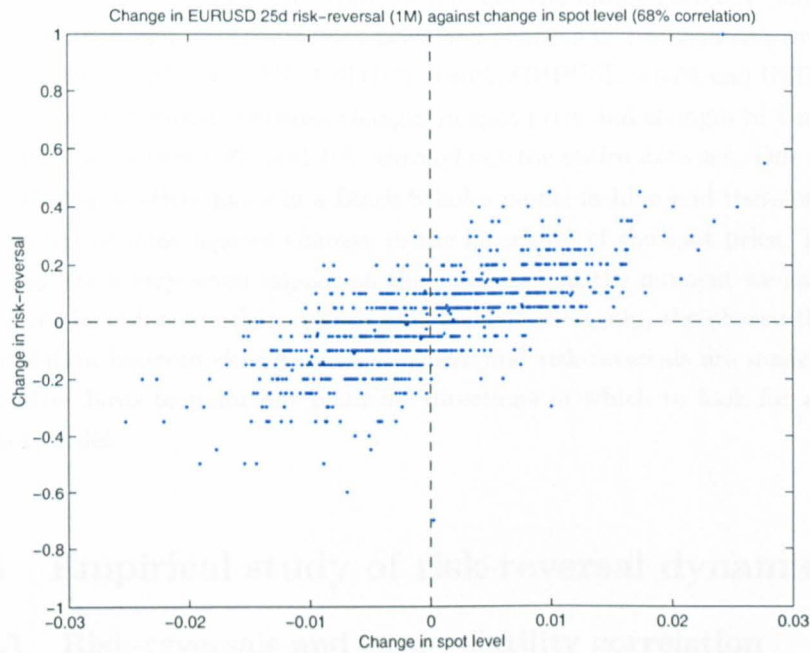


Figure 4.2: Daily change in risk-reversal against daily change in spot level

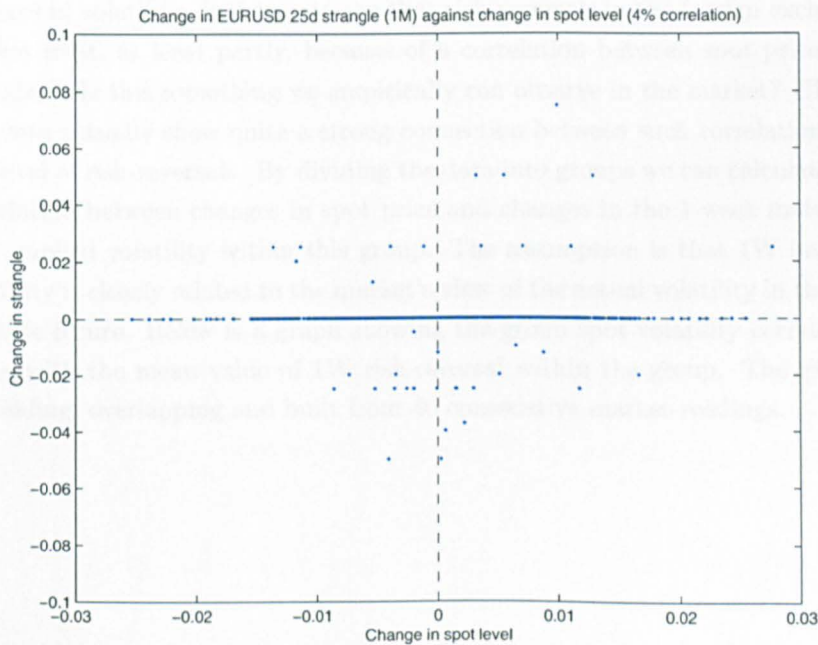


Figure 4.3: Daily change in strangle against daily change in spot level

dence on spot moves than the strangles do. For the major currency pairs the correlation between changes in spot price and changes in 1-month risk-reversal are given by: EURUSD +68%, USDJPY +66%, GBPUSD +54% and USDCHF +53%. The correlations between changes in spot price and changes in 1-month strangles stay between 0% and 10% throughout the entire data set. One could argue that spot prices move in a Black-Scholes model fashion and therefore the plot should be done against changes in the logarithm of the spot price. Doing this only has a very small impact on the result and at the moment we have no model of the risk-reversals to justify this behaviour. Again, the observation of a correlation between changes in spot prices and risk-reversals are made on a qualitative basis to point out plausible directions in which to look for a new market model.

4.3 Empirical study of risk-reversal dynamics

4.3.1 Risk-reversals and spot-volatility correlation

So far we have focused on the prices of risk-reversals and their connection to spot prices. We earlier also briefly discussed the reasons for the existence of risk-reversals – one of them being the correlation between changes in spot price and changes in volatility. Is it true to say that risk-reversals in the foreign exchange market exist, at least partly, because of a correlation between spot price and volatility? Is this something we empirically can observe in the market? Historical data actually show quite a strong connection between such correlation and the level of risk-reversals. By dividing the data into groups we can calculate the correlation between changes in spot price and changes in the 1-week maturity, 1W, implied volatility within this group. The assumption is that 1W implied volatility is closely related to the market's view of the actual volatility in the immediate future. Below is a graph showing the group spot-volatility correlation along with the mean value of 1W risk-reversal within the group. The groups are sliding, overlapping and built from 40 consecutive market readings.

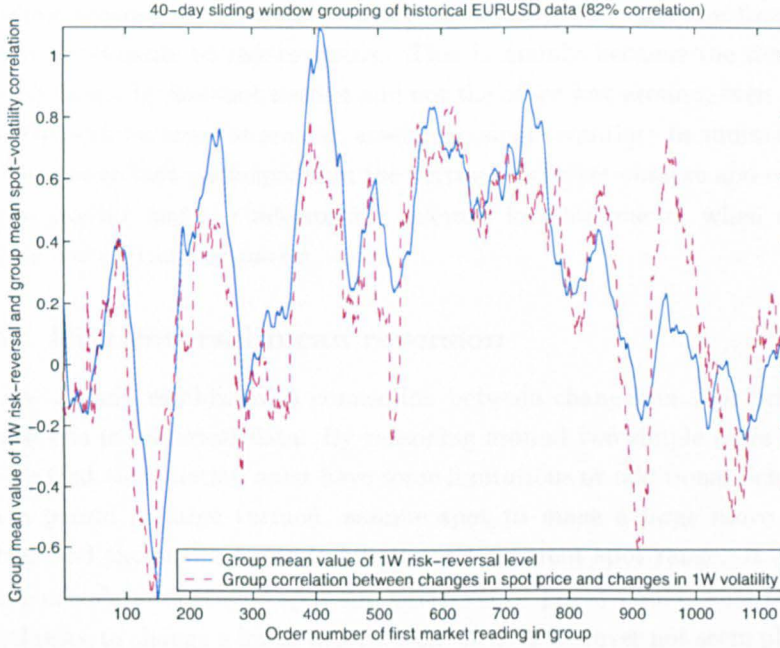


Figure 4.4: Evidence of risk-reversals being connected to correlation

Figure 4.4 shows a clear and positive relation between the spot-implied volatility correlation and the mean value of risk-reversals. This clear connection is observable for all our four main currency pairs. The graph correlation is for EURUSD +82%, USDJPY +82%, GBPUSD +80% and USDCHF +79%. It is also interesting to see that the correlation changes sign at many points in time and that this happens in parallel with a similar behaviour for the risk-reversals.

It is important to keep in mind that this plot shows the correlation between spot price and *implied* volatility. Of course implied volatility is not the same as actual volatility, in whatever sense we choose to define volatility, but it is a close indicator of where the market believes this volatility to be. The level of implied volatility here only affects the calculation of group spot-volatility correlation and we could therefore, without a loss in result, allow for a bias between implied volatility and actual volatility. The correlation measure is insensitive to such bias up to a linear transform as $\text{corr}(X, Y) = \text{corr}(\alpha + \beta \cdot X, Y)$.

As for all correlated observations we must also not forget to consider the causality of the relation. We cannot with this simple graph, at least not from a mathematical perspective, tell whether risk-reversals are a result of the state of the correlation between spot price and volatility or if this correlation is a result by some unknown phenomenon quantified by risk-reversals. It is however

a plausible assumption that the causality operates in the direction from spot-volatility correlation to risk-reversals. This is mainly because the derivative market is driven by the spot market and not the other way around, even though this in fact can be true for smaller assets like single equities. In summary this causality means that participants in the derivative market observe and consider the spot market and the information relevant for this market when making quotes in the derivatives market.

4.3.2 Risk-reversal mean reversion

We have already established a connection between changes in spot price and risk-reversals in historical data. By reasoning around two simple cases we can conclude that this relation must have some limitations or additional behaviour.

In a period of large turmoil, assume spot to make a large move in any direction and then calm down and settle at a different spot range. It is likely that the correlation between spot and risk-reversal prices would cause the risk-reversal price to change a lot in this turmoil. It does however not seem plausible that risk-reversals would stay at a historically deviating level once the period of turmoil has passed. (This is of course conditional on that the currency's economy is not expected to behave in an essentially different way after the large spot move.)

Another case to consider is a long term steady trend. Under this scenario the correlation to spot changes would make risk-reversals reach extreme levels after some time. Even though spot undergoes large changes, those changes are stable and anticipated by the market. It is plausible to assume that risk-reversals would reach some sort of steady state after some time.

In both the above cases the risk-reversals would have plausible behaviour if there was some sort of mean-reversion in its values. In the above two scenarios this mean-reversion would only need to be very small to remedy our concerns.

Can we find any support for mean-reversion of risk-reversals in real market data? In order to search for such evidence we must first assume a form, or rather a process, for which we want to test. Again we here only search for indications of mean-reversion and do not worry very much about the level of significance. We look at spot prices and 1-month 25-delta risk-reversal prices for EURUSD as this data belongs to a very liquid market. Assume the risk-reversal price process to have mean-reversion and to be related to spot prices in the following simple discrete way:

$$\Delta RR = \alpha \Delta S + \beta (\overline{RR} - RR) \Delta t + \text{error},$$

where $\overline{RR}$ is the level of mean reversion and Δt represents a business day. For simplicity we assume the market to have roughly 250 business days in a calendar year, which gives us $\Delta t \approx 1/250$. In order to test the above assumption we divide the data set into sliding subgroups. The reason for not looking at the entire data set at once is that we deem it implausible for those model constants to stay truly constant over a longer period of many years. We here choose a group size of 250 market readings which again is roughly equivalent to one calendar year. For each window of market data we solve the above linear system using a least-square technique to estimate the parameters α , β and $\overline{RR}$, and plot the data over time.

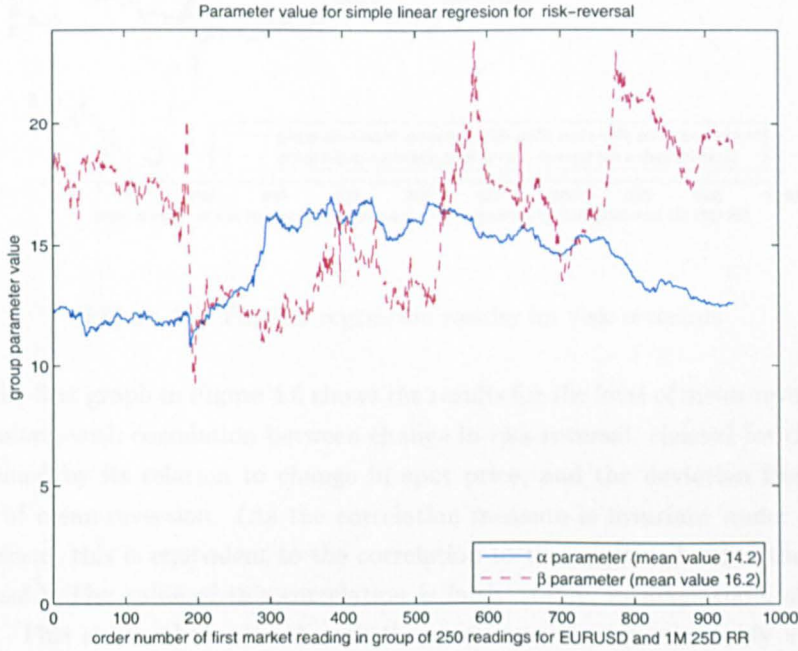


Figure 4.5: Risk-reversal regression showing stability in parameter magnitude

When plotting the results for the variables α and β we might not immediately see a clear pattern or a trend. But when we consider that the time series span about 4.5 years they seem, at least in magnitude, fairly stable over time. The daily change in risk-reversal is, on average, associated with a factor $\alpha \approx 15$ times the daily change of spot price. The mean-reversion per time-unit of the risk-reversal is associated to a factor $\beta \approx 15$ times the deviation from the level of mean reversion. This number for β would suggest that the 1-month 25-delta risk-reversal for the EURUSD would mean-revert with a half-life time of approximately 15 days ($\approx 365 \cdot \log(2) / \beta$).

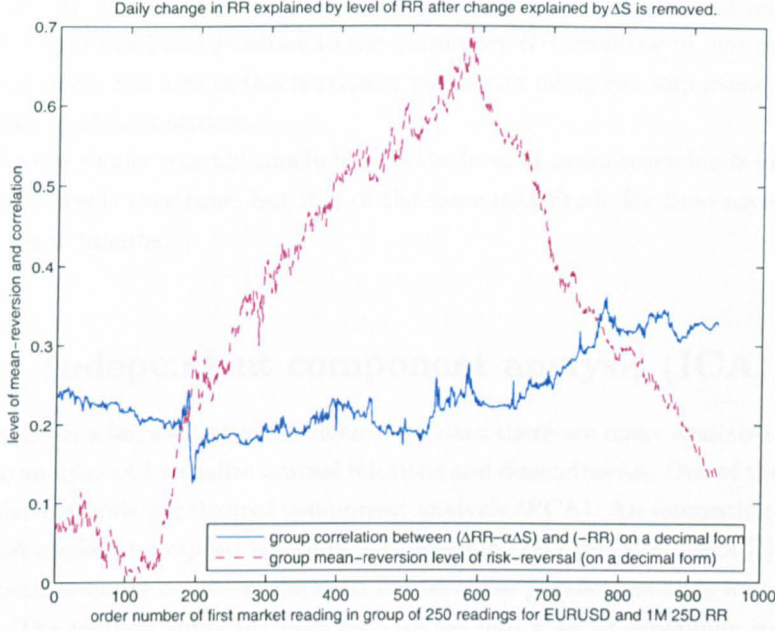


Figure 4.6: Further regression results for risk-reversals

The first graph in Figure 4.6 shows the results for the level of mean-reversion, $\overline{RR}$, along with correlation between change in risk-reversal, cleaned for change explained by its relation to change in spot price, and the deviation from the level of mean-reversion. (As the correlation measure is invariant under linear transform, this is equivalent to the correlation to the negative level of the risk-reversal.) The value of this correlation is fairly stable, in magnitude, around 25%. This shows that even though the mean-reversion is reasonably strong, there is a large part of the change in risk-reversals left to be explained by random noise, even after the correlation to changes in spot price have been removed.

The second graph in Figure 4.6, showing the calculated level of mean-reversion, $\overline{RR}$, is somewhat more troublesome. It is stable in neither level nor magnitude. A first interpretation would suggest that this parameter is essentially not constant and should perhaps also be modelled by a stochastic process. Support for this can be found in the simple observation that values of long-dated risk-reversals are not particularly constant. Another explanation is that with strong mean-reversion and much noise, the level of mean-reversion that will minimize the error will simply end up approximately at the mean value of the risk-reversal for the observed group. Looking at historical data of risk-reversal we see a higher level in the middle of the data set. One further aspect to consider is the fact that in our estimation we, in the term $\beta (\overline{RR} - RR) \Delta t$,

have $\beta \cdot \Delta t \approx 0.06$. This small value would make the linear system relatively badly conditioned and sensitive to the parameter $\overline{RR}$ resulting in, not only numerical noise, but also in this particular parameter being less important for the realisation of the outcome.

We can finally roughly conclude that the level of mean-reversion is not particularly stable over time, but it is of the same magnitude for time-scales of at least a few months.

4.4 Independent component analysis (ICA)

When given a large set of multidimensional data there are many available methods to analyze and visualize mutual relations and dependencies. One of the more popular methods is principal component analysis (PCA). An interesting study of PCA applied to implied volatility was presented by Cont & Fonseca [7]. This approach is based on the correlation between the parallel readings in a set of data. The method splits the original data set into a set of explaining multidimensional components onto which the original data is projected. The data can be recreated by the set of components, and their corresponding amplitude time series. To fully recreate the original data the number of components must be the same as the number of dimensions in the original data. Often very much of the observed data variance can be explained by fewer components than the original number of dimensions. These components are called “principal components” and this explains the origin of the name of the method.

One important characteristic of the PCA method is that all explaining components are mutually orthogonal in the original value space. The amplitude signals can however be highly correlated. This means that the results are not stable under linear transformation or even scaling of individual data series. The parallel data series need to be of the same magnitude and of the same form for the component separation to really make any sense from a dimensional perspective. In our case the data series are of different dimensions and often there are many different scales that are equally natural. As an example one can express volatility either on a decimal form or on a percentage form. Both forms will give a correct description but the components generated by PCA will be dependent on the choice of representing form.

A similar but more advanced method of component separation is the independent component analysis (ICA). The ICA method also tries to find a set of explaining components. Those components are, in contrast to PCA components, not necessarily mutually orthogonal in value space. The ICA procedure

4.4. INDEPENDENT COMPONENT ANALYSIS (ICA)

instead tries to find a set of explaining components with corresponding amplitude signals being as independent as possible.

Often when the dimensionality is high and one is looking for a relatively low number of explaining components, the PCA method can first be used to reduce the numbers of dimensions of the original data. In our case we must be careful with the PCA projection as our data set is on very different scales. Large nominal numbers can be assigned an unjustified high importance. A simple way of trying to get around this problem is to first normalize every dimension to a unit standard deviation. The explaining components will, after this normalisation, explain a change in each dimension relative to that dimension's original standard deviation. Without this normalisation PCA would only capture changes large in absolute terms and only capture very little of large relative changes if they happen to be small in absolute terms.

The actual mathematical procedure behind ICA can be very complicated and we have here used a free software package¹ for this component separation. This package is based on the MATLAB software and provides a user friendly way to separate data into independent components. For a more detailed description of the ICA method see work done by Hyvärinen & Oja [20] and [21].

The Figures 4.7-4.10 show the ICA components generated from normalized data projected onto the first five PCA components. The components are sorted in order of the kurtosis of the corresponding component's amplitude signal with the highest value first. In this case a low kurtosis means a more Gaussian-like signal behaviour and a high kurtosis means a higher degree of signal "burstiness". Data is given for the currency pairs EURUSD, USDJPY, GBPUSD and USDCHF.

Throughout the graphs we see a strong connection between changes in spot price and changes in risk-reversals. We see this as a high amplitude of both spot price and risk-reversals in the same component. This is especially true for the first or the second component in all four sets of components. Those components also show a smaller, but still significant, impact on implied ATM-volatility. The sign of this impact, relative to the sign of the spot price, is dependent on the average sign of the risk-reversal during the period of observation. It is this impact on implied ATM-volatility that is the very reason for the existence of risk-reversals in the first place.

This analysis of components is very interesting and is worth a chapter on its own. However we refrain from further analysis other than the observation that we again have found empirical evidence for an important connection between

¹MATLAB package from Helsinki University of Technology by Jarmo Hurri, Hugo Gävert, Jaakko Särelä, and Aapo Hyvärinen.

4.4. INDEPENDENT COMPONENT ANALYSIS (ICA)

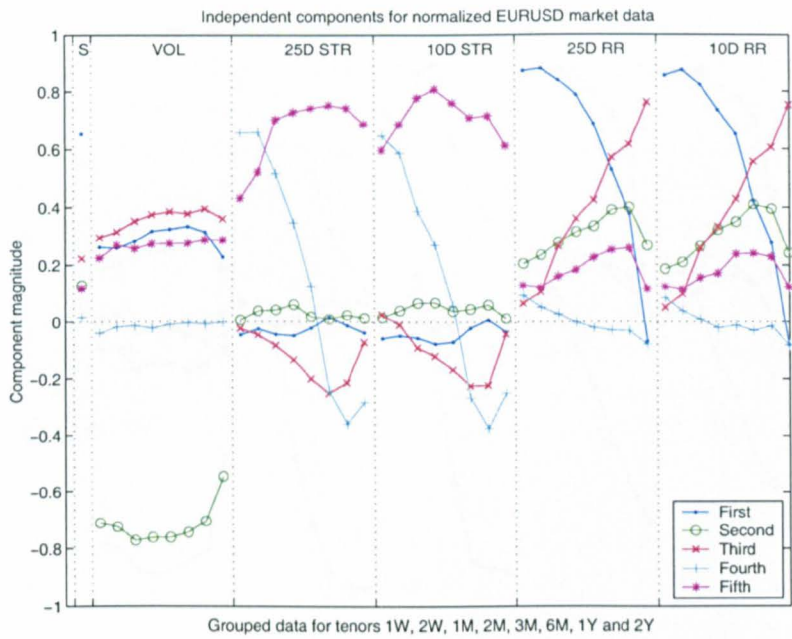


Figure 4.7: ICA components describing EURUSD dynamics

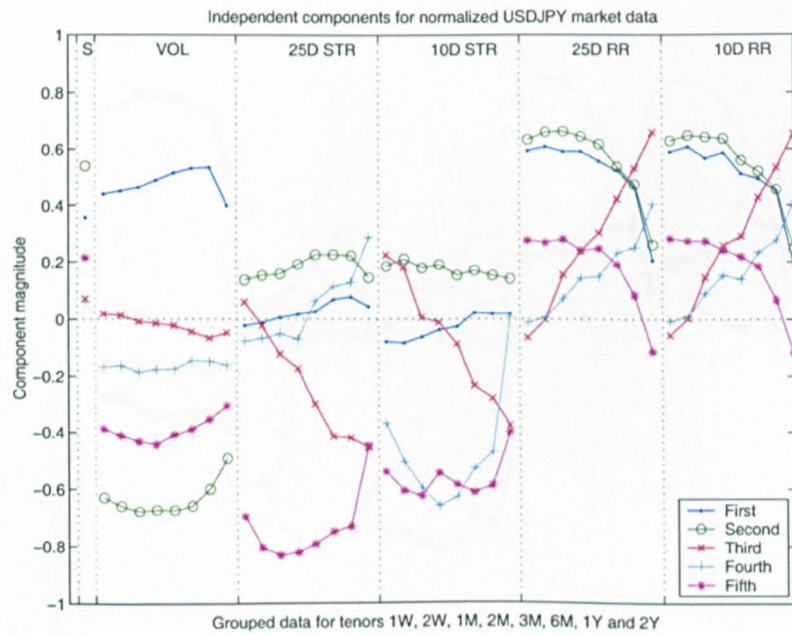


Figure 4.8: ICA components describing USDJY dynamics

4.4. INDEPENDENT COMPONENT ANALYSIS (ICA)

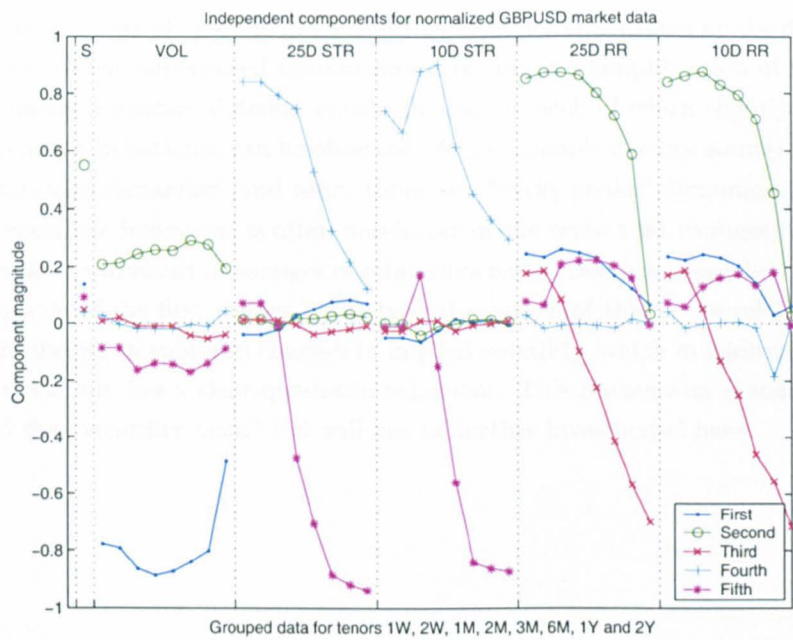


Figure 4.9: ICA components describing GBPUSD dynamics

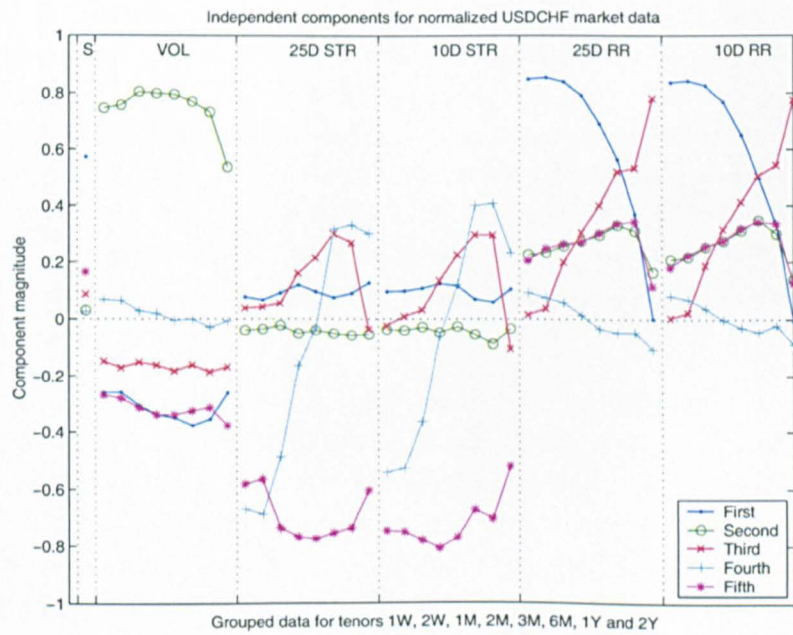


Figure 4.10: ICA components describing USDCHF dynamics

spot level, implied volatility and risk-reversals.

Finally we should keep in mind that this approach of breaking up the dynamics into linearly superposed components is of course a simplification of reality. The market witnesses different epochs in time, in each of which slightly different dynamic behaviours can be observed. As an example one can sometimes see “sticky delta dynamics” and other times see “sticky strike” dynamics. In the real world the behaviour is often non-linear in the sense that changes in some dimensions will result in changes of other dimensions that are dependent on the magnitude of the first change. One typical example of this is the relation between changes in spot and changes in implied volatility, which in addition to its linear relation, has a clear quadratic behaviour. This phenomena is sometimes called the “volatility bean” but will not be further investigated here.

Chapter 5

Modelling the market

We will here take a brief general look at different views of market modelling and pricing philosophy. A rough way of classifying the origin of models is suggested in Rebonato [37]. Models are here divided into “instrumental” and “fundamental” approaches.

A fundamental approach is based on looking at processes of underlying assets and trying to model what is observed. As long as a model for the underlying asset produces reasonable prices for traded instruments, the model that shows most similarities with the statistically observed market for the underlying asset should be chosen.

An instrumental approach is based on the reproduction of market properties of traded instruments derived from the underlying asset. The model used to model the underlying asset is rather seen as a tool to reproduce the sought properties of the traded instruments. We do not care so much about the similarities between the observation of the underlying asset and the process used to model it. Modelling the underlying asset with a process that differs much from the actual process can lead to problems when using the underlying asset to hedge an instrument.

If we were to describe the approach taken when we introduce the stochastic correlation model, it would probably fall mostly under the instrumental approach. It is the properties of the volatility surface that is the reason for extending the model. Despite this we also have a portion of the fundamental approach as we have seen that the suggested generalisation of the model also has support in empirical data. We here refer to the dynamics of the skew as an instrumental-style contribution and the connection between risk-reversals and the observed correlation between spot price and implied volatility as a fundamental-style contribution.

In general one can say that increasing the complexity of a market model, regardless of the approach, will give us more pricing flexibility. Such flexibility in pricing will allow us to better fit observed market prices. This will not necessarily give us any real contribution when it comes to hedging or pricing more complex options. If we model the market in a way that substantially differs from how the market really behaves, we risk losing the agreement between model and market prices as soon as the market changes in a dimension our model does not account for. In such a case not only pricing, but also hedging, would fail. This could be compared to changes in implied volatility in the Black-Scholes model.

5.1 Modelling the spot price

There is no real need to change the fundamental way we model the spot price. Apart from constant volatility the powerful log-normal model introduced by Black-Scholes [3] has been proved to satisfy most practical needs. Introducing stochastic volatility and stochastic correlation must be regarded as less of a deviance from the original model than a total change of the local return-distributions of spot price.

Even though we will not use this generalisation here, as a further development one could introduce market shocks by adding a jump term to the spot process.

5.2 Modelling stochastic volatility

Already when the Black-Scholes model was first introduced the assumption of a constant volatility was questioned. The market activity is not always the same and some periods in time are more turbulent than others. Periods with a very predictable impact on market activity, such as holidays and elections, can easily be taken into calculations when pricing options via the use of concepts like *trading days* or *business time* (see e.g. Chapter 4 in Hakala & Wystup [14]).

Even after the predictable component of non-constant volatility is removed, markets show a high degree of variation in volatility over time. The uncertainty of volatility is an integrated part of the everyday routine when trading options. In its most simple form, this risk is described by an option's so-called *vega*, which is the option's premium differentiated by implied volatility. As this parameter is assumed to be constant, hedging its risk actually means that we hedge outside

the model. By instead making volatility itself stochastic we move this risk inside our modelling framework.

Even though more modern methods like stochastic time-change exist today, the traditional approach to making volatility stochastic is to make the diffusion term in the spot price process a stochastic process itself. More correctly we say that we model the *instantaneous volatility*. This approach is not new and much work on the subject of stochastic volatility has already been published. One could here mention work by e.g. Wiggins [45], Hull & White [19] and Stein & Stein [42]).

One of the main reasons for the introduction of stochastic volatility was to account for the observed curvature of implied volatility, called the *volatility smile*. In mathematical terms the smile is a result of excess kurtosis, or so called fat-tails, in the distributions for returns of the underlying asset. The various implementations of stochastic volatility took many forms but they were all successful in producing both kurtosis and volatility smiles.

5.2.1 Different model processes

A problem when choosing a process to model the stochastic volatility is that the real process is not observable. It is a so called *hidden process*. The only observable process that is connected to volatility is the spot price. Theoretically we would be able to study high-frequency data and for small time periods calculate volatility and that way show the progress of volatility over time. Unfortunately this has been proved harder to do in reality and we will have to find another way to make our choice of process.

We could also try to make use of the market's aggregated view of volatility by looking at implied volatility. Implied volatility takes more aspects than just volatility into account. Such aspects could be attitude towards risk and market supply and demand. Even though implied volatility is not the same as volatility there is most likely a close relationship.

In order for the stochastic correlation framework to make any sense, the instantaneous volatility process must be described by a diffusion process. If not so, the concept of a correlation between driving Wiener-processes would fail. Before we try to choose a particular process to model the stochastic volatility we will look at some of the more popular diffusion models used today.

Ornstein-Uhlenbeck process

An Ornstein-Uhlenbeck process is a simple process suggested among others by Scott [41]. It is a simple process but it still shows important properties such as mean-reversion. Without mean-reversion volatility could reach implausible levels. The immediate drawback is the possibility of the process to take negative values. Negative values would be allowed mathematically but it would not make any real sense. A negative volatility would be equivalent to a change in sign of the correlation to spot price. This would give us less control when modelling stochastic correlation. Mathematically the SDE of the Ornstein-Uhlenbeck process is given by

$$d\sigma_{OU} = \alpha_{OU} (m_{OU} - \sigma_{OU}) dt + \xi_{OU} dW.$$

Logarithmic Ornstein-Uhlenbeck process

The pure Ornstein-Uhlenbeck process had the unwanted property of being able to reach negative values. By taking the exponential value of the process we can ensure strictly positive values and still keep the mean-reversion property. Mathematically the SDE of the Logarithmic Ornstein-Uhlenbeck process is given by

$$\begin{aligned} \sigma_{\text{LogOU}} &= e^{y_{\text{LogOU}}} \\ dy_{\text{LogOU}} &= \alpha_{\text{LogOU}} (m_{\text{LogOU}} - y_{\text{LogOU}}) dt + \xi_{\text{LogOU}} dW. \end{aligned}$$

The mapping from the Ornstein-Uhlenbeck process onto the exponential function has a positive curvature. Due to the Jensen inequality an increase in volatility of volatility, ξ_{LogOU} , will lead to an increase in the average level of volatility (see Appendix B in Øksendal [32]) Later we will see this effect when we study the impact on option premia of this parameter in a stochastic-correlation model. Apart from the expected increase in kurtosis, we will also see a general increase in the option premium similar to that which we get if we increase the reversion level m_{LogOU} .

CIR/Feller/Heston square root process

The CIR/Feller/Heston square root process has been shown to be attractive from an analytic perspective. It was first introduced by Cox, Ingersoll and Ross [8] but was probably mostly made famous by Heston [17] for his use of the model to reach a semi-closed formula for the price of vanilla options. The main drawback is that the process, for some parameter values, can reach zero.

5.2. MODELLING STOCHASTIC VOLATILITY

Unfortunately values needed to reproduce observed market prices often fall under this category. In itself a volatility very close to zero is not implausible and this could e.g. naturally happen during a longer holiday. What is implausible is that a volatility close to zero would exist after known effects, like holidays, are accounted for. Using the earlier mentioned concept of business time, the price process must always have a non-zero market activity.

Even if the square root model is analytically attractive it is numerically unattractive. The sharp change in volatility of the process close to the zero-level can make it numerically problematic. This model is actually not even Lipschitz but can still be shown to have a unique solution. Mathematically the SDE of the square-root process is given by

$$\begin{aligned}\sigma_{\text{Hstn}} &= \sqrt{y_{\text{Hstn}}} \\ dy_{\text{Hstn}} &= \alpha_{\text{Hstn}} (m_{\text{Hstn}} - y_{\text{Hstn}}) dt + \xi_{\text{Hstn}} \sqrt{y_{\text{Hstn}}} dW\end{aligned}$$

The mapping from the Ornstein-Uhlenbeck process onto the square-root function has a negative curvature. Due to the Jensen inequality an increase in volatility of volatility, ξ_{Hstn} , will lead to a decrease in the average level of volatility. One can see this phenomenon when plotting the impact on vanilla option premia when changing this parameter against strike price (see Section 3 in Heston [17]). Apart from the expected increase in kurtosis, we will also see a general decrease in option premium similar to the case in which we decrease the reversion level m_{Hstn} .

5.2.2 The freedom to choose a simple model

As instantaneous volatility is a hidden process it can be hard to make a choice about which model to use. Of course the chosen process must exhibit plausible and sensible behaviour, but apart from that we need other criteria to choose a model. We could either choose a model because of its analytic properties, as is often the case when the square root process is chosen, or we could instead make a choice from a numerical standpoint. It is by taking a numerical viewpoint that we have chosen to use the logarithmic Ornstein-Uhlenbeck process to represent instantaneous volatility in the remainder of this thesis.

We will show that as long as one of the above three mean-reverting diffusion process is chosen and volatility of volatility is reasonably small the actual choice of model is not very important. This can be done by choosing the free parameters in the three models so as to give the different processes

approximately the same local properties. By using Itô's formula to get non-transformed processes and then using the Taylor-expansion we can turn all three processes into pure Ornstein-Uhlenbeck processes. If we use the original Ornstein-Uhlenbeck process as a benchmark we can then choose the free parameters of the transformed processes to replicate the benchmark process. Under those choices of equivalent parameters the three process behave very similarly.

Equivalent parameters (Logarithmic Ornstein-Uhlenbeck process)

We use Itô's formula to get the SDE of the transformed process.

$$\begin{aligned}
 d\sigma_{\text{LogOU}} &= \sigma_{\text{LogOU}} \left(\alpha_{\text{LogOU}} (m_{\text{LogOU}} - \log(\sigma_{\text{LogOU}})) + \frac{1}{2} \xi_{\text{LogOU}}^2 \right) dt + \\
 &\quad + \sigma_{\text{LogOU}} \xi_{\text{LogOU}} dW \\
 &\approx \left(\alpha_{\text{LogOU}} (1 - m_{\text{LogOU}} + \log(m_{\text{OU}})) - \frac{1}{2} \xi_{\text{LogOU}}^2 \right) \times \\
 &\quad \times \left(\frac{\alpha_{\text{LogOU}} m_{\text{OU}}}{\alpha_{\text{LogOU}} (1 - m_{\text{LogOU}} + \log(m_{\text{OU}})) - \frac{1}{2} \xi_{\text{LogOU}}^2} - \sigma_{\text{LogOU}} \right) dt + \\
 &\quad + m_{\text{OU}} \xi_{\text{LogOU}} dW.
 \end{aligned}$$

Making the particular choices

$$\begin{aligned}
 \alpha_{\text{LogOU}} &= \alpha_{\text{OU}} \\
 m_{\text{LogOU}} &= \log(m_{\text{OU}}) - \frac{\xi_{\text{OU}}^2}{2\alpha_{\text{OU}} m_{\text{OU}}^2} \\
 \xi_{\text{LogOU}} &= \frac{\xi_{\text{OU}}}{m_{\text{OU}}},
 \end{aligned}$$

results in a first order equivalence to our benchmark process

$$d\sigma_{\text{LogOU}} \approx \alpha_{\text{OU}} (m_{\text{OU}} - \sigma_{\text{OU}}) dt + \xi_{\text{OU}} dW = d\sigma_{\text{OU}}.$$

In this case it is of interest to notice that we approximate the multiplier of the stochastic driver ($\sigma_{\text{LogOU}} \xi_{\text{LogOU}}$) by the constant ($m_{\text{OU}} \xi_{\text{LogOU}}$). The total process approximation is therefore slightly inaccurate in both the driving term and in the drift term.

Equivalent parameters (CIR/Feller/Heston square root process)

Similarly, we use Itô's formula to get the SDE of the transformed process

$$\begin{aligned} d\sigma_{\text{Hstn}} &= \frac{4\alpha_{\text{Hstn}} (m_{\text{Hstn}} - \sigma_{\text{Hstn}}^2) - \xi_{\text{Hstn}}^2}{8\sigma_{\text{Hstn}}} dt + \frac{1}{2}\xi_{\text{Hstn}} dW \\ &\approx \frac{4\alpha_{\text{Hstn}} (m_{\text{Hstn}} + \sigma_{\text{OU}}^2) - \xi_{\text{Hstn}}^2}{8\sigma_{\text{OU}}^2} \left(\frac{2\sigma_{\text{OU}} (4\alpha_{\text{Hstn}} m_{\text{Hstn}} - \xi_{\text{Hstn}}^2)}{4\alpha_{\text{Hstn}} (m_{\text{Hstn}} + \sigma_{\text{OU}}^2) - \xi_{\text{Hstn}}^2} - \sigma_{\text{Hstn}} \right) dt + \\ &\quad + \frac{1}{2}\xi_{\text{Hstn}} dW. \end{aligned}$$

Making the particular choices

$$\begin{aligned} \alpha_{\text{Hstn}} &= \alpha_{\text{OU}} \\ m_{\text{Hstn}} &= m_{\text{OU}}^2 + \frac{\xi_{\text{OU}}^2}{\alpha_{\text{OU}}} \\ \xi_{\text{Hstn}} &= 2\xi_{\text{OU}}, \end{aligned}$$

results in a first order equivalence to our benchmark process

$$d\sigma_{\text{Hstn}} \approx \alpha_{\text{OU}} (m_{\text{OU}} - \sigma_{\text{OU}}) dt + \xi_{\text{OU}} dW = d\sigma_{\text{OU}}.$$

This approximating process is slightly inaccurate in the drift term but is exact in the stochastic driving term.

Simulating the equivalent processes

To really see the similarities in the original processes we show a simulation of them all in parallel. We choose the model parameters of the benchmark Ornstein-Uhlenbeck process to $m_{\text{OU}} = 10.0\%$, $\alpha_{\text{OU}} = 0.3$ and $\xi_{\text{OU}} = 0.05$. The equivalent parameter values for the two transformed processes are calculated as described above and all three simulations use the same driving Wiener process.

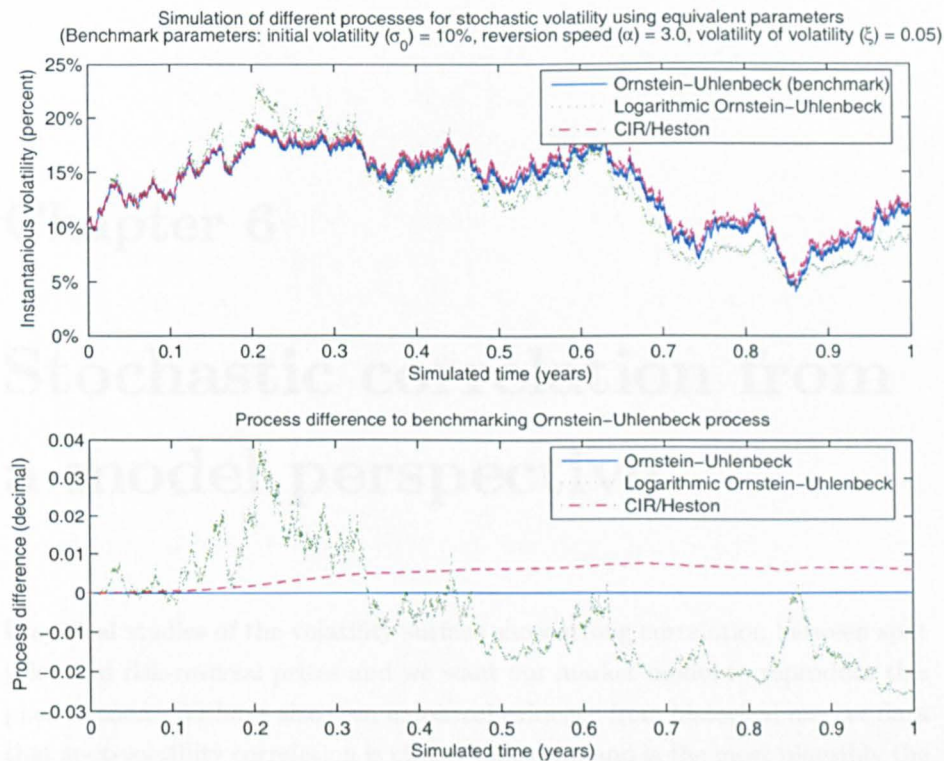


Figure 5.1: Simulation of three models for stochastic volatility

We see that during the period of one year the differences between the three processes stays within only a few percent of the absolute level. This difference is negligible in comparison to the change in value caused by the stochastic process itself.

Chapter 6

Stochastic correlation from a model perspective

Empirical studies of the volatility surface show strong correlation between spot price and risk-reversal prices and we want our market model to reproduce this phenomenon. We have also seen empirical evidence from historical market data that spot-volatility correlation is closely linked to, and is the most plausibly the explanation for, the existence of risk-reversals.

Another approach is to look at existing models. As we assume volatility to be stochastic we need a stochastic volatility model. What impact does spot-volatility correlation have on such a model? Because of this we now take a look at the well known Heston stochastic volatility model [17]. This model is chosen mainly because it has a closed, or at least semi-closed, form solution for the price of European vanilla options. This will very much reduce the computational effort of our numerical investigations. The way the market is modelled in this model is mostly a choice to allow for a closed form pricing formula and the realism of this model is debated.

The Heston model has its origin in the Black-Scholes model but adds an extra process for instantaneous volatility. This process introduces five new model parameters and leaves out the volatility parameter in the original Black-Scholes model. The new parameters are: initial instantaneous variance, long-term instantaneous variance, variance reversion-speed, volatility of instantaneous variance and the correlation between the driving Wiener processes of spot price and variance. By changing the model parameters we also change the premia of options priced under the model. This, in turn, changes the implied volatility smile generated by the model. To better understand the effects of the different parameters on the implied volatility smile one can plot the differential of the

smile with respect to the model parameters against strike price or option delta. Again this is a topic worth a section of its own but is here left out.

Briefly one can say that the initial instantaneous variance determines the short term implied volatility and the long-term mean instantaneous variance determines the long-term implied volatility. What expiries are considered short- or long-term is, in the Heston model, governed by the speed of mean-reversion. Changing volatility of instantaneous variance changes the convexity of the implied volatility smile. The last, and for our purposes the most important parameter, is the spot-volatility correlation. By changing the correlation between the two driving Wiener processes we can change the skew or overall slope of the volatility smile.

6.1 Heston model fitted to market data

We already know that the market quantifies volatility convexity by strangles and skew by risk-reversals. But how do those quotes relate to the model parameters of a stochastic volatility model like the Heston model?

To find an answer to this question we can fit the parameters of the model to historic market data. By doing so we find a set of parameters that reproduce the observed market prices. We can then plot those parameters against the real market data to get see if there is an obvious relationship.

The fitting is only done for one single tenor – the 1M tenor. Of course we could try to fit the model to all available data, but as we are mainly interested in the effects of the correlation parameter we try not to stretch the fitting procedure more than necessary. The actual fitting is done by numerically searching for model parameters that will minimize the sum of squared errors between the implied volatility of the model and historic market data. For every market reading the minimisation procedure starts at the same initial guess for the model parameters. As a result of the fitting only being done for one single tenor the parameters related to term structure, e.g. long-term mean variance and variance reversion speed, does not have a large impact on the fitting. In fact those parameters end up at having roughly the same values as the initial values throughout the fitting of the entire set of historical data. In Figure 6.1 we see a graph of the Heston parameters fitted to historical data for EURUSD.

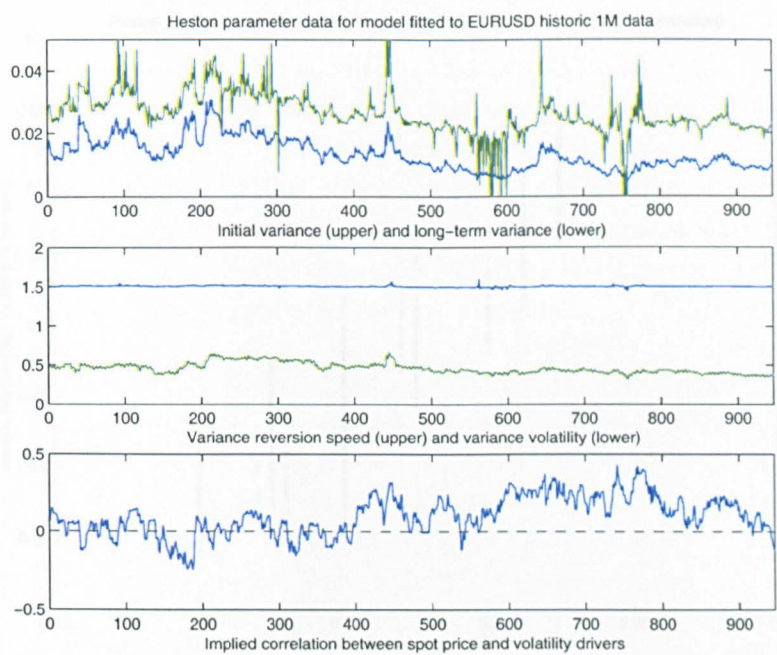


Figure 6.1: Parameter time series for market calibrated Heston model

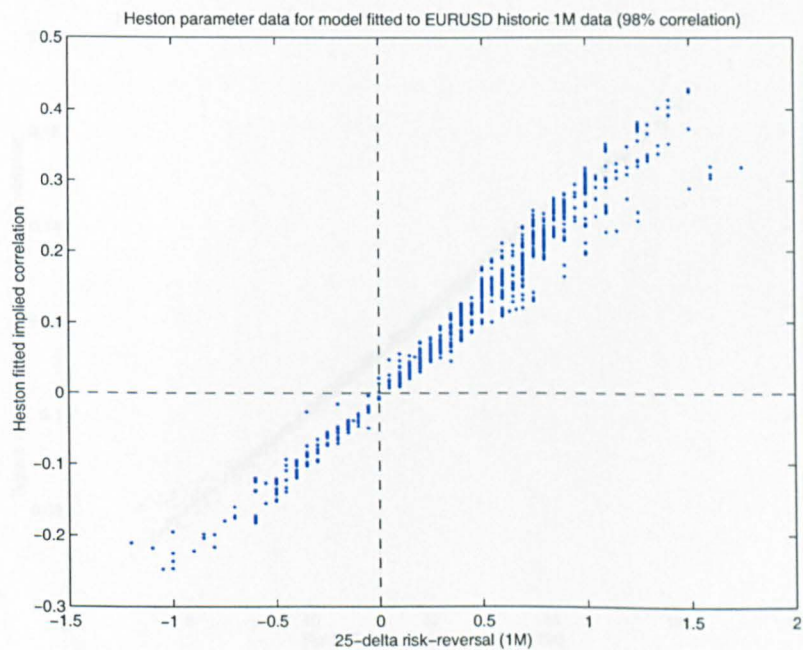


Figure 6.2: Risk-reversal and calibrated Heston correlation

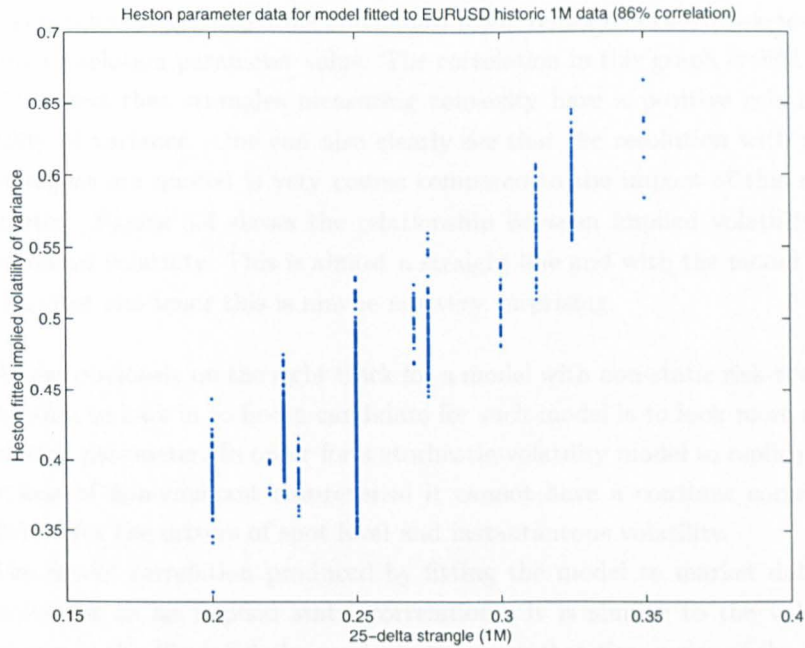


Figure 6.3: Strangle and calibrated Heston volatility of variance

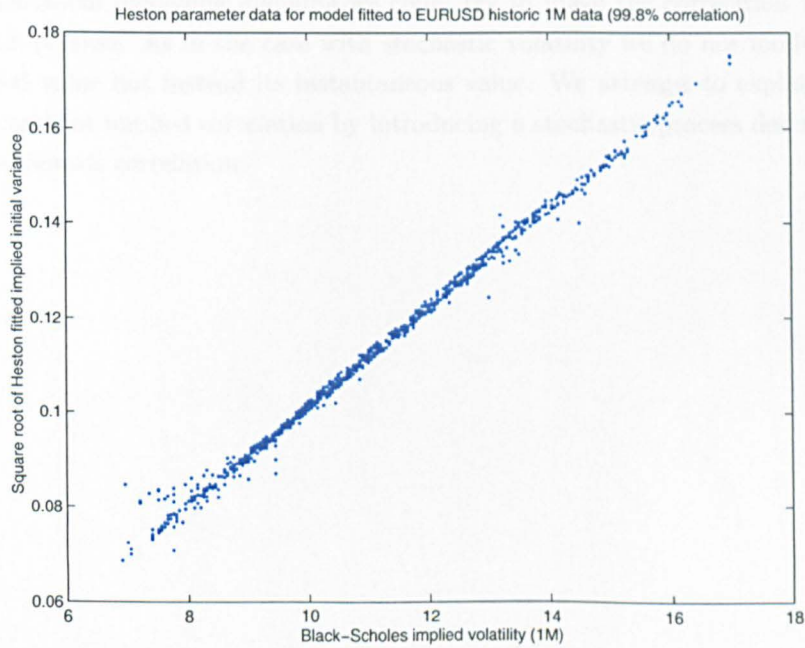


Figure 6.4: Implied volatility and square root of calibrated Heston initial variance

Figure 6.2 shows a staggering and almost linear relation between risk-reversals and the correlation parameter value. The correlation in this graph is 98%. Figure 6.3 shows that strangles measuring convexity have a positive relation to volatility of variance. One can also clearly see that the resolution with which the strangles are quoted is very coarse compared to the impact of this model parameter. Figure 6.4 shows the relationship between implied volatility and model initial volatility. This is almost a straight line and with the model being fitted to just one tenor this is maybe not very surprising.

We are obviously on the right track for a model with non-static risk-reversal. A direction to look in to find a candidate for such model is to look more at the correlation parameter. In order for a stochastic-volatility model to replicate the behaviour of non-constant risk-reversal it cannot have a constant correlation coefficient for the drivers of spot level and instantaneous volatility.

The model correlation produced by fitting the model to market data can be looked at as an implied static correlation. It is similar to the volatility parameter in the Black-Scholes model in the sense that time series of the model parameter, fitted to observed market prices, is not constant even though the underlying model assumes the parameter to be so.

In the same way that a stochastic-volatility model attempts to solve this mathematical modelling dilemma we could try to make the correlation a stochastic process. As in the case with stochastic volatility we do not model the implied value but instead its instantaneous value. We attempt to explain the non-constant implied correlation by introducing a stochastic process describing instantaneous correlation.

Chapter 7

Introducing the model

7.1 What we try to explain and quantify with the model

Earlier we explained that non-constant risk-reversals affect the prices of options. The various ways to account for this phenomenon only give a very coarse picture of the quantitative impact on option prices. With the stochastic correlation model, or rather a whole class of models, we want to create a tool to better quantify this effect. We wish to do this even when the payoff structure is very complicated and the methods used today do not really tell us how to approach such quantification. As also mentioned above, this rather means finding a model generating conditional distributions more in agreement with observed market prices.

A more fundamental question is to ask *why* the skew is not constant. The stochastic correlation is obviously just a mathematical ad-hoc answer and no fundamental explanation. We must dig deeper into the underlying economics and market psychology to explain why we see this kind of dynamics. This is left for further investigations.

7.2 The model

So far we have only argued for the introduction of stochastic correlation in general terms. To approach the idea of stochastic correlation further we must choose a model to represent our market and thus also our correlation. We will describe this model by a stochastic differential equation (SDE). The system is initially chosen to be a Markov diffusion process with coefficients invariant in

time but non-constant in state level.

To ensure that our general modelling framework is well defined and only take real value numbers our correlation process must only take values in the interval $[-1, +1]$. Without this fundamental restriction our stochastic framework cannot have strong solutions as the market model would get a non-defined behaviour with a non-zero probability. Apart from this important constraint we want to make our process as simple as possible.

7.3 Drift and diffusion of correlation process

The drift term of a diffusion process represents the deterministic behaviour of the process. Looking at historical data does not immediately reveal any obvious choice for the drift. We can however see a weak negative correlation between the change in risk-reversals and the level of risk-reversals. This negative correlation is present in data for all observed currency pairs. For shorter maturities (1W-2M) this correlation is about -25% and for longer maturities (1Y-2Y) this correlation is slightly lower (-15%). We have omitted the thorough investigation and settle with the conclusion that there is, however probably weak, mean-reversion in the risk-reversal data. Even though there is a strong relation between risk-reversals and spot-implied volatility correlation, we must not forget that risk-reversals are not equal to instantaneous spot-implied volatility correlation. But in a stochastic-correlation model our correlation process is our only way to control the risk-reversals. To give our market model the above observed mean-reversion we choose to give the correlation process the mean-reversion property as well. From a more theoretical view point there is also a valid reason for the introduction of mean-reversion. Remember that spot price and risk-reversals have a strong correlation. Without a mean reversion the risk-reversals would reach absurd levels every time there was a persistent trend in the spot price. Mean-reversion would allow trends but still make it possible to restrain risk-reversals to moderate levels.

As we know little more about the mean-reversion of the process than the very existence of it, there is no reason not to choose the simplest possible form for this reversion. A very simple form is a mean-reversion linearly proportional to the distance from the reversion-level like

$$\mu_\rho = \gamma(n - \rho_t),$$

where γ is the rate of mean reversion and n is the mean-reversion level.

The second property of a diffusion process is the magnitude of the diffusion

term, also known as volatility. This is a measure of random fluctuations of the process at a specific level. As for the drift, historical market data give few indications of what a suitable choice of diffusion would be. Again we use a more theoretical argument. We know that our correlation process must stay in the admissible interval $[-1, +1]$. As our choice of drift does not take care of this problem we should look for a remedy in the diffusion term. In a Black-Scholes world the spot price process will, with almost sure probability, never reach the zero level due to the linear decrease to zero of the volatility then the process approaching this level. We can in the same way let the volatility of our correlation process linearly go to zero as the process approaches the “no-go” regions. In our case this would mean zero volatility at -1 and $+1$. A simple non-trivial continuous function satisfying this volatility “cut-off” constraint is the quadratic polynomial

$$\sigma_\rho = \psi(1 - \rho_t^2),$$

where ψ is the volatility factor of the correlation process.

There is also a far-fetched interpretation of this choice of diffusion. It has its origin in statistical estimation based on limited information. When given a finite number of normally distributed pairs of observations we can calculate the sample correlation. As our available information is only a finite set of samples from a true relation between our observed data we can get nothing but an estimate of the true correlation. It can be shown that the standard deviation of this estimate is approximately proportional to $1 - \rho^2$ where ρ is the true value of the correlation. The far-fetched argument is that if the market only has a limited set of information to base the taken actions on, the choice of actions will be less uncertain and hence less volatile, when correlation is close to its extreme values. In contrast the uncertainty of the correlation estimation reaches its maximum when the true correlation is very close to zero.

When putting it all together we get the SDE with which we will model our correlation process.

$$\begin{aligned} d\rho_t &= \gamma(n - \rho_t) dt + \psi(1 - \rho_t^2) dW_t \\ \rho_0 &\in (-1, +1) \\ t &\in [0, T] \\ n &\in [-1, +1]. \end{aligned} \tag{7.1}$$

Here W_t is a standard Wiener process. We now need to prove that this function will not leave the admissible region. In fact we will be able to show that as long as $\rho_0 \in (-1, +1)$ the process, with probability 1, will not reach the boundaries

even without mean reversion.

We directly get into problems as the function does only satisfy the Lipschitz conditions (see more in Rogers & Williams [38]) within the region $[-1, 1]$. If the solution leaves this region we cannot use the standard arguments to guarantee its existence and uniqueness. We will prove that if we can show that $\rho_t \in (-1, +1), \forall t \in [0, T]$ then the process both exists and is unique.

7.4 Existence and uniqueness

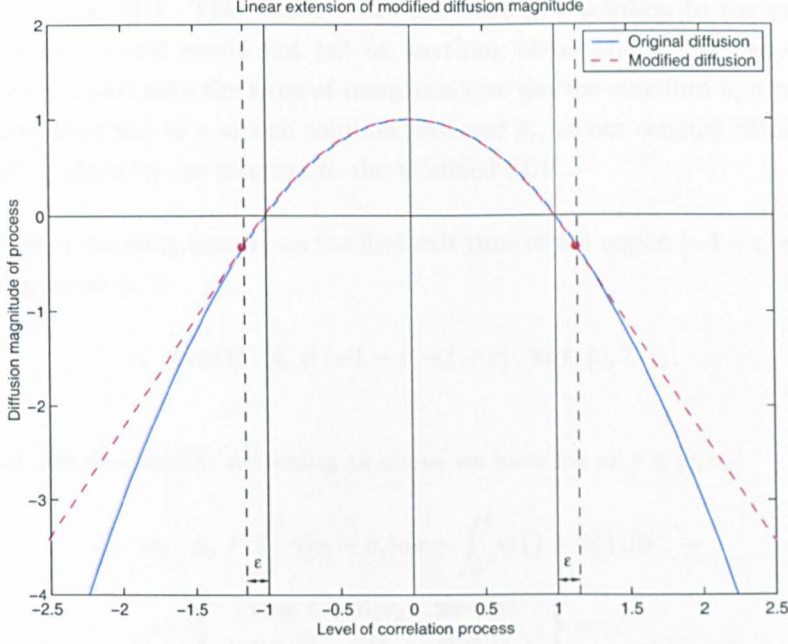
As mentioned above the problem lies in the quadratic nature of the diffusion term in (7.1). To directly be able to use standard arguments and ensure existence and uniqueness of a solution both the drift term and the diffusion term need to satisfy the Lipschitz conditions. The drift term is of a linear nature and causes no problem. The diffusion term however does not satisfy the required growth conditions and makes the SDE non-Lipschitz. (The same growth conditions prohibits the use of the more relaxed Yamada-Watanabe theorem to prove uniqueness.) Fortunately we will be able to show that the process is confined in the region $[-1, +1]$ and that this allows us to avoid the problem and still be able to use the Lipschitz framework.

We start by looking at a version of our original SDE slightly modified to satisfy the Lipschitz conditions. Define this new modified SDE as

$$\begin{aligned} d\rho_t^{\text{mod}} &= \gamma (n - \rho_t^{\text{mod}}) dt + \sigma^{\text{mod}}(\rho_t^{\text{mod}}) dW_t \\ \sigma^{\text{mod}}(\rho_t^{\text{mod}}) &= (1 - H(|\rho_t^{\text{mod}}| - 1 - \epsilon)) \psi(1 - (\rho_t^{\text{mod}})^2) + \\ &\quad + H(|\rho_t^{\text{mod}}| - 1 - \epsilon)(2(1 + \epsilon)(1 - |\rho_t^{\text{mod}}|) + \epsilon^2)\psi. \end{aligned}$$

Here $\epsilon > 0$ and $H(\cdot)$ means the Heaviside step function, which is used to define different behaviours for the diffusion term depending on whether or not the process is inside or outside the region $[-1 - \epsilon, +1 + \epsilon]$. When the process is inside the region the diffusion term is identical to the original diffusion term. Outside the region the diffusion term has been modified to a smoothly connected linear extrapolation of the original diffusion term.

It is of technical interest to notice that this modified diffusion is everywhere differentiable and is twice differentiable everywhere but in a finite number of points. The reduced differentiability in those finite points causes no problem for the Itô calculus (see Øksendal [32] p. 57).


 Figure 7.1: Diffusion term and its modification outside $[-1 - \epsilon, +1 + \epsilon]$

The modified SDE has an asymptotically linear, instead of quadratic, behaviour and therefore satisfies the Lipschitz condition. For the modified SDE we can now use standard arguments (see Roger & Williams [38]) to guarantee both existence and uniqueness of a solution. Denote this unique solution of our modified SDE as ρ_t^{mod} .

This solution to our modified SDE is not necessarily a solution to our original SDE as the diffusion term is not the same outside the region $[-1 - \epsilon, +1 + \epsilon]$. But if we could show that our process ρ_t^{mod} stays within this region the diffusion terms would be identical for all values of the process and ρ_t^{mod} would also be a solution to our original SDE. Let us for a moment assume that $\rho_t^{\text{mod}} \in [-1, +1]$ for $\forall t \in [0, T]$ almost surely or equivalently, with probability 1. Under this assumption we have

$$\begin{aligned} \rho_t^{\text{mod}} &= \rho_0 + \int_0^t \gamma(n - \rho_s^{\text{mod}}) ds + \int_0^t \sigma^{\text{mod}}(\rho_s^{\text{mod}}) dW_s = \\ &= \left\{ \begin{array}{l} \text{By assumption } \rho_t^{\text{mod}} \in [-1, +1] \text{ and} \\ \text{hence } \sigma^{\text{mod}}(\rho_t^{\text{mod}}) = \psi(1 - (\rho_t^{\text{mod}})^2) \end{array} \right\} = \\ &= \rho_0 + \int_0^t \gamma(n - \rho_s^{\text{mod}}) ds + \int_0^t \psi(1 - (\rho_s^{\text{mod}})^2) dW_s. \end{aligned}$$

We see above that under our assumption, ρ_t^{mod} would be a solution also

7.4. EXISTENCE AND UNIQUENESS

to our original SDE. This would prove existence of a solution to the original problem but it still would not tell us anything about the uniqueness of the solution. To deal with the issue of uniqueness we use the standard approach of assuming existence of a second solution, denoted $\tilde{\rho}_t$, to our original SDE apart from ρ_t^{mod} given by the solution to the modified SDE.

Define a stopping time τ_ϵ as the first exit time of the region $[-1 - \epsilon, +1 + \epsilon]$ for the process $\tilde{\rho}_t$

$$\tau_\epsilon \triangleq \inf (t : \tilde{\rho}_t \notin [-1 - \epsilon, +1 + \epsilon], \forall t \in [0, T]).$$

Now follow a similar reasoning as above we have for all $t \in [0, \tau_\epsilon]$

$$\begin{aligned} \tilde{\rho}_t &= \rho_0 + \int_0^t \gamma(n - \tilde{\rho}_s) ds + \int_0^t \psi(1 - \tilde{\rho}_s^2) dW_s = \\ &= \left\{ \begin{array}{l} \text{Using } t \in [0, \tau_\epsilon] \text{ we also} \\ \text{get } \psi(1 - \tilde{\rho}_s^2) = \sigma^{\text{mod}}(\tilde{\rho}_s) \end{array} \right\} = \\ &= \rho_0 + \int_0^t \gamma(n - \tilde{\rho}_s) ds + \int_0^t \sigma^{\text{mod}}(\tilde{\rho}_s) dW_s = \\ &= \left\{ \begin{array}{l} \text{The Lipschitz condition} \\ \text{guarantees a unique solution.} \end{array} \right\} = \rho_t^{\text{mod}}. \end{aligned}$$

We have here used the fact that as long as the process $\tilde{\rho}_t$ stays within the region it is also a solution to our modified SDE. But the modified SDE has only the unique solution ρ_t^{mod} so hence $\tilde{\rho}_t$ must be identical to this solution for $\forall t \in [0, \tau_\epsilon]$. This, together with the very definition of τ_ϵ , leads to

$$\left. \begin{array}{l} \tilde{\rho}_t = \rho_t^{\text{mod}}, \forall t \in [0, \tau_\epsilon] \text{ choosing } t = \tau_\epsilon \\ \tilde{\rho}_{\tau_\epsilon} \in \mathbb{R} \setminus (-1 - \epsilon, +1 + \epsilon) \end{array} \right\} \implies \rho_{\tau_\epsilon}^{\text{mod}} \in \mathbb{R} \setminus (-1 - \epsilon, +1 + \epsilon).$$

Using our initial assumption $\rho_t^{\text{mod}} \in [-1, +1]$ for $\forall t \in [0, T]$ almost surely we can now conclude for the stopping time τ_ϵ

$$\left. \begin{array}{l} \rho_{\tau_\epsilon}^{\text{mod}} \in \mathbb{R} \setminus (-1 - \epsilon, +1 + \epsilon) \\ \mathbb{P}[\rho_t^{\text{mod}} \in [-1, +1], \forall t \in [0, T]] = 1 \end{array} \right\} \implies \mathbb{P}[\tau_\epsilon > T] = 1.$$

Hence ρ_t^{mod} is also a unique solution to our original SDE for $\forall t \in [0, T]$. Left to show is that the solution to the modified SDE, ρ_t^{mod} , and hence via the above arguing also the solution to the original SDE, will stay in the region $(-1, +1)$.

7.5 Admissible values of correlation process

In a previous section we defined the model governing the instantaneous correlation. Of course this is not the only possible choice but it suits our purposes and agrees with the little empirical support available. Regardless of the actual choice of model a correlation process must stay in the closed region $[-1, +1]$ in order to make any sense. Intuitively the chosen process will stay in the admissible interval. When the process reaches the region's endpoints, the amplitude of the driving Wiener process gets cut off while the mean-reversion reaches its maximum amplitude and dominates the dynamics of the process. Intuition is however no more than a qualified guess, but it is no proof. Another hint that the chosen process actually will be contained within the admissible area is that with a volatility expressed as a power function of the state value, the power needs to be lower than $1/2$ in order for the process to reach the zero level with a non-zero probability (see Jäckel [23] p. 37). Our choice of volatility approaches a linear behaviour at the boundary points and will reach zero at the very boundaries. This gives a local behaviour very similar to that of the regular Black-Scholes spot model close to zero.

A further motivator for the assumption of an appropriate behaviour of this process, is the empirical results we get when numerically simulating the process using an Euler-scheme. With a given time span to simulate the process and a given size of the time steps, such simulation eventually will evolve to a state outside the admissible region $[-1, +1]$ with a non-zero probability. But if, when this happens and the simulation leaves the region, we refine our driving Wiener-process (possibly more than once) using a Lévy-construction, we are always able to find a smaller step size that will make our simulation stay inside the admissible region for the entire time span. Further, these types of simulation show that with a moderate time span, $T = 1.0$, the process will stay in the admissible region even for very large volatilities, $\psi = 10.0$, and no mean reversion at all, $\gamma = 0.0$. Such extreme values are however not plausible for simulating a real-life correlation as those values will cause the process to immediately go very close to one of the two borders and then stay very close to it for the rest of the simulation without actually touching the border.

To mathematically show that the process ρ_t^{mod} will only take admissible values we start by making the process discrete in time but still continuous in space. More precisely this is done using the Euler-scheme approximation. This scheme is also known as the Euler-Maruyama approximation and it converges to our solution strongly and uniformly as the step size is reduced. (see Theorem 10.2.2 p. 342 and Theorem 10.6.3 p. 361 in Kloeden & Platen [25]). By using simple linear interpolation between discrete points each approximation in the

series of processes can be turned into a t -continuous process.

We also note that every Itô-integral has a t -continuous version or modification. Generally we think of this version when we speak of stochastic integrals (see Øksendal p. 32 [32]). This gives us that the solution to our modified SDE can also be chosen to be t -continuous.

We will show that for each step in the approximating processes the probability of ending up outside the admissible set, when making the time steps smaller, goes to zero fast enough for the entire approximating process to stay in this set regardless of the choice of a finite time horizon T . The approach of the proof is very similar to the very construction of the Itô-integral.

First we need to define our approximating discrete functions. Choose step size $\Delta t = T/N$ where N is the number of partitions of our time space. For every $\rho_i \in [-1, +1]$ now define changes in our semi-continuous process as

$$\Delta \rho_i = \gamma (n - \rho_i) \Delta t + \psi (1 - \rho_i^2) \Delta W_i \quad (7.2)$$

where

$$\begin{aligned} t_i &= i\Delta t \text{ for } i \in \{0, \dots, N\} \\ \Delta W_i &= W_{t_{i+1}} - W_{t_i} = X_i \sqrt{\Delta t} \\ X_i &\sim \mathcal{N}(0, 1) \text{ and } i \neq j \Rightarrow \text{cov}(X_i, X_j) = 0. \end{aligned}$$

We then make the steps in time finer and finer by letting $N \rightarrow \infty$. For every choice of N , by a continuous interpolation between the discrete time steps, we can construct a truly continuous process. Our aim is to show that almost surely $\rho_i \in (-1, +1), \forall i \in \{0, \dots, N\}$ as $N \rightarrow \infty$.

If this can be shown we can use the fact that the t -continuous approximations converge uniformly to the t -continuous process ρ_t^{mod} solving our modified SDE. By letting $N \rightarrow \infty$ we know that the approximating process will be confined in the set $(-1, +1)$ at each point of the infinitely fine discrete approximation and by convergence also that $\rho_t^{\text{mod}} \in (-1, +1)$ for $\forall t \in T \cdot \mathbb{Q} \cap [0, T]$ which is a dense subset of $[0, T]$. (Here $\mathbb{Q}$ is the set of rational numbers.) The continuity in t of ρ_t^{mod} then leads to $\rho_t^{\text{mod}} \in [-1, +1], \forall t \in [0, T]$. Hence the ρ_t^{mod} is almost surely confined in the closed region $[-1, +1]$ for $\forall t \in [0, T]$. Left to show is that our semi-discrete process

$$\rho_{i+1} = \rho_i + \Delta \rho_i$$

will almost surely not leave the set $(-1, +1)$ for any $i \in \{0, \dots, N\}$ as $N \rightarrow \infty$.

To do this we need to prove that $\mathbb{P} \left[\bigcap_{i=0}^N \rho_i \in (-1, +1) \right] \rightarrow 1$ as $N \rightarrow \infty$. In order to show this we start by breaking the constraint of global confinement down into a series of simple confinements as follows.

$$\begin{aligned}
 & \mathbb{P} [\rho_N \in (-1, +1) \cap \rho_{N-1} \in (-1, +1) \cap \rho_{N-2} \in (-1, +1) \cap \dots \cap \rho_0 \in (-1, +1)] = \\
 &= \mathbb{P} [\rho_N \in (-1, +1) \mid \rho_{N-1} \in (-1, +1) \cap \rho_{N-2} \in (-1, +1) \cap \dots \cap \rho_0 \in (-1, +1)] \cdot \\
 & \quad \cdot \mathbb{P} [\rho_{N-1} \in (-1, +1) \cap \rho_{N-2} \in (-1, +1) \cap \dots \cap \rho_0 \in (-1, +1)] = \\
 &= \{ \rho_i \text{ is a Markovian process for every choice of } N \} = \\
 &= \mathbb{P} [\rho_N \in (-1, +1) \mid \rho_{N-1} \in (-1, +1)] \cdot \\
 & \quad \cdot \mathbb{P} [\rho_{N-1} \in (-1, +1) \cap \rho_{N-2} \in (-1, +1) \cap \dots \cap \rho_0 \in (-1, +1)] = \\
 &= \{ \text{Repeat the above procedure for all remaining intersections} \} = \\
 &= \mathbb{P} [\rho_N \in (-1, +1) \mid \rho_{N-1} \in (-1, +1)] \cdot \mathbb{P} [\rho_{N-1} \in (-1, +1) \mid \rho_{N-2} \in (-1, +1)] \cdot \\
 & \quad \cdot \mathbb{P} [\rho_1 \in (-1, +1) \mid \rho_0 \in (-1, +1)]
 \end{aligned}$$

After this break down we need to prove that $\mathbb{P} [\rho_{i+1} \in (-1, +1) \mid \rho_i \in (-1, +1)] \rightarrow 1$ as $N \rightarrow \infty$ for $\forall i \in \{0, \dots, N\}$. Under the condition $\rho_i \in (-1, +1)$ at time step i , the probability of not leaving the set after one extra single discrete time step is given by

$$\begin{aligned}
 & \mathbb{P} [\rho_{i+1} \in (-1, +1) \mid \rho_i \in (-1, +1)] = \\
 &= \mathbb{P} [-(1 + \rho_i) < \Delta \rho_i < 1 - \rho_i \mid \rho_i \in (-1, +1)] = \\
 &= \mathbb{P} \left[\frac{-(1 + \rho_i) - \gamma(n - \rho_i) \Delta t}{\psi(1 - \rho_i^2) \sqrt{\Delta t}} < X_i < \frac{(1 - \rho_i) - \gamma(n - \rho_i) \Delta t}{\psi(1 - \rho_i^2) \sqrt{\Delta t}} \mid \rho_i \in (-1, +1) \right] \geq \\
 &\geq \left\{ \text{Use } \gamma(-1 - \rho_i) \leq \gamma(n - \rho_i) \leq \gamma(1 - \rho_i) \text{ and } (1 - \rho_i^2) = (1 - \rho_i)(1 + \rho_i) \right\} \geq \\
 &\geq \mathbb{P} \left[\frac{-1 + \gamma \Delta t}{\psi(1 - \rho_i) \sqrt{\Delta t}} < X_i < \frac{1 - \gamma \Delta t}{\psi(1 + \rho_i) \sqrt{\Delta t}} \mid \rho_i \in (-1, +1) \right] \geq \\
 &\geq \left\{ \text{Use } -2 < (1 - \rho_i), (1 + \rho_i) < 2 \text{ and choose } N \geq 2\gamma T \text{ to ensure } \frac{1}{2} \leq 1 - \gamma \Delta t \right\} \\
 &\geq \mathbb{P} \left[\frac{-1/2}{2\psi\sqrt{\Delta t}} < X_i < \frac{1/2}{2\psi\sqrt{\Delta t}} \right] = \\
 &= 1 - 2\mathbb{P} \left[X_i > \frac{\sqrt{N}}{4\psi\sqrt{T}} \right]
 \end{aligned}$$

We clearly see that the last line must go towards 1 as $N \rightarrow \infty$ but this is not enough as we need to stay within the admissible set for every single step in our semi discrete process. In mathematical terms this constraint requires the product of all individual steps' probabilities of staying inside the admissible set to converge to 1 as we make the step size smaller. The following calculations

show this constraint to be satisfied.

$$\begin{aligned}
 1 &\geq \prod_{i \in \{0, \dots, N\}} \mathbb{P} [\rho_{i+1} \in (-1, +1) \mid \rho_i \in (-1, +1)] \geq \\
 &\geq \left(1 - 2\mathbb{P} \left[X_i > \frac{\sqrt{N}}{4\psi\sqrt{T}} \right] \right)^N \geq \\
 &\geq \left\{ \text{Choose } N \geq 16\psi^2 T \text{ and use } \mathbb{P}[X > x] \leq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ for } \forall x \geq 1 \right\} \geq \\
 &\geq \left(1 - \frac{2}{\sqrt{2\pi}} e^{-\frac{N}{32\psi^2 T}} \right)^N \geq \\
 &\geq 1 - N \frac{2}{\sqrt{2\pi}} e^{-\frac{N}{32\psi^2 T}} \rightarrow 1 \text{ as } N \rightarrow \infty.
 \end{aligned}$$

This shows that almost surely the limit of our semi-discrete process will stay within the set $(-1, +1)$. We have by this also finished the proof of existence and uniqueness of our original continuous process for instantaneous correlation (7.1).

There are also other ways to approach the proof of confinement of the correlation process. The process can be investigated after a transformation and then transformed back to the original process. This method has its origin in the numerical simulation of stochastic processes and is there used to make the amplitude of the driving Weiner-process constant in time and space. Among other things the transformation is originally used with some processes to get around the problem of a discrete approximation jumping out of an admissible set close to border of this set. By transforming the problem in a form where the process has a constant integrand for the Itô-integral, this integral can be solved exactly in each step and hence reduce numerical errors.

In fact the transformation shown in the next section is very useful for numerical simulations of our correlation process.

7.6 Transformation of correlation process

As with all processes only valid in a certain value set, Monte-Carlo simulations can sometimes result in non-admissible values (a good reference is Kloeden-Platen [25]). In the particular case of our correlation process the admissible set is $[-1, +1]$. If the process takes values outside of this set, it would result in imaginary values for the volatility process we adopt. To avoid this problem the process can often be transformed to another process that does not suffer

7.6. TRANSFORMATION OF CORRELATION PROCESS

from this problem. The interested reader can read more about this kind of transformation in Jäckel [23].

The original stochastic process is given by

$$d\rho = \gamma(n - \rho) dt + \psi(1 - \rho^2) dW,$$

and the idea is to find a transformation, $v = F(\rho)$, that results in a process with a constant integrand for the Itô-integral. The simulation is then instead carried out for the transformed process and the simulated results are then transformed back into the original process ρ . Jäckel shows that in order to cancel out the non-constant part of the integrand such a transformation in this case, because of Itô's lemma [32], needs to satisfy

$$\frac{dF}{d\rho} = \frac{1}{1 - \rho^2}.$$

For real values the solution to this is given by

$$v = F(\rho) = \operatorname{arctanh}(\rho) = \frac{1}{2} \log \left(\frac{1 + \rho}{1 - \rho} \right). \quad (7.3)$$

In this transformation the stochastic process is given by

$$dv = \left(\gamma \frac{n - \rho}{1 - \rho^2} + \rho \psi^2 \right) dt + \psi dW. \quad (7.4)$$

Note that in the above process we still use the non-transformed process ρ . As we want the process to be expressed strictly in our transformed variable we need the inverse transform

$$\rho = F(\rho) = \tanh(v) = \frac{e^{2v} - 1}{e^{2v} + 1}. \quad (7.5)$$

When simulating the process (7.4) the most efficient approach is to use this inverse (7.5) in every step and use the original variable ρ .

Expressed strictly in the transformed variable the process (7.4) becomes

$$dv = [\gamma \cosh(v) (n \cosh(v) - \sinh(v)) + \psi^2 \tanh(v)] dt + \psi dW.$$

It is worth noticing that the transformed process v is free to evolve and take any possible real value. The inverse transform (7.5) will, regardless of the transformed process' value, map the entire real line onto the admissible set $[-1, +1]$. Where the endpoints -1 and $+1$ are only reached for the extreme

7.7. EXPECTED VALUE AND LONG TIME RUN DISTRIBUTION OF THE CORRELATION PROCESS

values $v \in \{-\infty, +\infty\}$.

7.7 Expected value and long time run distribution of the correlation process

The correlation process is a stochastic process and we can calculate an expected value for it. This can be useful when analyzing the effect of the process. We start by writing the process defined by (7.1) on its integral form

$$\rho_T = \rho_0 + \int_0^T \gamma (n - \rho_t) dt + \int_0^T \psi (1 - \rho_t^2) dW_t.$$

The last integral is a martingale and its expected value is zero. The expectation over the first integral can be moved inside the integral using Fubini's theorem (see Rogers & Williams [38]).

$$E[\rho_T] = \rho_0 + \int_0^T \gamma (n - E[\rho_t]) dt$$

On an ordinary differential equation (ODE) form the expectation of the process is given by

$$\frac{dE[\rho_T]}{dT} = \gamma (n - E[\rho_T])$$

with the initial condition

$$E[\rho_0] = \rho_0.$$

The solution to the ODE is given by

$$E[\rho_T] = n - (n - \rho_0) e^{-\gamma T}.$$

For non-path dependent options the final outcome is not as much affected by the actual path taken by the correlation process but rather by the time average of correlation. We call this time average the “effective correlation” between time 0 and T and it is calculated as

$$\overline{\rho_T} = \frac{1}{T} \int_0^T \rho_t dt.$$

To find the average value of the effective correlation as a function of maturity we calculate the expected value of the function. Again by using Fubini's theorem we arrive at

$$E[\overline{\rho_T}] = n + \frac{1}{\gamma T} (n - \rho_0) (e^{-\gamma T} - 1). \quad (7.6)$$

7.7. EXPECTED VALUE AND LONG TIME RUN DISTRIBUTION OF THE CORRELATION PROCESS

This time average effective correlation would roughly describe the correlation value a non-path dependent option will be affected by. To see the impact on this value of the initial correlation value, ρ_0 , we differentiate (7.6) with respect to ρ_0 . The result is given by the expression

$$\frac{dE[\rho_T]}{d\rho_0} = \frac{1}{\gamma T} (1 - e^{-\gamma T}),$$

that no longer is a function of ρ_0 . If we plot this sensitivity to ρ_0 as a function of standard tenor values for T , we find something that looks like the risk-reversal components we earlier saw in Figures 4.7-4.10 when analysing ICA components. There is a clear resemblance between the graph in Figure 7.2 and the ICA component that for each currency pair has the highest impact on spot level and risk-reversals.

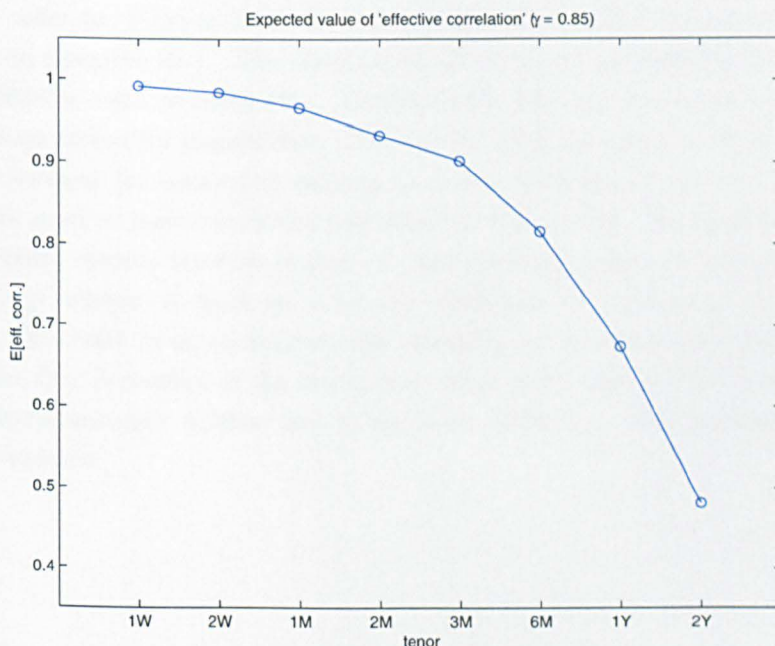


Figure 7.2: Initial correlation value impact on effective correlation

The same approach used to calculate the mean value can sometimes also be used to calculate the expectation of the squared process. This could then be used to calculate the variance of the process. Unfortunately this proves to be somewhat harder in this particular case and is therefore left for further investigations.

We can also look at the distribution of the long time run of the process (7.1).

7.7. EXPECTED VALUE AND LONG TIME RUN DISTRIBUTION OF THE CORRELATION PROCESS

One way of doing this is to look at the infinitesimal generator of the process. The method is described by Fouque *et al.* [13], and the infinitesimal generator is here given by

$$L = \frac{1}{2}\psi^2 (1 - x^2)^2 \frac{d^2}{dx^2} + \gamma(n - x) \frac{d}{dx}.$$

The adjoint operator of this infinitesimal generator is

$$L^* = \frac{1}{2}\psi^2 (1 - x^2)^2 \frac{d^2}{dx^2} + (x(\gamma - 2\psi^2) - \gamma n) \frac{d}{dx} + (\gamma - \psi^2)$$

To get the long time run distribution of the process we need to solve the homogenous equation

$$L^*P(x) = 0.$$

In order for $P(x)$ to really be a probability distribution the solution also needs to integrate to 1. The equation could be solved numerically and there are different ways of doing this. Unfortunately the equation often results in numerous numerical instabilities. One way to solve the equation is to use an Euler forward (or backward) scheme on a two dimensional system. In this case we need to have two initial conditions at some point. We know that the probability density function is zero on the boundary points -1 and $+1$. For our Euler scheme to work we need the conditions to be given at the same point. We could, in an appropriate way, try different values for the initial value and the first derivative of the function at some point until we find a solution that simultaneously reaches zero in the both endpoints. This method is not recommended.

7.7. EXPECTED VALUE AND LONG TIME RUN DISTRIBUTION OF THE CORRELATION PROCESS

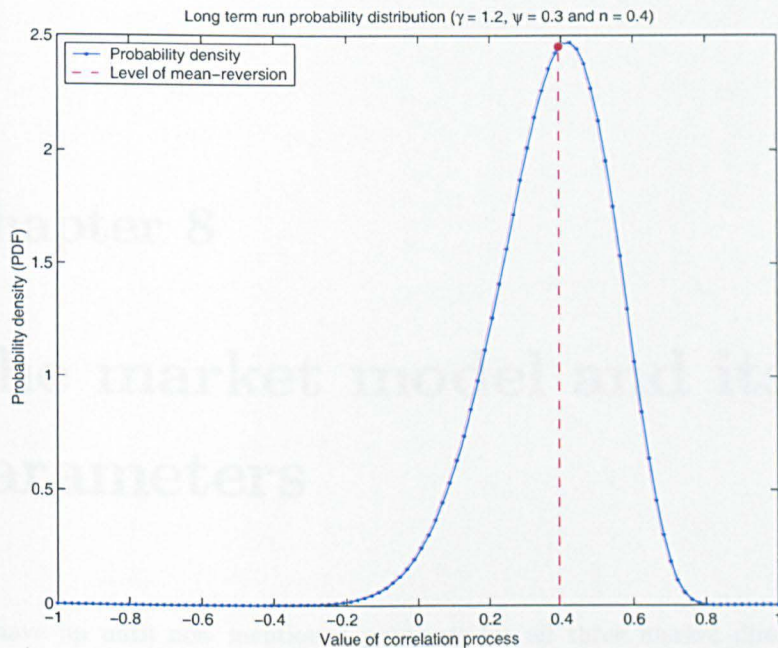


Figure 7.3: Stationary probability density using a minimisation technique

An easier method is to discretise the differential equation using a finite differences approach and try to find its null space (see e.g. Heath [16]). For every element in the null-space the differentials will sum up to zero and hence be an approximation to the homogeneous solution. We have N points in our discretisation and N constraints. The problem is that the resulting matrix has full rank and therefore only has the trivial zero-vector solution. We get around this by adding an extra constraint that the discrete integral of the probability must be 1. The system is now over-determined and we should be able to solve it using a least square approach to get as close as possible to the solution. This however proves to be very hard as the involved matrix has a huge condition number, which makes the solution very numerically unstable.

A way to still be able to solve the problem is to use a minimisation function over the over-determined system. Figure 7.3 shows the resulting distribution using this technique with 100 discrete points. The graph agrees well with a histogram of the same distribution derived using a Monte-Carlo approach.

Chapter 8

The market model and its parameters

We have up until now mentioned processes for all three market dimensions in our framework - spot price, instantaneous volatility of the spot price and instantaneous correlation between spot and volatility.

For spot level we choose the regular Black-Scholes model with the extension of adding the processes describing the other states of the market. Those states are volatility and correlation between spot price and volatility.

Earlier we used the Heston market model and its corresponding model of the stochastic volatility process. As also mentioned earlier, the main reason for this choice is that we arrive at a semi-closed pricing formula for regular European options. The process is not chosen because we believe the square-root model describes the real world better than any other model. In fact the square-root process used in the Heston model can result in numerical difficulties and negative variance during simulation. We will for those reasons not use the square-root process to represent volatility but instead choose the more well behaved logarithmic Ornstein-Uhlenbeck process.

The instantaneous correlation will from now on be modelled as previously described in (7.1).

8.1 Partial differential equation

We start by showing the general PDE for a derivative dependent on three processes. Begin by defining our three state processes by

$$\begin{aligned} d\theta_1 &= \mu_1 dt + \sigma_1 dW^{(1)} \\ d\theta_2 &= \mu_2 dt + \sigma_2 dW^{(2)} \\ d\theta_3 &= \mu_3 dt + \sigma_3 dW^{(3)}. \end{aligned}$$

Using the an equilibrium argument we can derive a PDE for the price, $\Pi(\theta_1, \theta_2, \theta_3)$, of a derivative dependent on our three state variables. Following the approach described in e.g. Appendix 19B in Hull [18] we get

$$\begin{aligned} &\frac{\partial \Pi}{\partial t} + (\mu_1 - \lambda^{(1)}\sigma_1) \frac{\partial \Pi}{\partial \theta_1} + (\mu_2 - \lambda^{(2)}\sigma_2) \frac{\partial \Pi}{\partial \theta_2} + (\mu_3 - \lambda^{(3)}\sigma_3) \frac{\partial \Pi}{\partial \theta_3} + \\ &+ \frac{1}{2} \left(\sigma_1^2 \frac{\partial^2 \Pi}{\partial \theta_1^2} + \sigma_2^2 \frac{\partial^2 \Pi}{\partial \theta_2^2} + \sigma_3^2 \frac{\partial^2 \Pi}{\partial \theta_3^2} \right) + \\ &+ \rho_{12}\sigma_1\sigma_2 \frac{\partial^2 \Pi}{\partial \theta_1 \partial \theta_2} + \rho_{13}\sigma_1\sigma_3 \frac{\partial^2 \Pi}{\partial \theta_1 \partial \theta_3} + \rho_{23}\sigma_2\sigma_3 \frac{\partial^2 \Pi}{\partial \theta_2 \partial \theta_3} - r\Pi = 0 \end{aligned}$$

Here the parameters $\lambda^{(i)}$ represent the market price of risk for the involved processes. Those prices are not required to be constant and can themselves be processes. As our first process, θ_1 , is a foreign currency we have a drift given by $\mu_1 = \mu_S S$ and a diffusion given by $\sigma_1 = \sigma_S S$. As for the Black-Scholes equation the market price of asset risk is given by $\lambda^{(1)} = (\mu_S - r^d + r^f)/\sigma_S$. The other prices of market risk are slightly more complicated as they are actually processes. We will look closer at those prices later.

If we choose a logarithmic Ornstein-Uhlenbeck process to model our volatility the actual market model is given by

$$\begin{aligned} dS &= \mu S dt + e^Y S dW^{(1)} \\ dY &= \alpha(m - Y) dt + \xi dW^{(2)} \\ d\rho &= \gamma(n - \rho) dt + \psi(1 - \rho^2) dW^{(3)}. \end{aligned}$$

Under this market model the premium PDE is given by

$$\begin{aligned} &\frac{\partial V}{\partial t} + (r - q) S \frac{\partial V}{\partial S} + (\alpha(m - Y) - \xi\lambda^{(2)}) \frac{\partial V}{\partial Y} + (\gamma(n - \rho) - \psi(1 - \rho^2)\lambda^{(3)}) \frac{\partial V}{\partial \rho} + \\ &+ \frac{1}{2} \left(e^{2Y} S^2 \frac{\partial^2 V}{\partial S^2} + \xi^2 \frac{\partial^2 V}{\partial Y^2} + \psi^2 (1 - \rho^2)^2 \frac{\partial^2 V}{\partial \rho^2} \right) + \\ &+ \rho e^Y S \xi \frac{\partial^2 V}{\partial S \partial Y} + \rho_{13} e^Y S \psi (1 - \rho^2) \frac{\partial^2 V}{\partial S \partial \rho} + \rho_{23} \xi \psi (1 - \rho^2) \frac{\partial^2 V}{\partial Y \partial \rho} - rV = 0 \end{aligned} \quad (8.1)$$

8.2. MUTUAL CORRELATION

The resulting PDE (8.1) is hard to handle and looking for simplifications can help us making practical use of it. The three driving processes are all correlated to each other and in particular the correlation between the two first processes is governed by the state value of the third process. To get rid of one correlation variable we could assume Y and ρ to only be correlated via their correlation to S . (What this really means will be further explained in the next section.)

However there is no direct mathematical problem in giving the drivers of Y and ρ a freely chosen correlation. The situation is complicated as ρ is correlated to S and the correlation between Y and S is given by the current value of the process ρ itself. We will in the next section show some empirical support for the assumption of Y and ρ only being correlated via S .

As a diverting discussion on the topic of simplifications we could mention a drastic reduction of the complexity. One could completely drop the third process representing the correlation and instead make it dependent on the other two state variables by using $\rho_{12} = \rho_{12}(S, Y)$. We could e.g. use a simple function as $\rho_{12} = 2 \arctan(a(S - b)) / \pi$ or $\rho_{12} = (\exp(a(S - b)) - 1) / (\exp(a(S - b)) + 1)$ for some constants a and b . A function of this form would give a positive relation between risk-reversals and spot price just as observed in the market. The drawback of this simplification is that the correlation, and hence the risk-reversal, would not have the mean-reverting property we can observe empirically. Also, this type of simplification should probably not be advocated as by doing so we try to explain two separate phenomena by only one dimension. The introduction of stochastic volatility is such an important generalisation that, if we choose to only use two dimensions, the second dimension should probably be dedicated solely for it.

8.2 Mutual correlation

The three-factor market process described in the previous section is not well defined only by the involved obvious parameters like volatility along with level and speed of mean reversion. In order for the framework to be fully defined we also need the initial values and the correlations between the driving factors. Assume $W^{(1)}$, $W^{(2)}$ and $W^{(3)}$ be our model's driving Wiener-processes. Below we show how those driving processes can be constructed from the three

8.2. MUTUAL CORRELATION

independent Wiener-processes denoted, $W^{(a)}$, $W^{(b)}$ and $W^{(c)}$.

$$\begin{aligned}
 dW^{(1)} &= dW^{(a)} \\
 dW^{(2)} &= \rho_{12}dW^{(a)} + \sqrt{1 - \rho_{12}^2}dW^{(b)} \\
 dW^{(3)} &= \rho_{13}dW^{(a)} + \pi_2dW^{(b)} + \sqrt{1 - \pi_2^2 - \rho_{13}^2}dW^{(c)} \\
 \mathbb{E}[dW^{(1)}dW^{(2)}] &= \rho_{12}dt \\
 \mathbb{E}[dW^{(1)}dW^{(3)}] &= \rho_{13}dt \\
 \mathbb{E}[dW^{(2)}dW^{(3)}] &= (\rho_{12}\rho_{13} + \pi_2\sqrt{1 - \rho_{12}^2})dt
 \end{aligned}$$

In the above relations the correlation ρ_{12} is actually our correlation process ρ_t and is itself stochastic. We could assume that the processes for volatility and correlation to be correlated only by their mutual correlation to the spot process. This would mean $\pi_2 = 0$. We want to find some economic or empirical support for this assumption. A direct economic argument is not very easy to give but on the other hand there is no direct argument that speaks for an extra contribution to the correlation either. One could argue that as the spot price is the only truly observable process and both volatility and correlation are hidden processes generated by the spot market, it would be plausible to believe that the two hidden processes do not have any further mutual communication apart from their spot dependency.

The empirical approach is done via a crude study of actual market data where we form groups, or “rolling windows”, of 50 consecutive daily market readings. We run through out the entire data set by shifting the group one day at a time. For every group we will calculate the correlations between changes in the three market factors. We start with our original market model.

$$\begin{aligned}
 dS &= \mu S dt + e^Y S dW^{(1)} \\
 dY &= \alpha(m - Y) dt + \xi dW^{(2)} \\
 d\rho &= \gamma(n - \rho) dt + \psi(1 - \rho^2) dW^{(3)}
 \end{aligned}$$

As the available data is discrete we rewrite the processes on a discrete form. We also make the very crude assumption that for every group of market readings the above factors connecting and nesting the processes, as functions of each others, are instead approximately constant within the group. Those constants could e.g. be the mean value within the group but we never need to worry about the actual value as long as it is assumed to be constant. The resulting processes are now no longer nested but are separated and we have marked the replacing

8.2. MUTUAL CORRELATION

nesting mutual dependent variables with a bar on top of the original name.

$$\begin{aligned}\Delta S &= \mu \bar{S} \Delta t + \bar{e}^V \bar{S} \Delta W^{(1)} \\ \Delta Y &= \alpha (m - \bar{Y}) \Delta t + \xi \Delta W^{(2)} \\ \Delta \rho &= \gamma (n - \bar{\rho}) \Delta t + \psi (1 - \bar{\rho}^2) \Delta W^{(3)}\end{aligned}$$

We also assume that the correlation between $W^{(1)}$ and $W^{(2)}$ has the constant value $\bar{\rho}$. With these assumptions the drifts of the three processes are constant and will not affect the correlations. To simplify our further steps we just drop the drift terms and replace the stochastic increments by the three random variables Y_1 , Y_2 and Y_3 . We also replace our driving increments with a linear combination of three identical and independent random variables with standard normal distribution. These random variables are X_1 , X_2 and X_3 . The standard deviation of the very stochastic increments themselves are $\sqrt{\Delta t}$. As all the factors in front of the stochastic increments are constant we can replace them with the three constants σ_1 , σ_2 and σ_3 . The new system is

$$\begin{aligned}Y_1 &= \sigma_1 X_1 \\ Y_2 &= \sigma_2 \left(\rho_{12} X_1 + \sqrt{1 - \rho_{12}^2} X_2 \right) \\ Y_3 &= \sigma_3 \left(\pi_1 X_1 + \pi_2 X_2 + \sqrt{1 - \pi_1^2 - \pi_2^2} X_3 \right).\end{aligned}$$

Here π_1 and π_2 are just constant coefficients describing the linear construction of the third stochastic driver. Let the correlations between the random variables Y_1 , Y_2 and Y_3 be denoted by ρ_{12} , ρ_{13} and ρ_{23} . The first correlation ρ_{12} is assumed to follow a stochastic process and other two correlations are given by

$$\begin{aligned}\rho_{13} &= \frac{\sigma_1 \sigma_3 \pi_1}{\sigma_1 \sigma_3} = \pi_1 \\ \rho_{23} &= \frac{\sigma_2 \sigma_3 \rho_{12} \pi_1 + \sigma_2 \sigma_3 \sqrt{1 - \rho_{12}^2} \pi_2}{\sigma_2 \sigma_3} = \rho_{12} \rho_{13} + \sqrt{1 - \rho_{12}^2} \pi_2.\end{aligned}$$

Rewriting the second equation gives us

$$\pi_2 = \frac{\rho_{13} - \rho_{12} \rho_{13}}{\sqrt{1 - \rho_{12}^2}}.$$

After this long and in many aspects crude and approximating simplification we want to find empirical evidence for π_2 being zero (or at least very close to zero). We go back to our original processes and we now have to tackle the problem that only one of our market processes, i.e. the spot price, is directly ob-

8.2. MUTUAL CORRELATION

servable. Both the processes for instantaneous volatility and the instantaneous correlation are hidden processes. The only observable market data relating to those two processes are implied volatility and risk-reversals. We again choose to replace the non-observable processes with data for the volatility 1-week maturity called tenor 1W.

We start by noticing that if π_2 really was zero then we would have

$$\rho_{23} | \pi_2=0 = \rho_{12}\rho_{13}.$$

This would result in the correlation between Y_1 and Y_2 being very similar (at least in shape due to the factor ρ_{13}) to the correlation between Y_2 and Y_3 .

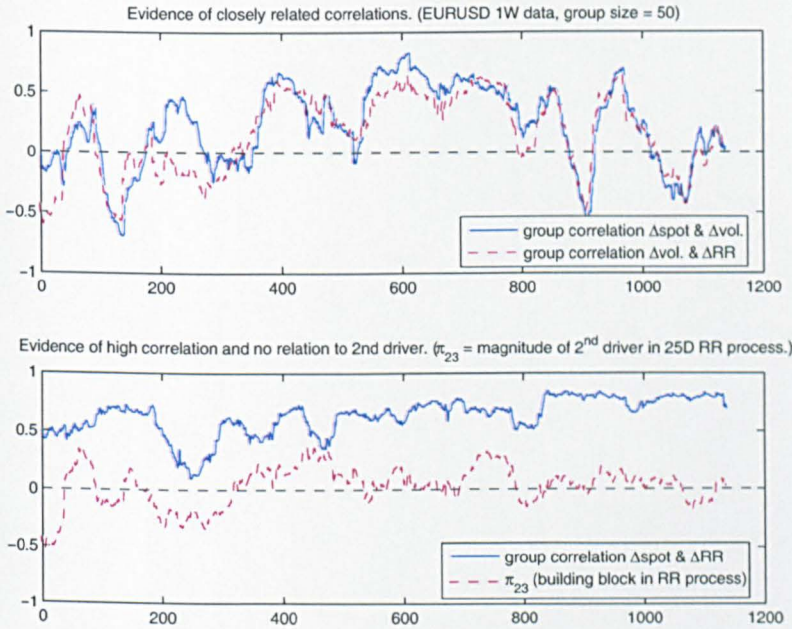


Figure 8.1: Relations and correlations for market factors

The upper graphs in Figure 8.1 show an high agreement for $\text{corr}(\Delta S, \Delta vol)$ and $\text{corr}(\Delta vol, \Delta RR)$. The only place where the agreement is low is in the area around reading number 250. This area coincide with a low $\text{corr}(\Delta S, \Delta RR)$. The same type of pattern can be seen for all other currency pairs in the data set.

At the end of this part we should again stress the fact that this is a non-rigorous statistical justification of our assumptions about the correlations between the driving factors. Nevertheless the study shows that the assumptions are not implausible. As the empirical study is model independent this should

8.2. *MUTUAL CORRELATION*

make an empirical contribution to the entire class of stochastic correlation models.

Chapter 9

Markovian properties and market completeness

9.1 Assumptions of Markovian properties

In simple terms a Markov process is a process where the probability of the future distribution is only dependent on the present state. If our market model was an Itô-diffusion process we would know that our market would have the Markov property (see Section 7.1 in Øksendal [32]). Our problem is that our correlation process, ρ_t , does not satisfy the necessary Lipschitz condition. This is the same problem we had when proving existence and uniqueness of the solution. In that section we showed that the solution, with probability 1, is equal to the solution of the modified SDE which satisfies all necessary conditions. As the solution to this modified SDE is a regular and well behaved Ito-process it also has the Markov-property. The conclusion is that, with probability 1, also our correlation process will have the Markov-property.

9.2 Market completeness

A market is free from arbitrage if there exists an equivalent martingale measure [40]. In theory a market is complete if every contingent claim is attainable by trading in available assets. Further, an arbitrage free market is complete if and only if there exists a unique martingale measure (see Section 4.3 in Bingham & Kiesel [1]).

In the case of stochastic correlation all market activities are described by the market model's three market processes S_t , σ_t and ρ_t . If T is our economic

horizon, each process at time $t \in [0, T]$ is a random variable measurable in the σ -algebra $\mathcal{F}_t$ and the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t=0}^T$ holds all relevant market information. If we were only able to trade the underlying asset S_t there would be room for more than one equivalent martingale measures and the market would not be complete. The problem is that the states for σ_t and ρ_t are not directly traded. In order to complete the market we need to find other assets that together with the underlying asset will fully generate the filtration $\mathbb{F}$. This can be done by adding two options as non-redundant assets to the market. Those options must hold a direct dependency to the non-traded processes to expose the part of the filtration orthogonal to, and hence not exposed by, the underlying asset.

9.3 Market price of risk

The issue of market price of risk is often not paid much attention when searching for models explaining what is observed in the real market. One can often read quick detours like “we assume that we are in a risk-neutral measure”, without hesitating to assume that the market-price-of-risk is a result of the chosen measure rather than letting the measure be a result of the market-price-of-risk. The whole phenomenon is directly addressed in Chapter 10 of Björk [2] where we can read: “Who chooses the martingale measure? The market!”.

Indeed the market chooses the pricing measure and this is a result of the aggregated attitude towards risk by the market participants. Even though the end result is the same, it seems more reasonable to assume that the market participants form an attitude towards risk before they start trading options. The difference might seem to be of a philosophical nature but if the attitude towards risk is formed before we look at and trade options this indicates that the resulting market-price-of-risk would have a quite simple mathematical form.

Is it really plausible to believe that the market chooses a price of risk that just happens to make the risk-neutral measure generated by a simpler market model the same as the one chosen in a more complicated market model to reproduce observed market prices? If we think this is not the case we should not use the market-price-of-risk as a means to match observed market prices but instead choose a simple form for it and instead make the market model more general. This is in some sense a critique of pricing models that try merely to find a risk-neutral process that reproduces observed option prices. Referring back to our previous discussion on market models in general this would advocate the so called *fundamental approach* to modelling the market.

The risk-neutral measure is a result of the cost of being perfectly hedged. If the market is incomplete, the market price of risk represents a real or fictious

9.3. MARKET PRICE OF RISK

price accepted by someone else (or ourselves) to accept the non-traded risk. If the market really has a conscious relation to the trading of this risk it is reasonable to believe its price to be very regular, if not even constant.

In the stochastic volatility model proposed in Heston [17] the market price for volatility risk is chosen to be linear with instantaneous variance. This gives us the possibility to include the market price of risk in the other model parameters. By inclusion we can then rewrite the risk neutral process onto the same form as the original process in the objective measure. The new parameters are called “risk neutralized” parameters. There is a not straightforward argument behind this choice of market price of risk. The most plausible reason for this choice must however be the fact that the resulting PDE can be solved in a semi-closed form.

In a stochastic correlation framework we have three sources of uncertainty and hence three market-prices of risk. Only the underlying asset is directly traded so this risk can be fully hedged away at a known cost. The volatility and correlation risks are not directly traded and we therefore need the market prices for those risks in order to price derivatives. What market prices of risk should we choose for the volatility and correlation risks? Following the view about simple forms of risk suggested above, the market price of risk is not directly connected to risky market factors but to the driving factors of those parameters. Heston seems to ignore, or at least not mention, the fact that a certain part of the volatility risk actually can be hedged by spot because of its correlation to spot. This would introduce both the market price for spot risk, often denoted $\lambda^{(S)} = (\mu - r)/\sigma$, and the spot-volatility correlation ρ in the market price for risk. In a stochastic correlation model this correlation is not constant. If we want our price of risk to be of a simple constant form we need to include the spot-volatility correlation state value in the market price of volatility. For the market price of correlation risk the corresponding correlation to spot price is constant.

To better understand the market price of risk we look back at the construction of our model’s driving processes, as described in the previous section. Assume the market price of risk for the fundamental independent driving processes $W^{(a)}$, $W^{(b)}$ and $W^{(c)}$, given by $\lambda^{(a)}$, $\lambda^{(b)}$ and $\lambda^{(c)}$, to be constant. The market price of risk for our correlated derived process $W^{(1)}$, $W^{(2)}$ and $W^{(3)}$ would then

9.3. MARKET PRICE OF RISK

be given by

$$\begin{aligned}\lambda^{(1)} &= \lambda^{(a)} \\ \lambda^{(2)} &= \rho_t \lambda^{(a)} + \sqrt{1 - \rho_t^2} \lambda^{(b)} \\ \lambda^{(3)} &= \rho_{13} \lambda^{(a)} + \pi_2 \lambda^{(b)} + \sqrt{1 - \rho_{13}^2 - \pi_2^2} \lambda^{(c)}.\end{aligned}$$

Under the earlier assumption that $\pi_2 = 0$, the third market price of risk $\lambda^{(3)}$ is slightly simplified.

The significance of the market-price-of-risk is of course not negligible and is directly linked to market participants' risk-aversion and the absence of arbitrage. But as long as we only think of it as a way to account for the fact that investors are risk-averse, it will neither substantially change nor further explain prices of options. It can be compared to interest rates that, with the obvious exception of fixed-income markets, most often is assumed to have a deterministic behaviour. Without a dramatic change in our beliefs about the market's relation to risk, and hence the market-price-of-risk, it will only result in a different risk-neutral drift and just give us another parameter to calibrate. This is an area of further research and for all pricing done in later sections we have used the simplification $\lambda^{(2)} = \lambda^{(3)} = 0$.

Chapter 10

Qualitative analytic approximations

As qualitative approximations we here mean techniques to determine the general impact of the model rather than the quantitative impact. The line between a qualitative or a numerical/quantitative approximation can be unclear. A qualitative model can sometimes be refined into giving very accurate numerical results and a numerical scheme can sometimes be simplified into just giving quick numerical results. This is the reason why some methods listed here could also be regarded as numerical techniques.

This has not been a main area of our research and this section should rather be looked at as ideas of further research.

10.1 Asymptotic pricing and series expansions

The technique of asymptotic pricing and series expansions is only valid for very small values of crucial model parameter values and is a perturbation technique. By first setting characteristic parameters in the model to zero, the model sometimes turns into an easier problem which is easy to solve. Such a parameter could for example be the volatility of volatility in a stochastic volatility model, where setting this parameter to zero essentially turns the model into the standard Black-Scholes model. By representing the solution to the PDE as a series expansion in the characteristic parameter we can group terms in the series together and instead solve a number of simpler problems.

The interested reader is referred to Chapter 5 of Fouque *et al.* [13] or Chapter 3 of Lewis [27].

In mathematical terms the method is known as *perturbation theory* for PDEs.

Even though the technique can be generalised into many dimensions we often only see the 1-dimensional case being used in finance. This version is also called *singular perturbation* and is what is used in the work in two above references.

In the case of stochastic correlation the perturbation technique would take its origin in the Black-Scholes PDE, which would then be expanded in the two new dimensions of volatility and correlation. This means that perturbation must be done in two rather than in one dimensions. Even though this is theoretically possible and only a few terms in the expansion probably would be sufficient, it could be much more complicated than a singular perturbation.

10.2 Hull-White approach to correlation value in a stochastic volatility model

The approach presented by Hull & White [19] is based on the observation that in a stochastic volatility model the prices of vanilla options are only dependent on the mean variance of volatility up to maturity. If volatility, or rather variance, is non-correlated to spot price and we accept volatility risk with no extra risk premium we can calculate the price of the vanilla option as the expected value of the Black-Scholes price under the probability distribution of mean variance.

The method could, to some extent, be used in a stochastic framework. We could use a stochastic volatility model as a starting point and handle the correlation in the same way as volatility is handled by Hull & White. This would result in the stochastic volatility model being used to price options over a distribution of values for the correlation parameter. The assumption of no extra risk premium for correlation risk is no major problem. The assumption of an independent correlation process is a larger restriction as we empirically have shown that the opposite is true in the foreign exchange market.

10.3 Fourier transform techniques

The Fourier transform technique should probably be regarded as a numerical approximation rather than a qualitative one. The aim is to use the Fourier transform to reduce the PDE into a new one but of a lower dimensionality and where the new PDE is a function of a transform variable. Probably the most famous use of this technique in finance is the pricing equation for vanilla options under stochastic volatility proposed in an article by Heston in 1993 [17]. The article describes how, by choosing a specific process for stochastic volatility, one can use the Fourier transform to reach a semi-closed pricing formula for vanilla

10.3. FOURIER TRANSFORM TECHNIQUES

options. A comprehensive explanation of the technique is given in e.g. Lewis [27].

For such choices of market models and options where the Fourier technique results in more or less analytic formulae, pricing can be done very quickly. The vastly improved speed over slower numerical schemes is a great advantage in model calibration. There are unfortunately only a very limited set of market models that lead to analytic formulas and this has resulted in some criticism of the method. In order to reach such formulae one is tempted to sacrifice agreement with the observed market and chose an implausible market model for the sole purpose of ending up with a suitable PDE.

As the original idea uses the transform of option payoffs at maturity, only European style options can easily be priced by this method. This might at first not seem to be of practical interest in a market with liquidly traded vanilla options and hence where prices for all European style options are already available. However there is a great advantage in being able to price also vanilla option prices quickly when calibrating the model to an existing market.

It is not sure that the Fourier transform would lead to a quicker pricing scheme under a stochastic correlation model. If no analytic solution can be found to the transformed PDE it must be solved numerically. This is also very likely the case for a stochastic correlation model. On the upside the transformed PDE is of lower dimensionality and is much quicker to solve, but on the downside the PDE must be solved for a range of values of the transform variable.

Chapter 11

Numerical approximations

11.1 Monte-Carlo simulations

This approach is by far the simplest to implement and an implementation that produce results can quickly be put together. Even though there are many techniques to make the method more efficient it is very computer intensive. On the upside it “scales linearly” and can easily be implemented in a multiprocessor environment with virtually no overhead computational cost. Another issue worth mentioning about the Monte-Carlo method is the natural uncertainty of the results produced.

11.2 Tree or lattice models

A simple and popular model in the one-dimensional Black-Scholes world is the so called tree models (see e.g. Hull [18]). This class of models can be generalised into higher dimensions. In the case of a stochastic correlation model we would require three space dimensions and one time dimension. As this would give us a tree in four dimensions, it would be a very memory intensive approach to implement as the tree grows like $N_{steps}^{time} \times N_{steps}^{spot} \times N_{steps}^{volatility} \times N_{steps}^{correlation}$. A well known problem is also that a tree model needs very many time steps in order to avoid numerical instabilities in the solution. This could make the otherwise very quick model unpractical. On the upside the models intrinsic property of dropping the boundary nodes during back-propagation, allows us to simply ignore any problems that might come from the boundary conditions.

11.3 Finite differences

As we have the general PDE (8.1) of the problem, solving this numerically could still be possible despite having three space dimensions and one time dimension. As a rule of thumb Monte-Carlo methods are said to be among the fastest techniques for problems with a dimensionality higher than four. If we choose to numerically solve the PDE a popular class of methods are the finite differences methods. If we choose the more stable implicit type of finite differences schemes we will have to solve very large systems of linear equations. We could use an iterative technique like Gauss-Seidel to solve such linear system. Another possibility is to use an so called alternating direction implicit scheme (ADI). These types of implicit schemes are quick but generally have problems solving PDEs with non-zero cross derivatives .However there are versions of the ADI schemes that, at some computational cost, can handle also this type of PDEs (see work by Craig & Sneyd [9]). Before we start implementing any of the possible finite differences schemes we must find a way to handle the boundary conditions.

Chapter 12

Finite differences

The general idea of finite difference techniques is to approximately solve a PDE by discretising it in both space and time. The differentials in the PDE are locally approximated by using adjacent points. As the PDE we are trying to solve only has differentials up to the second order, it is enough for the finite-differences approximation to only include one extra discrete point in each space direction. In our case we also need to take second-order cross-derivatives into account so we will also have to include points one step away in each of the two involved dimensions. Without going into much further details, for each node in our finite differences approximation we only need access to all the eight closest neighbouring nodes. This is a key point in the speed of finite differences methods.

The PDE, together with the discrete approximations of the differentials, form a well defined set of relations between the points in the discrete solution. As both the original PDE and the approximation are linear, the finite difference approach leads to a linear matrix relation for each time-step of our approximation. Depending on the way we choose to approximate our differentials this matrix relation can be either a straight-forward matrix multiplication (explicit schemes) or a linear matrix system (implicit schemes). A more thorough introduction to finite differences can be found in Iserles [22] or Travella & Randall [44].

Assume that we want to solve a PDE in d space dimensions and one time dimension. We choose to discretise our i th dimension into N_i discrete points and our time dimension into M discrete points. Our entire solution, at each of our M time steps, has $N_1 \times N_2 \times \dots \times N_d$ points. The matrix connecting two solutions at different adjacent time-steps will have $(N_1 \times N_2 \times \dots \times N_d)^2$ elements. It is obvious that the size of this matrix rapidly grows as either the

dimensionality or the precision of the solution increases.

When using an explicit scheme the exploding number of elements is not really a problem as we only need to care about the non-zero elements, which are all easy to reach. However for an implicit scheme we need to solve a linear system. In one space dimension this is easy as the resulting matrix is tri-diagonal. Such matrices are both easy to store and solve (see Numerical Recipes in C [35] Section 2.4). On the other hand, for implicit schemes, this simple structure quickly gets lost and matters get more complicated in higher dimensions. The resulting matrix is, thanks to the local nature of the approximations of differentials, very sparse but is no longer a band-matrix and is hence not trivial to solve.

A simple handling of an, otherwise very large, matrix can actually be maintained when using some iterative solution approximations like the Jacobi, Gauss-Seigel or other similar versions of those methods. This because those methods are essentially based on iterated explicit methods, which again only need the easily reached non-zero elements in the matrix.

The reason to choose the more complicated implicit method rather than a less complicated explicit method is a matter of convergence and stability. An explicit method requires finer time-steps to reach the same precision and stability as an implicit method. There is however a class of finite difference schemes that combines the best of the two worlds. Schemes in this class are called Alternating Direction Implicit (ADI) schemes and they break up a complex multi-dimensional implicit time-step into a number of, well approximating, simple one-dimensional time-steps, which are more easily solved. The reason to choose an ADI method over an implicit method using an iterative solver is a matter of computational effort as the ADI scheme is much faster (see p. 115 in Overhaus *et al.* [33]).

12.1 Alternating direction implicit (ADI)

The general idea behind the ADI scheme is to start with a step using a fully explicit scheme for all but one dimension. The remaining dimension is treated implicitly. The result is in practice a one-dimensional implicit step and it hence results in a simple tri-diagonal matrix system. We then treat the other dimensions one at a time. The dimension's contribution stemming from the initial explicit step is subtracted and instead this part of the total contribution is treated as a new implicit step. This creates artificial intermediate steps, which in the end are no longer needed. Those extra steps should not be thought of as

12.2. BOUNDARY CONDITIONS

intermediate steps in time but rather as a successive refinement of the result of the main single time step.

In the case of no cross-derivatives or any contribution proportional to the value of the function (zeroth-order terms), in the end we have gradually replaced all of the explicit step by many implicit steps. If the PDE has non-zero cross-derivatives or any contribution proportional to the function, this part of the step is never treated implicitly. The full step then includes an explicit part which reduces the convergence to that of a fully explicit step. If the coefficients in the PDE are constant the cross-derivatives can be removed by a linear transform (see p. 77 Clewlov & Strickland [6]). In our case the coefficients of the PDE are not constant and we, therefore, cannot use this otherwise very efficient transform to get rid of the cross-derivatives.

The complication of cross-derivatives and proportional terms can, however, be overcome by essentially running the same ADI step one extra time. In this second run the previously explicit contribution is time-centred by using the result of the first run as an approximation to the final result of the step. This method for increased precision is very similar to the “predictor-corrector” method often used for ordinary differential equations. For a full explanation of the handling of cross-derivatives in ADI see Craig & Sneyd [9] or Overhaus *et al.* [33] p. 110.

12.2 Boundary conditions

An extreme but simple way to deal with boundary conditions is to make the grid of nodes so big that the boundary nodes will have no relevant impact on the final solution. The chance of the underlying process actually reaching the boundary is very small and one can in such a setup more or less ignore the behaviour at the boundary nodes. The technique can be wasteful and inefficient and should therefore only be thought of as a last resort. The exception is if we have strong mean reversion. In such a case the same phenomena can be observed for quite moderate grid sizes. Again it is all dependent on the probability of actually reaching the boundary.

A general problem at boundary points is the lack of surrounding nodes in the direction of the boundary itself. One technique to calculate higher order derivatives at a boundary point where we only have node points on one side is the up-winding technique, which changes the used points of centre and hence avoids the need for node points outside the limited grid. The main constraint on this method is that the drift is working in the direction towards the boundary.

12.3. STABILITY, CONVERGENCE AND SPURIOUS OSCILLATIONS

For a mean reverting process, like the volatility and correlation processes in our case, the drift is working away from the boundary. Fortunately the entire concept is saved by the fact that we solve the PDE backwards in time. From a numerical point of view, time is running backwards and the drift is actually working towards the boundary points. The drawback is that up-winding at the boundary leads to non-tridiagonal solutions if we need to use a second order derivative boundary condition. The exception is the often used *zero-gamma* condition where the second order derivative is assumed to be zero.

For vanilla options and forward starting options the zero-gamma condition is true for both the upper and lower boundary in the spot direction. For one-touch options it is true on the opposite side of the barrier to touch. Fortunately we know that the price on the barrier is given by 1.0 for a pay-at-hit option and $e^{-r(T-\tau)}$ for an pay-at-expiry option. Here r is the risk-free interest rate, T maturity time and τ is the time of hitting the barrier.

When it comes to boundary conditions of volatility and correlation the situation is more complicated.

The fact that the payoffs of all ordinary options are only dependent on spot level at maturity but are not dependent on volatility or correlation suggests the solution to the PDE will have linear boundary behaviour in those dimensions, at least for the time period close to maturity. A linear behaviour is equal to a zero-gamma boundary constraint. As mentioned above, the fact that both volatility and correlation are mean-reverting, will reduce the impact of those variables. The conclusion is that even though a linear behaviour at the boundary in those dimensions might not be perfect, any resulting error introduced by it will be small.

A zero-gamma boundary condition for stochastic-volatility using a ADI method was also suggested in Overhaus *et al.* [33] p. 112.

12.3 Stability, convergence and spurious oscillations

Using the above mentioned ADI scheme with cross derivatives, we let θ denominate the “implicitness” of the scheme. Choosing $\theta = 0$ gives a fully explicit scheme, $\theta = 1$ is a fully implicit scheme, and the choice $\theta = \frac{1}{2}$ gives us the famous Crank-Nicolson scheme. We also specify the implicitness for the two-pass-calculation of the cross derivatives as θ_{cross} . If x_i represents any of the d spatial dimensions, the convergence if the scheme is $O((\Delta x_i)^2) + O((\Delta t)^2)$ for $\theta = \theta_{cross} = \frac{1}{2}$ and $O((\Delta x_i)^2) + O(\Delta t)$ for any other choices of θ or θ_{cross} . For $d \leq 3$ the scheme is unconditionally von-Neumann-stable for $\theta \geq \frac{1}{2}$

12.3. STABILITY, CONVERGENCE AND SPURIOUS OSCILLATIONS

when $\theta_{cross} = \frac{1}{2}$. More details about the derivation of stability and convergence properties are given in Craig & Sneyd 5 [9].

The von Neumann stability condition is sufficient to avoid exploding solutions and oscillations of growing magnitude in time. It is however not enough to guarantee absence of transient oscillations, also known as *spurious oscillations*. This type of oscillations can occur in the presence of discontinuities in the initial values. In our case this means sharp changes in the value at expiry (e.g. barrier-style options). The phenomena can also be seen when the PDE we are solving is *convection dominated*. The term is borrowed from fluid dynamics and means that the diffusion term is relatively small in comparison to the convection term. In financial language this means that the volatility is small in comparison to the interest rate differential.

To avoid the spurious oscillations we need to satisfy other constraints (see Duffy [11] Section 17.4). The first such condition is the so called Péclet condition and is, with a slight modification to introduce the foreign interest rate, given by

$$\Delta S < \frac{\sigma^2 S}{r - q}.$$

Another condition is derived in Zvan *et al.* [47]. This condition is

$$\frac{1}{(1 - \theta) \Delta t} > \frac{\sigma^2 S^2}{(\Delta S)^2} + r.$$

Meeting this condition guarantees that new points do not end up over or under the immediately adjacent points. If there is a slope in the function to be calculated, small oscillations can actually occur without violating what the second condition is guaranteed to prevent. In other words the second condition does not guarantee absence of oscillations but only that the oscillations are small enough not to change the mutual magnitude relation between adjacent points.

We must not forget that the above two conditions are derived for the spot dimension only. Theoretically similar constraints need to be derived for the other dimensions as well. In practice however this is not needed as the payoffs of options are essentially always only dependent on the spot price dimension and hence discontinuities can also only exist in this dimension. It is worth mentioning that the above conditions do not tell us anything about the behaviour stemming from the cross derivatives. But as long as discontinuities only occur in one single dimension, cross derivatives will not be affected.

Empirical tests show that no visible oscillation is present as long as the above

12.4. RECYCLING FINITE-DIFFERENCES RESULTS

conditions are met. The parameter settings in the same tests also show that the second constraint causes more problems than the first. If we are willing to trade precision for stability, we can always satisfy the second condition by switching from the Crank-Nicolson scheme ($\theta = \frac{1}{2}$) to a fully implicit scheme ($\theta = 1$).

Another way to suppress the spurious oscillations is to smoothen the initial values in order to get rid of the discontinuities. One might be somewhat reluctant to do this as a change in payoff essentially would mean that we change the very definition of the option in question. A similar method that lets us keep the original payoff structure exists, and is easily introduced in the ADI framework. The method is called the Rannacher method and simply requires the first few critical time-steps to be fully implicit (see Pooley *et al.* [34]). The fully implicit method is less accurate but does not suffer from unwanted oscillations. After the first few time steps the diffusion has naturally smoothened our function and we can switch from explicit to implicit steps without suffering from spurious oscillations.

Even though this issue is very important, further investigations falls outside the scope of this thesis.

12.4 Recycling finite-differences results

When calculating option prices numerically using the finite-differences method we do not only get the result at one single point but simultaneously also at every point in the used mesh grid. All points represents different initial values for the different state variables but prices the option using the same payoff and boundary conditions. When studying pricing we might use many of those points to plot graphs showing the relation between prices and different parameters. This is easy if the parameters we would like to investigate are the initial values of the state variables. But what if we want to look at different parameters?

One solution of course is to reprice the entire framework many times and for each calculation change the specific parameter we are looking at. This is very costly and leaves almost all our information about prices unused. I here propose a method of “recycling” the available data to solve for option prices with different payoff and boundary conditions that I have yet so far not seen in the available literature. The method is based on the following observation:

Assume the solution $\Pi(X, \dots)$ to be the premium of an option. Let $\mathcal{L}\Pi = 0$ be a linear pricing PDE where all differentials with respect to a variable X only occur multiplied by the same variable X raised to the same power as of the

degree of the differential. Also assume that the variable X does not occur in any other places. If $\Pi(X, \dots)$ is a solution to the PDE, so will also $\Pi(X/p, \dots)$ be.

In economic terms this is equivalent to a change of numeraire and as long as the spot price has the same relative behaviour to the spot price level, as is true for most used models, the procedure is model independent. The only possible problem to be aware of is changes in behaviour at boundary conditions.

12.4.1 Call and put options

When looking at regular vanilla options we often want to know the value for different strike prices. We want a solution to the PDE with the same boundary conditions but at maturity where we want a different payoff. Let $\Pi(S, K_0, \dots)$ be the price of an option with current spot level at S and with the payoff $\max(S_T - K_0, 0)$. In a finite-difference solution we have this value for many different values of S but all with the same strike level K_0 . We can here use the above observation in the following way; $\Pi(S_0, K_0 \cdot S_0/S, \dots) = S_0/S \cdot \Pi(S, K_0, \dots)$ and let $K = K_0 \cdot S_0/S$. We have here recycled prices for different spot levels S with the same strike level K_0 to generate prices with the same spot level S_0 with different strike levels K .

12.4.2 One-touch options

For a One-Touch option changing the strike level changes the boundary conditions rather than the at-expiry payoff. We are originally given prices at different initial strike levels but all with the same one-touch barrier level. We can here use the different initial levels of spot price, S , to instead push the strike level, K_0 , to different levels, K , but keep the initial spot level constant at S_0 . This is done in the following way; $\Pi(S_0, K_0 \cdot S_0/S, \dots) = \Pi(S, K_0, \dots)$ where $K = K_0 \cdot S_0/S$. The procedure is similar to the one used for call and put options with the slight difference that for One-Touches we do not use a scaling factor for the prices.

12.4.3 Forward starting call and put options

There are many types and definitions of forward starting options, or forward setting options as they are also called. Here we look at one of the simplest types where the strike price of a call or put option is determined by the spot level at a future time point called the "forward setting date". Generally the strike could be set at the future spot level multiplied by any fixed constant multiplier. A commonly chosen construction is to simply set the strike level at the same level as the spot price at a previously agreed future setting date. This is equivalent to

12.4. RECYCLING FINITE-DIFFERENCES RESULTS

choosing a fixed multiplier equal to 1.0 and such option is often called a “forward setting at-the-money option”. Once we have reached the forward-setting date the strike level is known and the option turns into an ordinary call or put option up until expiry of the option.

In the finite-difference framework at the forward-setting date we need the prices of the ordinary option that is determined at that point. This price is not only dependent on the values of all state variables in each node but, most importantly and by definition, also on the prevailing spot level. In our particular case we need the price of at-the-money options priced under a multitude of initial values for the different state variables. This might seem hard but in fact, with a slight modification, we can use the previously mentioned way of re-using call or put option prices.

Start by letting a finite-difference engine run the pricing of a regular call or put option from expiry back to our forward-setting date and halt it there. We now hold prices for a regular option for all values of our state variables included in the used mesh grid. Among all the nodes there is a subset with spot price levels equal to the strike level of the initial option. All those nodes hold the prices of at-the-money options and hence have the right values. All other nodes are incorrectly priced as they hold prices for in- or out-of-the-money options. The important observation is that for every combination of values for the state variables, excluding the spot level, there exists a node that holds the correct price for an at-the-money option and with the same values for all state variables but the spot level. By using the previously described method, this node can now be re-used for all nodes sharing the same values for the non-spot states. Once this is done all nodes in the mesh grid hold the price of an at-the-money option priced using the correct values for the state variables.

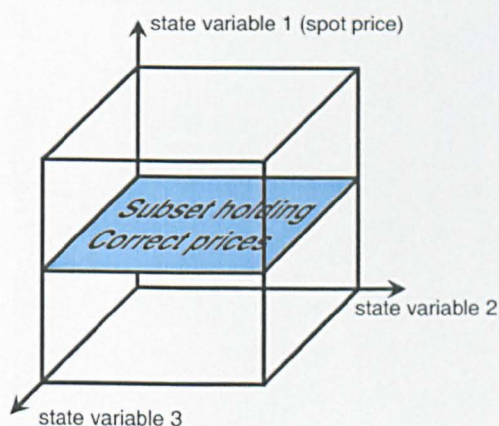


Figure 12.1: Stylised scheme of node subset holding correct prices

12.4. RECYCLING FINITE-DIFFERENCES RESULTS

This procedure can, with a slight modification, be used with other setups than the pure at-the-money forward-setting style. By choosing other nodes as the correctly priced subset, other multipliers than 1.0 can be constructed.

Chapter 13

Results

We have now developed a model and argued for its introduction. We have also looked at different ways of pricing options under this model. Out of those ways the ADI-scheme is the most computationally efficient for most practical contracts. In this section we will not look very much at the actual numerical results but rather at the pricing results produced by the model.

To look at the various cases of path dependency we will study three important option types. Those types are vanilla options, one-touch options (also called American digital options) and forward starting vanilla options. The three option types, in principle, represent three different flows of probability. Vanilla options represents the flow of probability from the present to the future. One-touches represent the flow of probability from one side of the barrier level over to the other side of the barrier level. Forward starting options represent the flow of probability from a future time point to another future time point.

13.1 Vanilla options

Vanilla options represent one of the simplest types of options but also the most important. This contract does not only constitute the vast majority of all traded volatility dependent derivatives, but is also an important instrument for hedging and an important source of market information. In theory vanilla options are independent of the path of the underlying asset and their prices are only dependent on the risk-neutral probability distribution at the expiry date. This essentially makes this type of option independent of the specific market dynamics we wanted to capture with the introduction of stochastic correlation. It does not mean that the properties of the correlation process will not affect the pricing

13.1. VANILLA OPTIONS

of vanilla contract. It means that prices are not affected by the type of changes the stochastic correlation makes to the paths leading up to expiry, as long as the terminal distribution at expiry is the same.

We will here look at call option premia, or rather the equivalent Black-Scholes implied volatilities, produced by the stochastic correlation model. To understand the impact on pricing of the various parameters in pricing we will also plot the differential of premium with respect to those parameters.

Throughout this section all pricing of vanilla options have been made using the model parameter values seen in the below list. The chosen values produce premia for vanilla options close to those which we on average observed in the market for EURUSD. We have here chosen an approximate market of ATM volatility = 10.0%, STR25 = 0.25% and RR25 = 0.35% equal for all the tenors 1M, 3M and 1Y. Calibration is done by using a numerical search and the results are seen in Table 13.1.

variable description	notation	value
spot	S_0	100.0
expiry time	T	3/12 or {3/12, 6/12, 1}
domestic interest	r	0.0
foreign interest	q	0.0
initial log-volatility	Y_0 or $Y0$	$\log(0.099)$
log-vol. reversion level	$\bar{Y}$ or $YMean$	$\log(0.089)$
log-vol. reversion speed	α or $YRev$	9.4
log-vol. volatility	ξ or $YVol$	1.8
initial correlation	ρ_0 or $corr0$	0.1
corr. reversion level	$\bar{\rho}$, n or $corrMean$	0.0
corr. reversion speed	γ or $corrRev$	0.7
correlation volatility	ψ or $corrVol$	0.85
driver 1-3 correlation	ρ_{13} or $drw13corr$	0.73
price-of-risk driver 1	λ_1	0.0
price-of-risk driver 2	λ_2	0.0
price-of-risk driver 3	λ_3	0.0

Table 13.1: Parameter values used for stochastic correlation model

It is interesting to see that the correlation for the stochastic drivers 1 and 3, denoted ρ_{13} or $drw13corr$, is calibrated to a magnitude of 70%. This is the same magnitude of correlation we empirically saw between spot price and risk-reversals.

Using the above values the model produces the following Black-Scholes equivalent implied volatilities.

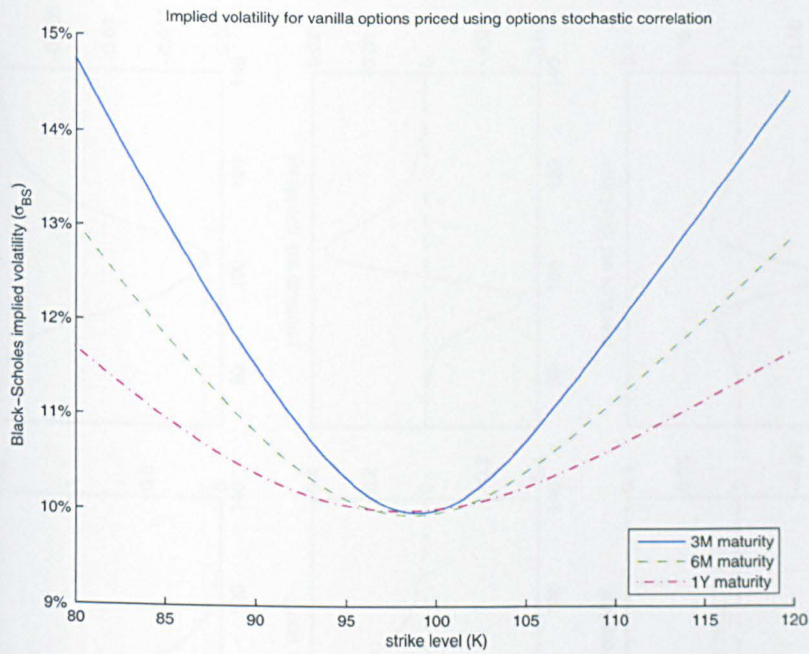


Figure 13.1: Volatility smile under stochastic correlation

The resulting volatilities displayed in Figure 13.1 show the same type of curvature that is produced by a pure stochastic volatility model.

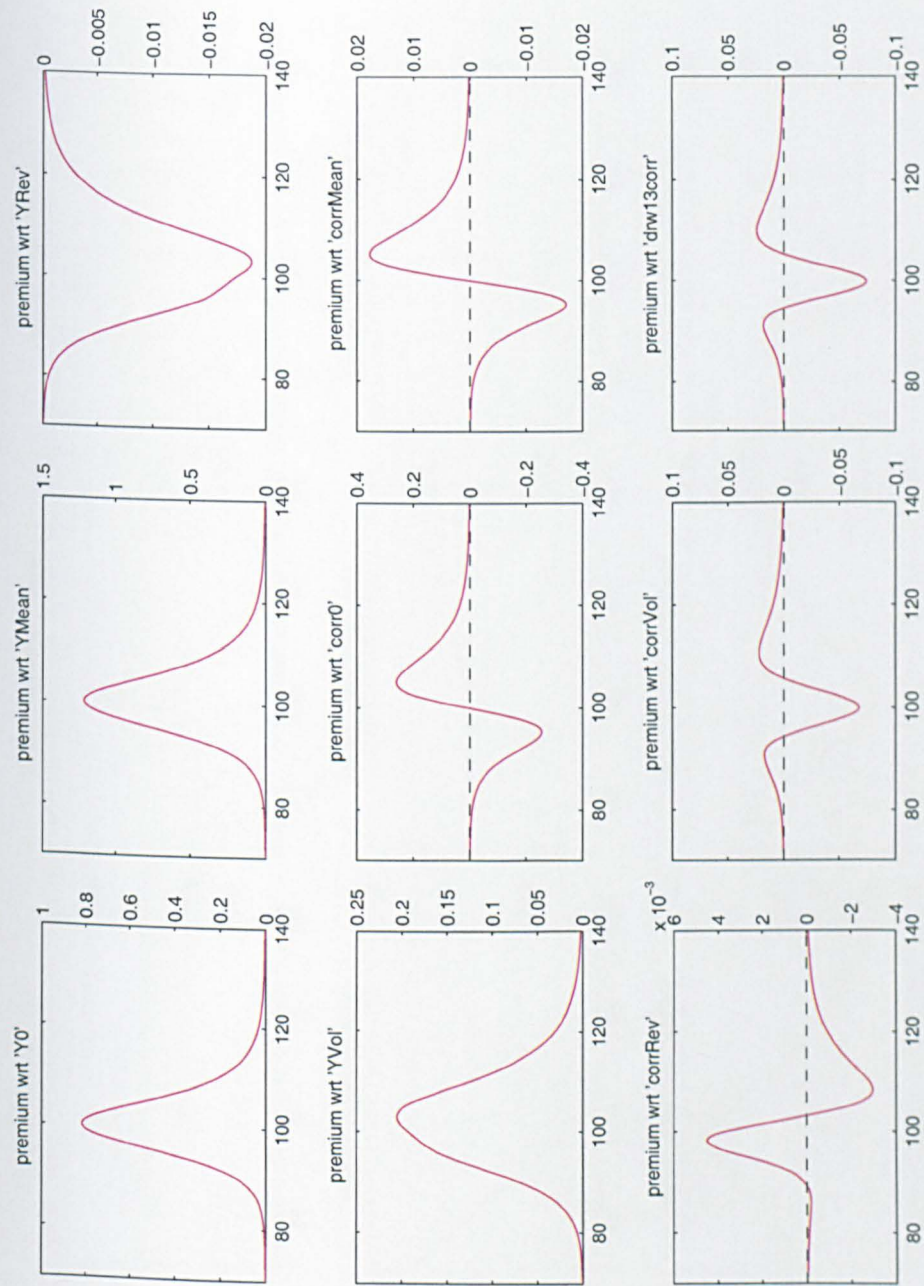


Figure 13.2: Parameter impact on vanilla option premium against strike level

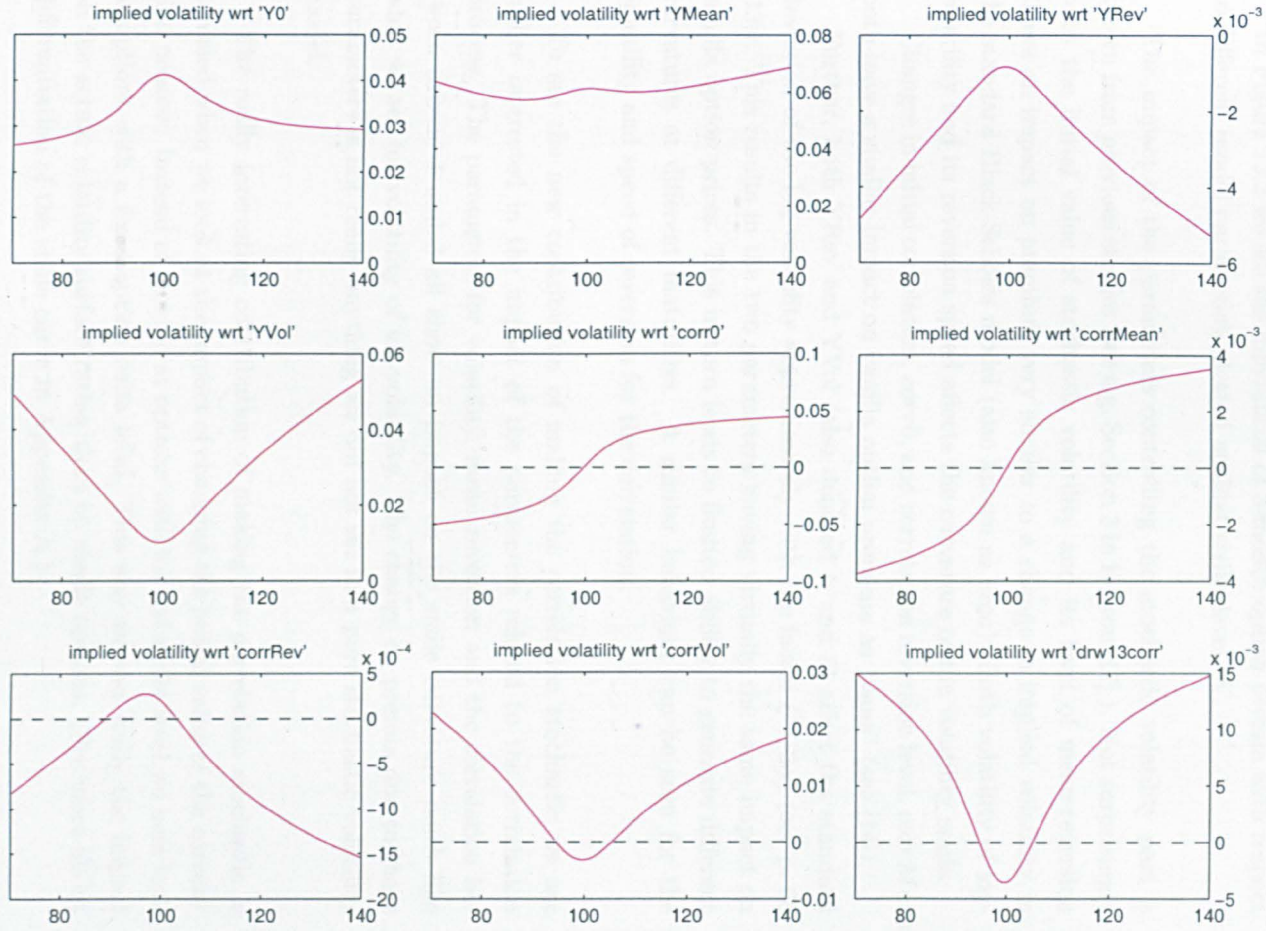


Figure 13.3: Parameters' impact on vanilla implied volatility plotted against strike level.

13.1. VANILLA OPTIONS

In Figure 13.3 we see the differential of 3-month option premia with respect to different model parameters plotted against strike level, K .

The impact of the parameters controlling the stochastic volatility part is known from previous studies (see e.g. Section 3 in Heston [17]). Not surprisingly both the initial value of stochastic volatility and its level of mean-reversion shows an impact on premium very similar to a change in implied volatility in the standard Black-Scholes model (also known as *vega*). Both volatility of log-volatility and its reversion speed affects the curvature of the volatility smile.

Changes in initial correlation, $corr0$, and correlation reversion level, $corrMean$, both have a similar impact on vanilla option premium as “vana” (see [14]).

Further, both $YRev$ and $YVol$ (also denoted α and ξ) affect the standard deviation of the log-volatility approximately with the factor $\xi^2/(2\alpha)$ (see p. 68 [13]). This results in the two parameters having virtually the same impact on vanilla option prices. This in turn leads to limited ability to generate different curvatures at different maturities. A similar behaviour can be seen for the volatility and speed of reversion for the correlation.

To see the new contribution of making the correlation stochastic we are rather interested in the impact of the parameters related to the correlation process. The parameters for volatility, mean reversion and the correlation between drivers 1 and 3 all show an impact on the smile curvature much like what we see for volatility of log-volatility. The change on premia due to those parameters is not really anything we did not see in a pure stochastic volatility model.

The really interesting contribution of making the correlation stochastic is revealed when we look at the impact of changing the initial value of the correlation process. Instead of looking at options with a fixed strike level we here look at options with a fixed option delta level. This way we can study the impact on the actual volatility surface rather than on single options. (See more about differentiation of the smile curve in Appendix A.)

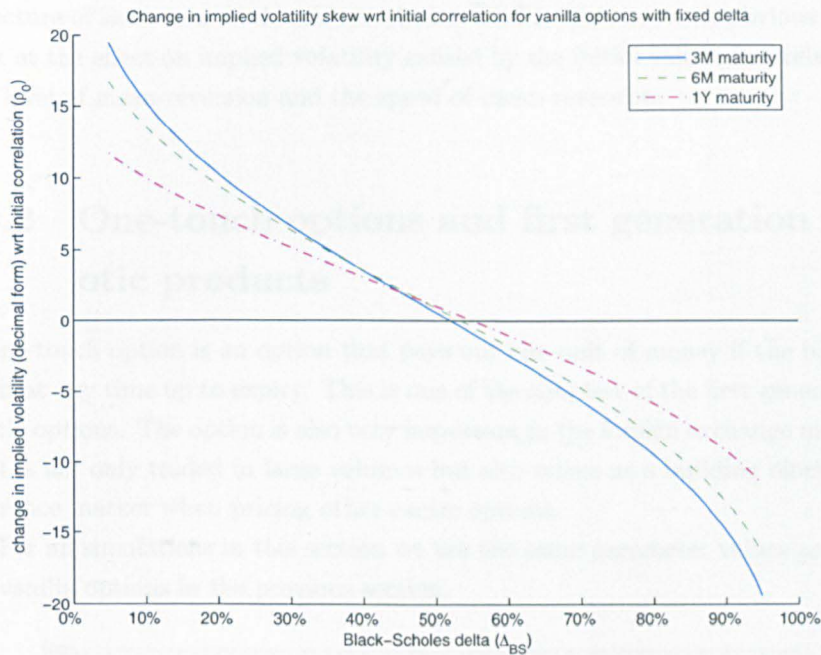


Figure 13.4: Change in volatility smile due to change in initial correlation value

The resulting impact of changing the initial value for the correlation process is shown in Figure 13.4. An almost identical graph describes the impact on the volatility smile if we instead change the long-term mean value for the correlation process, $\bar{\rho}$. The only difference is that the magnitude of change in volatility is the opposite and is instead larger for long-dated options than for short-dated options.

In a pure stochastic volatility model a similar type of change in skew can be controlled by the static correlation between spot price and volatility. The big difference in our case is that this value is not static but a stochastic process. It is, in some sense, this impact on the volatility skew that is the core of the entire model. As the correlation is made a process in a stochastic correlation model such a model will generate dynamics on the implied volatility surface of the same form as we see in Figure 13.4. By changing the static value for the correlation between driver 1 and 3, also denoted '*drw13corr*' or ρ_{13} , the dynamic change in volatility skew can be made correlated to the spot price.

We can finally make a comment on the term structure of the volatility skew generated by the model. In the same way that in a stochastic volatility model we can control the term structure of volatility by changing the initial value and mean-reversion properties of the local volatility, we can control the term

structure of skew in a stochastic correlation model. This becomes obvious if we look at the effect on implied volatility caused by the initial value of correlation, the level of mean-reversion and the speed of mean-reversion.

13.2 One-touch options and first generation exotic products

A one-touch option is an option that pays out one unit of money if the barrier is hit at any time up to expiry. This is one of the simplest of the first generation exotic options. The option is also very important in the foreign exchange market as it is not only traded in large volumes but also serves as a building block and reference marker when pricing other exotic options.

For all simulations in this section we use the same parameter values as used for vanilla options in the previous section.

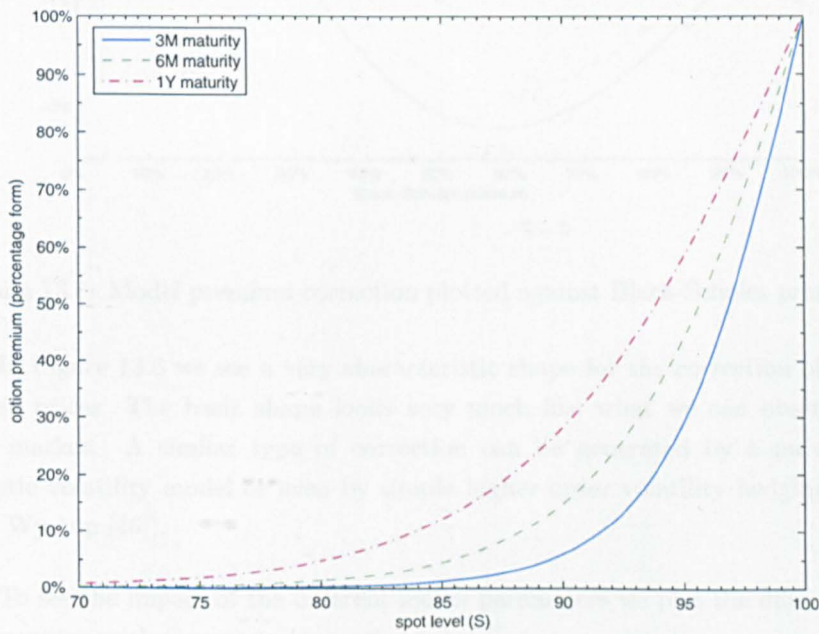


Figure 13.5: One-touch option premium against spot level

A popular way to present the premium of one-touch options is to compare it to the price generated by the Black-Scholes model. This is done by plotting the difference in option premium to the Black-Scholes option premium, using the at-the-money volatility. Rather than plotting the difference against strike level, as we can see in Figure 13.5, one usually plots it against the Black-Scholes

13.2. ONE-TOUCH OPTIONS AND FIRST GENERATION EXOTIC PRODUCTS

option premium. The benefit of this way of showing the correction is that one can easily compare the correction for different currencies, different maturities and different levels of volatility. The idea resembles the idea of plotting vanilla implied volatilities against the option's delta value.

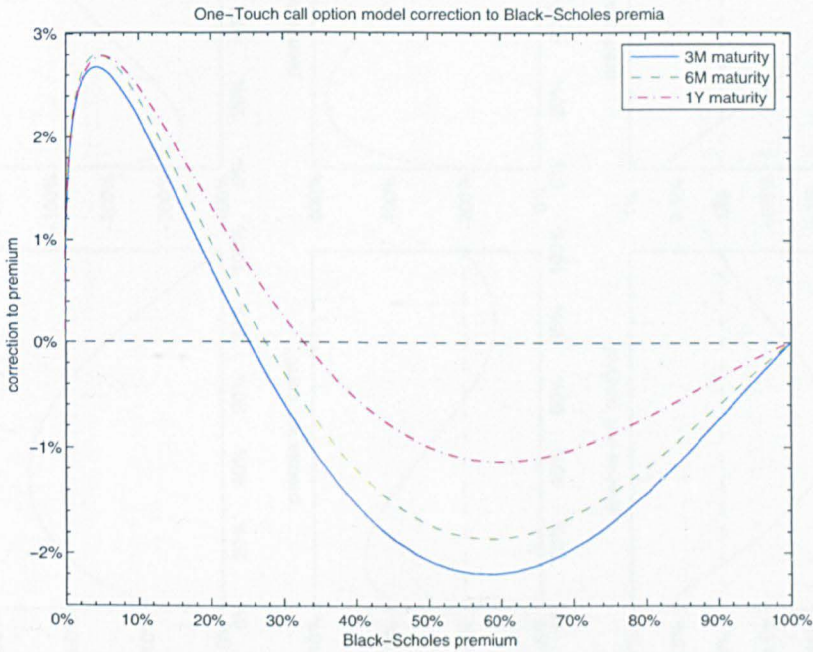


Figure 13.6: Model premium correction plotted against Black-Scholes premium

In Figure 13.6 we see a very characteristic shape for the correction of one-touch prices. The basic shape looks very much like what we can observe in the market. A similar type of correction can be generated by a pure stochastic volatility model or even by simple higher order volatility hedging (see e.g. Wystup [46]).

To see the impact of the different model parameters we plot the differential of premium with respect to parameter value.

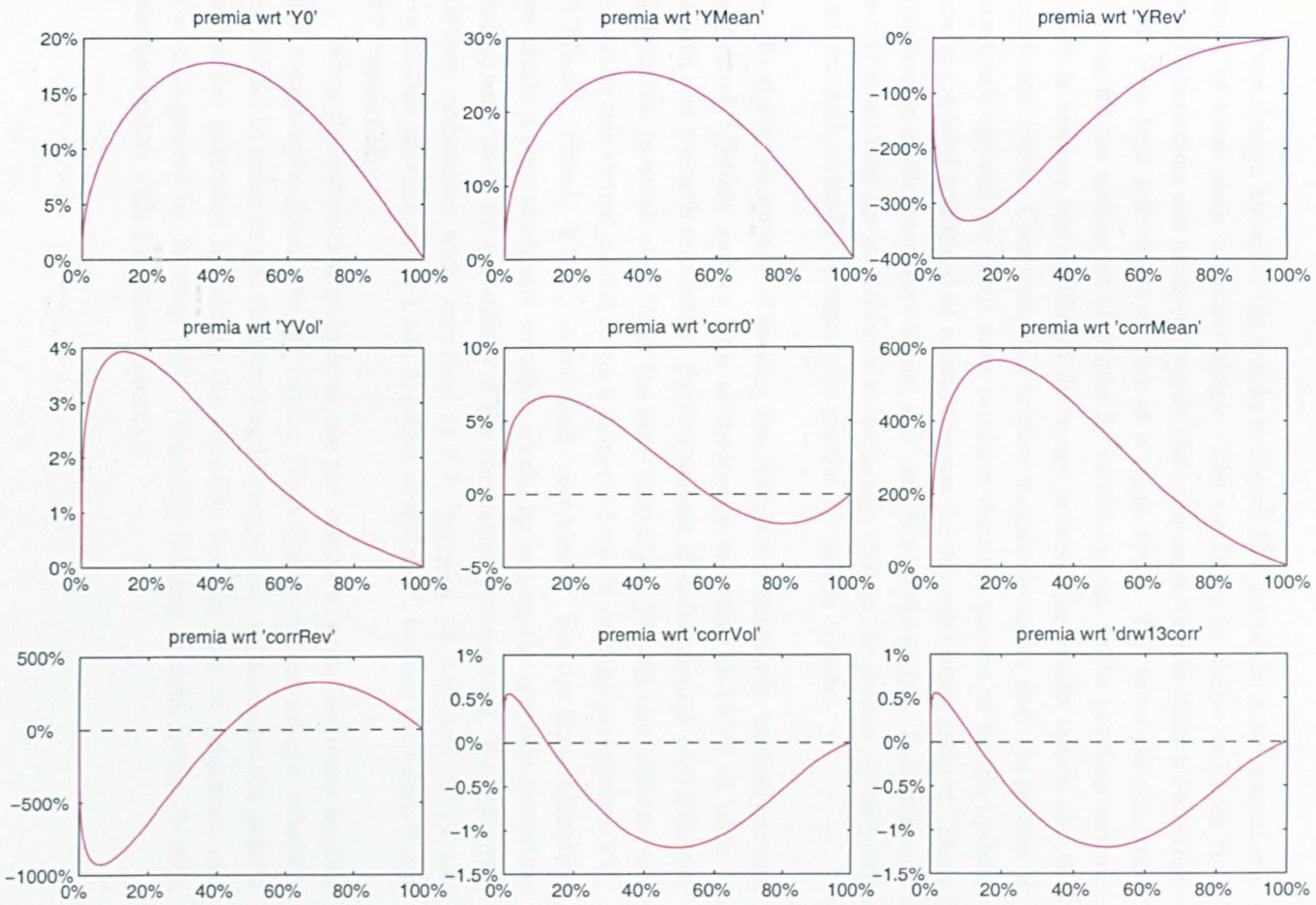


Figure 13.7: Parameter impact on model one-touch premium plotted against Black-Scholes premium

13.2. ONE-TOUCH OPTIONS AND FIRST GENERATION EXOTIC PRODUCTS

Even though not all of the graphs in Figure 13.7 are of the same magnitude, many of them show a similar shape. The similarity in shape indicates that some parameters can produce roughly the same correction as other parameters. This is at least true when we look at a single tenor. The same pattern could be seen for the similar set of graphs for vanilla options in the previous section. There is however one significant difference between the vanilla option and the one-touch option. Comparing the relative magnitude shows that the premia of one-touch options are much more sensitive than the premia of vanilla options are to the level and speed of mean-reversion for the correlation process. Those parameters are denoted '*corrMean*' and '*corrRev*' in the graphs. This difference would mean that the model could substantially change the premia of one-touch options with virtually no impact on premia for vanilla options.

To study the impact of making the correlation stochastic we first assume a stylised volatility surface with at-the-money volatility at 10.0% at both 1-month and 3-month maturities. Further assume 25-delta strangle at 0.25% and 5-delta risk-reversals at 0.0% for the same maturities. We will now calibrate our model to this virtual market using a numerical search over the parameters '*Y0*', '*YMean*', '*YRev*', '*YVol*', '*corr0*' and '*corrMean*'. For our first calibration we create a pure stochastic volatility model by setting the *corrRev* (reversion speed) and '*corrVol*' (volatility) of the correlation process to 0. We make two further calibrations with '*corrRev*' at 0.7, '*corrVol*' at 0.85 and choose the correlation between driver 1 and 3, called '*drw13corr*', to have the values 0 and 0.7 respectively.

When the calibration is done we use the results to price one-touch options and compare the premia to the regular Black-Scholes model using a volatility at 10.0%. In order to put the resulting in perspective, we also plot the premia correction generated by adding the volatility hedging cost of volgamma and vanna suggested by Wystup [46]. (Volatility hedging is done using 25-delta vanilla options with 3-months maturity.)

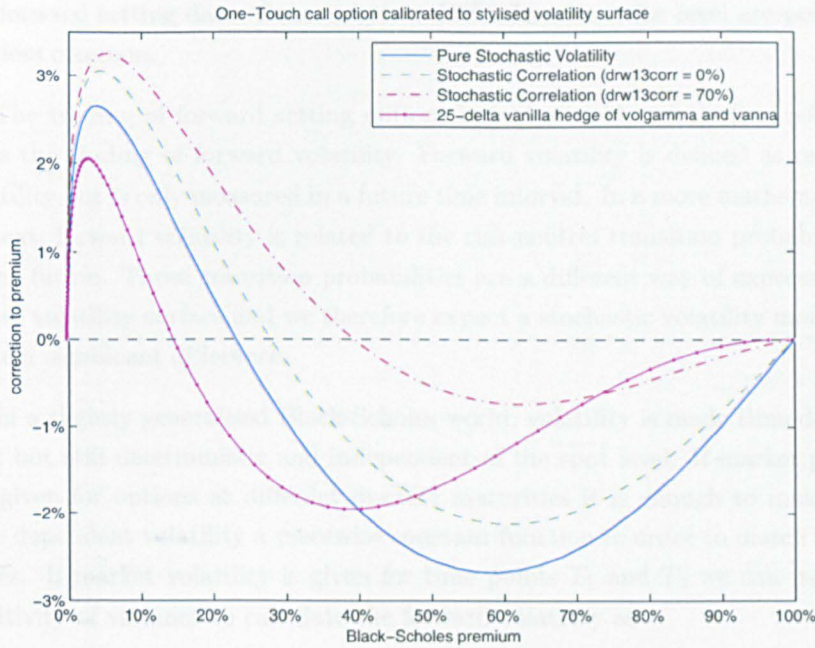


Figure 13.8: Model comparison for one-touch premium correction

The premium corrections for one-touch options under different models calibrated to the same volatility surface can be seen in Figure 13.8. We see that the general shape of higher premium for low-probability strikes and lower premium for high-probability strikes goes through the entire set of corrections. Among the three stochastic-correlation corrections the pure stochastic-volatility calibration generates the lowest premia. Making the spot-volatility correlation stochastic, by making its volatility non-zero, increases the option premium for all strike levels. Making the same correlation process positively correlated to the spot price process increases premium for all strike levels further.

We have essentially showed that by making the correlation between the spot price process and the volatility process stochastic we can generate different one-touch premia without changing the premia of vanilla options.

13.3 Forward setting options

A forward setting, or forward starting option, is a regular vanilla option where the strike level is set at a previously agreed future time point. This time point is called the forward setting date and must naturally be a date prior to the expiry date. In its most common form the option's strike level is set to the spot level at

13.3. FORWARD SETTING OPTIONS

the forward setting date. Other ways to determine the strike level are possible but less common.

The trading of forward setting options is fairly liquid and is often referred to as the trading of forward volatility. Forward volatility is defined as regular volatility but is only measured in a future time interval. In a more mathematical context, forward volatility is related to the risk-neutral transition probabilities in the future. Those transition probabilities are a different way of expressing a future volatility surface and we therefore expect a stochastic volatility model to make a significant difference.

In a slightly generalised Black-Scholes world, volatility is made time dependent but still deterministic and independent of the spot level. If market prices are given for options at different discrete maturities it is enough to make the time dependent volatility a piecewise constant function in order to match those prices. If market volatility is given for time points T_1 and T_2 we can use the additivity of variance to calculate the forward volatility as

$$\sigma(T_1, T_2) = \sqrt{\frac{\sigma^2(0, T_2)T_2 - \sigma^2(0, T_1)T_1}{T_2 - T_1}},$$

where $\sigma(T_A, T_B)$ is the volatility prevailing between times T_A and T_B .

This way of handling forward volatility can be slightly confusing if we have more than one traded option with the same maturity but with different strike prices. Due to the volatility smile such options almost always have different implied Black-Scholes volatilities and the natural question in such a case is what strike level, and hence volatility, to choose. The answer is far from trivial as the presence of a volatility smile contradicts the assumptions of volatility being independent of spot level made in the first place. However the general practice, at least for at-the-money forward setting options, is to only look at at-the-money options.

The pricing of forward-starting options is closely linked to the pricing of so called cliquet or ratchet options (Section 7.3 in Hakala & Wystup [14]). Those products belong to an area of ongoing research and as there is not a commonly agreed treatment of forward volatility in the presence of a volatility smile many banks have their own treatment. We shall not go further through existing methods on how to price forward volatility but instead look at the impact of stochastic correlation on such prices.

The general idea here is to calibrate our model to the vanilla market described by the volatility smile and then price forward starting options under

13.3. FORWARD SETTING OPTIONS

this calibration. Before we look at the impact of stochastic correlation we take a look at the simpler case of just stochastic volatility. Assume we look at a 3 month option starting 3 months from now and expiring in 6 months. We further assume the market to have an at-the-money implied volatility of 10.0%, 25-delta strangle of 0.25% and 25-delta risk-reversal at 0.0% at both the 3 month and the 6 month tenors. Using the above described technique we calculate the forward volatility to be 10.0%. We now want to calibrate our model to this stylized market.

In order to remove the stochastic correlation part we set both mean-reversion of correlation and volatility of correlation to 0.0. These choices totally remove the effect of the parameters for correlation mean-reversion level and correlation between the drivers 1 and 3. The six remaining parameters are chosen by a numerical search in order to match our six market values. The calibration produces a virtually perfect match to the chosen market values. Once the calibration is done we use the resulting model parameters to price our forward starting option. It turns out that the implied volatility of this forward starting option is 9.56%. We immediately note this number to be lower than the volatility of 10.0% we got when only looking at at-the-money options.

The same study is now repeated but with different values for the strangles. At both the 3-month and the 6-month tenors we have at-the-money volatility at 10.0%, 25-delta risk-reversal at 0.0%.

STR25 at both 3M and 6M	3M to 6M forward volatility
0.00%	10.00%
0.25%	9.61%
0.50%	9.31%
0.75%	9.06%
1.00%	8.86%

Table 13.2: Strangle impact on forward volatility under pure stochastic volatility

The results in Table 13.2 show a clear pattern. In a stochastic volatility model the forward volatility is monotonically decreased from the theoretical level of 10.0% as curvature of the volatility surface, here measured by the 25-delta strangle, is increased. The conclusion is that under given at-the-money volatilities a stochastic volatility model can be calibrated to either smile curvature or forward-volatility. The model cannot be calibrated to both of them at the same time. We will now show that this constraint is not true for a stochastic correlation model.

Assume the same stylized market as we did for the stochastic volatility case. Use the stochastic correlation model and fix the parameters 'corrRev' at 0.7,

13.3. FORWARD SETTING OPTIONS

'*corrVol*' at 0.85. and let '*drw13corr*' take values from -90% to $+90\%$. The remaining parameters are found via a numerical calibration to the market.

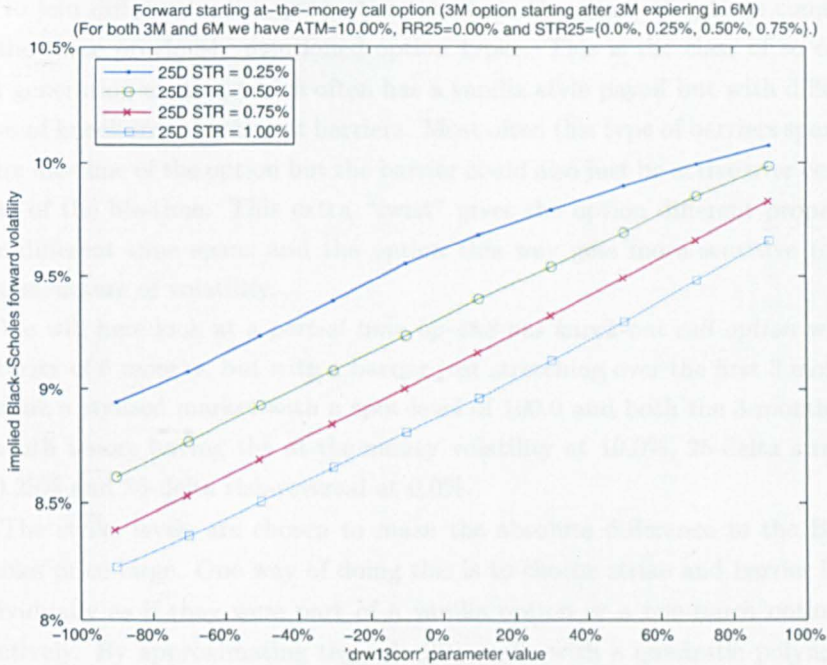


Figure 13.9: Impact of parameter '*drw13corr*' on implied forward volatility

The resulting forward volatilities are shown in Figure 13.9. We again see forward volatilities lower than the 10.0% suggested by a Black-Scholes model under time-dependent volatility. Along with this inheritance from the pure stochastic volatility model, we see a clear dependence on the correlation parameter '*drw13corr*'. A higher correlation between spot price and risk-reversal leads to a higher forward volatility and vice versa. This is an area where a stochastic correlation model shows a clear contribution.

We here clearly see a product where the vanilla market is not enough to calibrate the model. If we know the market price for a forward setting option we can use this price to calibrate the value of the '*drw13corr*' correlation parameter. In a less liquid market one could use an empirical value for this correlation and instead calculate the forward volatility.

13.4 First generation products with a twist

After having looked at three very typical and distinct option types we now try to join different option properties together. One class of options combines all the three previously mentioned option types. This is the class of so called first generation exotics, which often has a vanilla style payoff but with different types of knock-in or knock-out barriers. Most often this type of barriers span the entire life-time of the option but the barrier could also just be active over certain parts of the life-time. This extra “twist” gives the option different properties over different time spans and the option this way gets more sensitive to the term-structure of volatility.

We will here look at a *partial time up-and-out knock-out call option* with a maturity of 6 months, but with a barrier just stretching over the first 3 months. Assume a stylised market with a spot level of 100.0 and both the 3-month and 6-month tenors having the at-the-money volatility at 10.0%, 25-delta strangle at 0.25% and 25-delta risk-reversal at 0.0%.

The strike levels are chosen to make the absolute difference to the Black-Scholes price large. One way of doing this is to choose strike and barrier levels individually as if they were part of a vanilla option or a one-touch option respectively. By approximating the volatility smile with a quadratic polynomial in option delta, fitted to our stylised market data, we can easily find the strike level with the biggest absolute difference in premium. For a 6-month option this approximately turns out to be strike level 108.8 on the upside and 92.4 on the downside and both differences are positive. In order to maximize the total difference in premium we want to choose the barrier level to further increase the premium difference. We can do this by finding a barrier level with a low knock-out probability relative to Black-Scholes. By looking at our results in the previous section, we see that for the stochastic correlation model such a barrier level is found at a hitting probability of approximately 60%. With our market data this translates to a barrier level of approximately 102.6 for a 3-month one-touch option.

We again choose ‘*corrRev*’=0 and ‘*corrVol*’=0 for the pure stochastic volatility case and ‘*corrRev*’=0.7, ‘*corrVol*’=0.85 and ‘*drw13corr*’=0% for the first stochastic correlation case and ‘*drw13corr*’=70% for the second stochastic correlation case.

The numerical values of the partial time up-and-out knock-out option premia are compared to the theoretical Black-Scholes premium (see e.g. Section 2.10.3 in Haug [15]).

The obvious observation is that all calibrations result in a higher option premium than the Black-Scholes price. This is not surprising as the Black-Scholes

13.4. FIRST GENERATION PRODUCTS WITH A TWIST

model calibration	premium (upside strike, call)	difference to Black-Scholes
Black-Scholes	0.0141	± 0.0000 ($\pm 0\%$)
pure stochastic volatility	0.0336	+0.0195 (+138%)
stoch. corr. ($\rho_{13} = 0\%$)	0.0353	+0.0212 (+150%)
stoch. corr. ($\rho_{13} = 70\%$)	0.0263	+0.0122 (+86%)

model calibration	premium (downside strike, put)	difference to Black-Scholes
Black-Scholes	0.3642	± 0.0000 ($\pm 0\%$)
pure stochastic volatility	0.3809	+0.0167 (+5%)
stoch. corr. ($\rho_{13} = 0\%$)	0.3826	+0.0184 (+5%)
stoch. corr. ($\rho_{13} = 70\%$)	0.4150	+0.0509 (+14%)

Table 13.3: Model comparison for partial knock-out options

model does not take the volatility smile into account. What instead could be surprising is the similarities between the pure stochastic volatility calibration and the non-correlated stochastic correlation calibration.

The big difference appears first when the stochastic correlation is made correlated to the spot price. We here see a big difference in how the stochastic correlation affects the two different strike levels. The correlation makes the upside call option decrease in premium and the downside put option increase in premium.

We conclude that it is not the introduction of stochastic correlation itself that makes the largest impact on pricing. The largest impact comes from making the stochastic correlation correlated to the spot price.

Chapter 14

Model calibration to observed market prices

A severe problem associated with a model that does not replicate observed market prices is that it theoretically would mean an arbitrage opportunity. This in turn suggests repeated trading of this arbitrage opportunity, or at least taking the largest position allowed for by regulations, until market forces have removed the market mispricing. Even though a trader is never supposed to act in this way huge amounts of money have been lost in the market after this kind of behaviour. Few model builders are willing to shoulder such risk.

If exotic options are priced using a more advanced model than used for pricing vanilla options and the products are very different, would it really matter that the advanced model cannot fully reproduce the vanilla prices? Under regular conditions this might not be a very big concern. However problems might occur when the properties of the exotics become more extreme. For example, call and put options with a knock-out barrier must have the same price as a regular vanilla option if the barrier is pushed very far away from today's spot price. Such relations indicate that even if a model is only used for pricing exotic options it must be able to reproduce vanilla prices. This is not only a constraint in order to avoid arbitrage but is also a constraint in order to avoid inconsistencies and discontinuities in pricing. If the model does not calibrate to the market the user must always be very careful and know when to use the model and when not to.

Having concluded that we desire our model to be able to reproduce observed market prices we will now look at the calibration of the model. In general we need one constraint for each free parameter in a market model. Such a constraint could be a direct choice of a value for a parameter. The chosen

value could either be a pure guess or the result of statistical analysis. A more common type of constraint is a price of an option we want the model to be able to reproduce. The constraints of this type are less direct and parameter values must be found via a numerical search algorithm. If such a search finds more than one set of possible variables satisfying the given constraints we must make a choice based on further criteria. However it is more likely that no exact solution can be found at all. In this case the numerical search should be replaced by a numerical minimisation algorithm trying to minimise the error between prices given by the model and prices given by the constraints. A popular setup is to minimise the sum of squared errors.

Numerical minimisation can be a very computer intensive task as the minimisation algorithm needs to calculate option prices using a large set of different model parameters. If option pricing itself is very slow, as is often the case if we use a numerical pricing method, minimisation can take many hours to perform. It is obvious that an intelligent choice of minimisation algorithm is important. There are also many techniques available to speed up the minimisation procedure. One such technique is to reduce the numerical precision in pricing early in the minimisation process and then increase precision as time passes. If we had access to an analytic pricing formula the calibration process would be much quicker.

In our case we use an ADI finite differences scheme implemented in C++ for pricing and the MATLAB command `"lsqnonlin"` for minimisation over option prices. Minimisation over six to nine parameters takes between 10 minutes to 10 hours depending on precision and choice of options.

The presented stochastic correlation model has in total 12 parameters and three initial process states. The initial spot price, S , can be directly observed in the market and should not be used for calibration. If we assume constant and deterministic interest rates the same is also true for both the domestic and the foreign interest rates, r and q . With the exception of the three market prices of risk, all other parameters could in theory be found by statistical investigations. In reality this approach does only give us information about the past market history and gives us no real information about the future. This reasoning has an analogue in the Black-Scholes world where one most often calibrates against observed option prices to get an implied volatility rather than using past spot data to calculate the historical volatility. By calculating the implied volatility we make use of the aggregated information about the future incorporated by the market into the option price. In a similar way we could choose the 12 non-observable variables by a numerical calibration against 12 liquid option prices observed in the market.

It is however important to see the difference between calibrating to market prices and capturing market dynamics. In an ideal theoretical world with a market fully obeying the model the two things are the same. In reality they are not. We could use the 12 free variables to calibrate against 12 arbitrary options observed in the market. If a majority of those options are of the same type, the chances are that other options of this type also will be priced close to what we observe in the market by the model. In this case we have not really explained anything new, but only used the model to interpolate between the market prices we calibrated to.

To make true use of the model's potential we should calibrate against different types of options, each of which is dependent on different market dynamics. This way the model will have less accuracy in pricing specific types of options than if we had calibrated solely against the same specific type. On the other hand we could expect to get a similar, but possibly lower, accuracy over a whole range of different option types and hopefully also for types we did not calibrate against in the first place. This will of course not be true for options having significant sensitivity to other dynamics than what is covered by our model. Such options would not be correctly priced by any model not involving those other dynamics in its fundamental construction.

Chapter 15

Practical use of the model

15.1 Products used for hedging new exposure

At the foundation of the theory of option pricing is the concept of replication. This means the ability to meet a contract obligation with no risk by dynamically trading the underlying assets. In a Black-Scholes world this is made possible by a risk-free bank account and a liquid market in the underlying asset. But what are the underlying assets in a stochastic correlation model?

The problem is not new and similar problems arise for other models based on e.g. stochastic volatility and jump-diffusion processes. In the same way that the problem is solved in those cases, hedging the factors in a stochastic correlation model would include trading in other options. But what other options are suitable?

From a more mathematical perspective the risk exposure comes from the three independent driving factors rather than from the underlying market processes driven by them. This way of looking at the problem is important when we handle the market price of risk. At the stage of hedging, the market price of risk is assumed to be known so we can ignore this issue and only care about the three market processes — spot, log-volatility and correlation (see Appendix 19B in Hull [18]).

A simple method to hedge out this exposure is to pick as many different assets as we have factors. In our case there are three driving factors and hence we need three assets where one asset, naturally, is the underlying asset. In theory those assets could be options chosen more or less randomly. Once the assets are chosen we calculate their exposure to the underlying drivers. We now want to choose the amounts of the different assets that result in all exposure to

15.1. PRODUCTS USED FOR HEDGING NEW EXPOSURE

the driving factors being cancelled out. The problem forms a system of linear equations, which can be solved by simple linear algebra. Unless we have been really unlucky with our choice of assets, resulting in the resulting matrix being singular, the system has a unique solution.

If we lived in an ideal world with a liquid market without transaction costs, the above method would work fine for dynamic hedging. This is far from reality and other aspects need to be taken into account. Such aspects could be transaction costs and stability of the hedge.

By stability we here mean that the exposure to a factor does not change much when the market moves. In mathematical terms this means low second order derivatives with respect to factor values. This includes both curvature and cross derivatives.

The aspect of transaction costs will be further handled in the next section but in short it means the need to consider the cost of rehedgeing by looking at the bid/ask-spread of different assets.

Not surprisingly, a natural choice of asset turns out to be the underlying asset itself. It satisfies both the above criteria as its second order derivatives are zero and the transaction costs are virtually also zero. Hedging the volatility and correlation risk is not as straight forward as here we need to trade options. In order to keep the transaction costs down it is desirable to only use vanilla options. We get a quite good view of what type of vanilla options we need by looking at the premium sensitivity to the model parameters 'Y0' and 'corr0' in Figure 13.3. Without going into too much detail here we sketch the outlines of possible hedging trades.

To take a position in volatility we can buy options with more or less any strike level we want to. But if at the same time we want to minimize the exposure that the trade has to changes in correlation, we must either buy at-the-money options, which has no exposure to correlation, or make two trades with cancelling exposure to correlation. As the exposure to correlation has different signs on each side of the at-the-money level, such a cancelling trade could be to buy two options roughly equally spread around the at-the-money level.

In order to take a position in correlation we must buy out-of-the-money options, but doing so will also give an exposure to volatility. Once again we can make use of the difference in sign for the correlation exposure at different strike levels and solve this problem by a cancelling trade. By buying an option with a strike level above the at-the-money level and selling an option with a strike level roughly equally far below the at-the-money level, we cancel the volatility exposure but have a positive exposure to correlation.

15.2. TRADER WILL STILL ONLY HEDGE DELTA AND VEGA

Any residual exposure to the underlying asset generated by hedging volatility and correlation is simply hedged in the same way as ordinary exposure to the underlying asset is hedged. By hedging of the exposure to the underlying asset and the use of put-call parity (see Section 7.4 in Hull [18]) we can turn the above volatility hedging trade into a *straddle*¹ or a *strangle*². In the same way we can turn the correlation hedging trade into a *risk-reversal*³.

The conclusion is that trading-structures that are often used in a Black-Scholes framework with a volatility surface, also work fine for a stochastic correlation framework.

If the trader is a market maker, a more realistic attitude to the hedging of volatility risk and correlation risk using a stochastic correlation model would probably be closer to today's attitude to the hedging of implied volatility using a Black-Scholes based environment. Exposure to implied volatility, known as *vega*, is most often not really hedged out by entering special trades. Instead the trader's exposure is fed into the prices the trader shows to customers in the everyday trading flow. If a trader is long volatility he or she can lower the bid and ask prices shown to customers. This will, in an active market, lead to the trader's customers, on average, buying more options than they are selling. If the trader instead is short volatility, prices are increased, leading to an opposite customer behaviour. Once the vega exposure reaches an acceptable level, prices are returned to be centred around a more neutral price level.

Specific trades can sometimes be undertaken if a price adjustment does not lead to the desired trading flow or if a large exposure is the result of an otherwise advantageous trade. In such a case the trader will try to hedge the exposure with a trade that "looks cheap" at the current market levels.

15.2 Trader will still only hedge delta and vega

How would a new model like a stochastic volatility model fit in with an existing trading framework?

Generally financial houses trading in derivatives are, and should be, very conservative to changes in pricing models. This is a natural, but also regulated, attitude to model risk and the human factor. New models do not only mean that large parts of the software infrastructure must be rewritten but also, at least initially, increase the risk of human errors.

¹ A long at-the-money call option and a long at-the-money put option

² A long out-of-the-money call option and a long out-of-the-money put option

³ A long out-of-the-money call option and a short out-of-the-money put option

15.2. TRADER WILL STILL ONLY HEDGE DELTA AND VEGA

However on the trading floor there is a quite large margin for traders to use their own judgment and market experience to divert from the standard model if they do not find the resulting prices satisfying. A trader could use a non-standard model to both price and hedge certain products as long as the total value and risk of the entire portfolio do not deviate too much from what is given by the old model.

This is a classic clash between theoretical quantitative work and practical trading. There must be a good incentive for a trader to divert from the way of working and thinking that he or she is used to. In a trader's terms this primarily means "making more money" or secondarily "being exposed to less risk".

The chances are that a more advanced model will be used only for pricing certain options where the traditional models have proved inadequate. But once the option is traded, it will disappear in the trading book and be handled by the traditional models. But would this matter?

Models today are often based on the Black-Scholes model with a generalised implied volatility. Vanilla options are often priced under a volatility surface and exotic options are often priced under a volatility term structure. For both types changes in the at-the-money volatility can easily be calculated and results in a *volatility risk* also known as *vega*. This risk is very important but can easily be hedged by trading other options

For options priced under a volatility surface the impact of changes in the volatility skew can be calculated by a simple perturbation of the surface. In this framework the hedging of this risk is restricted to the trading of other options priced under a volatility smile. This might not be a problem as those options are often very liquid. But what is a problem is the fact we cannot handle and control the impact on changes in volatility skew for options not priced under a volatility surface. This was one of the very reasons for the introduction of the stochastic correlation model in the first place.

So essentially we could conclude that for options priced under a volatility smile there will be no difference, whereas for options not priced under a volatility smile we end up with more accurate prices but with the same unhedged exposure to the volatility skew.

In a Black-Scholes world dynamic hedging is not only keeping us insensitive to changes in the underlying asset. As a secondary effect dynamic hedging will continuously gain or lose wealth due to constant reheding where the option premium has curvature. In mathematical terms this is a result of what is called *quadratic variation*. Even though we, in reality, cannot rehedg truly continuously, the underlying asset can be traded at very high frequency if needed. In the real world this is called *gamma trading* and it results in gains and losses

very much like the theoretical quadratic variation.

Quadratic variation has a similar impact on all driving market factors in a stochastic correlation model, not only the underlying asset. If we do not dynamically trade all driving factors in the model we will not only have a non-hedged exposure, but we will neither pick up the gains or losses stemming from quadratic variation involving those factors. In theory this would mean that a traded option, or a portfolio of options, priced using a stochastic correlation model for pricing but hedged using a Black-Scholes model, will systematically end up with a position diverging from the delivery obligation. The position will not only deviate due to unhedged exposure, but also the expected value of the portfolio will deviate due to lack of gains and losses from quadratic variation.

Leaving the theory behind one must remember that dynamic hedging of neither volatility nor correlation is possible due to the large bid/ask-spread for the hedging products. Dynamic hedging of those factors would simply bleed wealth at a totally unacceptable rate. This would at least be true for a single product, which will be covered briefly in the next section.

15.3 Bid/ask spreads and transaction costs

One of the fundamental assumptions in the Black-Scholes model is the assumption of a friction-free market. This assumption does not fully agree with reality where friction, among other things, can be observed as limitations in liquidity and the presence of transaction costs. Transaction costs can have many components but the important part is the marginal cost of a transaction. A more or less negligible fraction of the marginal cost is related to different operational costs but the dominating part of transaction costs comes from the difference in buy and sell prices on the market. This difference is called the *bid/ask-spread* and corresponds to the so called “round-trip” transaction cost, which is the loss made by buying an asset at the higher ask price and immediately selling the asset at the lower bid price. If we imagine a neutral price in the middle of the bid/ask-spread, this would mean a theoretical loss of half the spread every time we buy or sell the asset.

For many of the large foreign exchange markets the BID/ASK-spread is only a few *pips*⁴ wide. This corresponds to a relative bid/ask-spread, and hence transaction cost, in the order of 0.01% which is virtually negligible for practical frequencies of reheding.

⁴A “pip” is the smallest allowed precision a traded asset can be marked in on a market. This is often the 4th decimal place.

15.3. BID/ASK SPREADS AND TRANSACTION COSTS

In Leland [26] a simple approach to estimating the impact of transaction costs on option pricing is suggested. The approach is based on the expected losses due to transaction costs and is incorporated in basic option pricing as a modification to the volatility term. It is further assumed that no extra change is made for the increased uncertainty in hedging. For the Black-Scholes market model such a correction is⁵

$$\hat{\sigma}^2 = \sigma^2 \left(1 + \frac{k}{\sigma} \sqrt{\frac{2}{\pi \Delta t}} \right).$$

Here σ is the original volatility and $\hat{\sigma}$ the modified volatility corrected for transaction costs. The “round-trip” transaction cost, in relative terms, is denoted k and Δt represents the time interval between rehedges.

A similar correction would be necessary for all driving factors. This means that in a stochastic correlation model we need to correct the volatility of log-volatility and volatility of correlation as well.

For vanilla options the bid/ask-spread for a liquid option market is, measured in implied volatility, of the magnitude of 0.10% – 1.00%. Even though this spread is not necessarily a result of it, this is very close to the correction Leland proposed based on an at-the-money volatility of 10.0% and the above mentioned spread for the underlying asset. Translating this into a premium spread for the options themselves, a 25-delta option gets a relative spread in the region of 1% – 10%. This is a few hundred times larger than the spread of the underlying asset.

This does not only suggest less frequent rehedges of volatility and correlation compared to the frequency of hedging for the underlying asset. It also suggests that the larger spread should also indicate larger bid/ask-spreads for options sensitive to volatility of log-volatility or volatility of correlation. One can in general terms say that convexity to a traded market factor is always associated with rehedges and is therefore penalised by transactions costs regardless of whether the convexity is positive or negative.

It is worth mentioning that here we have not considered the fact that the log-volatility and correlation in turn have a correlation to the underlying asset and, at least partly, could be hedged by the underlying asset itself. A high such correlation would hence result in a lower transition cost for hedging those factors. This area is however left for further research.

It is also in place to remind the reader that the impact of transaction costs, which are related to the *second* order derivatives of premium with respect to

⁵ Before I heard of Leland's work I derived the same volatility correction on my own.

15.3. BID/ASK SPREADS AND TRANSACTION COSTS

the market factors, must not be confused with the market-prices-of-risk, which are related to the *first* order derivatives of premium with respect to the market factors and have a different source.

It is important to remember that the above theoretical framework applies only to the case where a sole option is being replicated only with instruments traded on the open market. In reality a trader holds a whole book of options that most often has, at least some, net convexity to all involved market factors. If a new option holds an opposite convexity to one or more market factors, the book's net convexity towards those factors will decrease by adding this option. Trading the new option will actually reduce the need for future rehedgeing and hence reduce the expected loss from transaction costs. As a result the trader is willing to make the trade at a less favourable premium than the above framework would suggest.

The reasoning with hedging costs netting each other out due to opposite convexity is not always straightforward as we must consider the change in net convexity up to the option's maturity date and not only at the local time point.

Chapter 16

Further areas of research

16.1 Term structure of variables

Despite the reasonably large number of free parameters in the model, it will never be able to fully calibrate to a liquid market with options traded at many strike-levels and with many maturities. As discussed earlier, the ability to calibrate to the market is important in order to avoid pricing problems.

In order to calibrate to the market we need to generalise our model. One way of doing this is to make the model parameters functions of time. The method is widely used to generalise the original Black-Scholes model where this is almost always done by making implied volatility a piecewise constant function of time. For every new liquidly traded tenor, a new piece is added to this function to match observed market prices.

As we have seen, a stochastic correlation model is not capable of generating different curvature of the smile, or equivalently strangles, at different maturities. In order to do so we could make the parameters responsible for this curvature, e.g. volatility and reversion speed of log-volatility, a function of time. Such a generalisation would also work for a pure stochastic volatility model.

In the original setup of a stochastic correlation model, the at-the-money volatility and the volatility skew have a simple term structure generated by initial value, reversion level and reversion speed of the volatility and correlation processes. This term structure is of a very primitive nature and does not allow for more than fitting short-term and long-term values. In order to have a finer precision in time also those parameters must be given a more explicit term-structure. This could for example be done by using constant values for process volatility and reversion speed but making the reversion-level a function of time.

By generalising the model we increase the number of free parameters in

order to meet the increase in the number of price constraints. Such an increase will also increase the number of parameters to calibrate. If we try to calibrate all variables at once we will have to calculate prices for all constraints we try to match for each numerical iteration. This would be a very time consuming procedure and finding efficient methods for calibrating would be important. An alternative way to calibrate the entire system at once could be to first calibrate parameters describing early parts of the term structures and then successively work our way through all the parameters describing the later parts of the term structures. The benefit stems from the fact that for each part of the calibration we only need to price a smaller set of price constraints. Even though each partial calibration would be quicker the downside is that we would have to run a larger number of calibrations. We leave for future research to decide which of the two methods is quicker.

16.2 Local parameters

An further generalisation of the above mentioned term-structures would be to make certain parameters local. Local parameters means that parameters are a function of both time and space. The most well known example is the local-volatility model suggested by Derman & Kani [10].

In a stochastic correlation volatility model the most general parameters would be functions of time, spot price, volatility level and correlation level. We do not necessarily need to use this enormous number of degrees of freedom. In the universal volatility model [29], volatility is made a local parameter as a function of time and spot level not using the possibility of also making it a function of volatility level. We mentioned earlier that local volatility was introduced to give the volatility surface a certain type of dynamics. Apart from this rather special reason to make parameters local, the main reason for its introduction is to make the model fully calibrated against observed market prices.

As a stochastic correlation model already has all the desired dynamics of the volatility surface, this will not require the introduction of local parameters. In our desire to match observed market prices we must remember that there are only a limited set of liquidly traded options. A traditional local volatility model needs access to a continuous set of option prices and hence the volatility surface must be interpolated. As we showed in the previous section it would be enough to make parameters a function of time in order to match the limited set of liquidly traded variables. The conclusion is that there are no obvious reasons to use local parameters in a stochastic correlation model.

If despite all this we choose to use local parameters, we have to face the

problem of model calibration. We would in such case have to interpolate the volatility surface to make prices continuous. The number of free parameters makes calibration very problematic and suggests that methods other than numerical minimisation over the set of free variables must be used. Out of curiosity we can mention that many financial institutions have developed more or less efficient methods for calibrating advanced models, but are keeping the details of those methods secret to maintain a competitive edge.

16.3 Jumps and events in all three dimensions

As in all markets almost nothing can be predicted and the market is sometimes subject to large shocks. Even though a standard diffusion process can generate large movements in short time spans, in reality we see larger and more frequent shocks than would be predicted by a diffusion alone. One way to introduce shocks to the market is to add a jump component to the diffusion. Such a component is often be modelled by a Poisson process. A famous example is the model suggested by Merton [31] where the ordinary Black-Scholes model is generalised by the introduction of jumps in the spot price. If our model has further dimensions than spot level, there is nothing that says that we cannot also have jumps in those dimensions. In a stochastic correlation model this would mean jumps in both volatility and correlation.

16.4 Cross currencies

An intrinsic nature of the foreign exchange market is the freedom to move wealth between different assets. In more direct words this means the freedom to exchange money from one currency into another. This results in many cross-currency exchange rates between less liquid currencies derived from more liquid exchange rates¹. This way derivatives could be dependent on an illiquid rate that is still welldefined via other liquid rates. This is true for multicurrency products like e.g. basket options. As there is rarely a market in cross-volatilities, hedging can be impossible.

Another problem is that it is hard to take smile effects into account for cross-currencies. By setting up a model for two liquid currency pairs we, at least in theory, could make the cross-currency a well defined process. The idea is that if the two original processes can reproduce the correct smile properties,

¹An example of how a cross-rate could be calculate is;
 $CADSEK = EURSEK / (EURUSD \times USDCAD)$

16.4. CROSS CURRENCIES

the derived cross-process would be a better model of reality. It would probably also give a better idea about how to hedge the dynamics of the illiquid cross volatility surface.

The problem lies not only in the large number of free model parameters, but also in the fact that the cross-rate itself would introduce further parameters via correlations between the involved driving processes. Finding values for those parameters will probably be very hard if the cross-rate is illiquid in the first place.

Appendix A

Differentials against option delta

We know that to get an overview of a volatility surface (associated to vanilla options) it is sometimes convenient to plot premium or volatility against the option's theoretical Black-Scholes delta value. This makes it easier to compare results for sets of options with different properties like time to maturity or the spot level of the underlying. Also much of the intuition and experience about changes in the volatility surface are related to values of option delta rather than absolute levels of strike. In some sense we can see this way of looking at options as if we used a different coordinate system. When studying the impact of different model parameters on the volatility surface and its dynamics it is desirable to show the results in this other coordinate system.

It might be in its place to mention that this coordinate system is not very useful and even counter-intuitive when it comes to looking at already traded options or other options with fixed strike levels. The problem lies in the fact that while an option will have the same absolute strike level during its life time, the delta of the option will not stay the same. As a matter of fact this delta will change as soon as the premium of the option changes. Having knowledge about and keeping all other parameters fixed, it is enough to know an option's strike level, K , and its premium, Π , to calculate its delta. In practise this means finding the option's implied Black-Scholes volatility and then calculating its Black-Scholes delta value.

It is the connection between the option's premium and its delta value that makes working in coordinates of delta less straight forward. A point with a specific delta value in the coordinate system will represent different options when the system changes. This way of looking at surfaces is related to the area

of differential geometry and the following should not be looked on as a formal handling of the matter.

We want to study the impact of variable, a , on the option property, P , using a Black-Scholes delta coordinate. If we assume all other parameters but the variable, a , and strike level, K , in the system to be fixed we get $P = P(a, K)$. The property, P , could for example be the option premia or its corresponding Black-Scholes implied volatility calculated using some numerical scheme. We assume that we have access to both $\partial P/\partial a$ and $\partial P/\partial K$. Using the chain rule for derivatives the impact on P of variable a can now be written as

$$\frac{dP}{da} = \frac{\partial P}{\partial a} + \frac{\partial P}{\partial K} \frac{dK}{da}.$$

In a strike price coordinate system we have $dK/da = 0$ and the system therefore contains nothing unknown to us. However in delta coordinates this is no longer true and dK/da is unknown. Instead of keeping the strike level, K , constant, the Black-Scholes delta value, Δ_{BS} , must be kept constant, or equivalently, have a zero derivative with respect to variable a . Again using the chain rule, this constraint can be written

$$0 = \frac{d\Delta_{BS}}{da} = \frac{\partial \Delta_{BS}}{\partial a} + \frac{\partial \Delta_{BS}}{\partial K} \frac{dK}{da},$$

and can be used to substitute the unknown dK/da in the original expression. This substitution gives

$$\frac{dP}{da} = \frac{\partial P}{\partial a} - \frac{\partial P}{\partial K} \frac{\partial \Delta_{BS}}{\partial a} \left(\frac{\partial \Delta_{BS}}{\partial K} \right)^{-1}.$$

As the expression for the Black-Scholes delta, and hence also $\partial \Delta_{BS}/\partial K$, is known, only the factor $\partial \Delta_{BS}/\partial a$ in the above expression is unknown. This problem can be solved as long as there is a known relation between the property, P , and option premium, Π , and strike level, K . If the property we look at is in fact the option premium itself this relation is trivial. Using such a relation we can write

$$\frac{\partial \Delta_{BS}}{\partial a} = \frac{\partial \Delta_{BS}}{\partial \Pi} \frac{\partial \Pi}{\partial P} \frac{\partial P}{\partial a}.$$

If we choose our property, P , as the Black-Scholes implied volatility the factor $\partial \Pi/\partial P$ becomes the inverse of the Black-Scholes vega. Putting it all together and explicitly writing the parts we know, gives the general following formula describing the change in property, P , with respect to an arbitrary vari-

able, a , when keeping the Black-Scholes delta value constant.

$$\begin{aligned}\frac{dP}{da} &= \frac{\partial P}{\partial a} - \frac{\partial P}{\partial K} \frac{\partial \Pi}{\partial P} \frac{\partial \Delta_{BS}}{\partial \Pi} \frac{\partial P}{\partial a} \left(\frac{\partial \Delta_{BS}}{\partial K} \right)^{-1} = \\ &= \frac{\partial P}{\partial a} \left(1 - \frac{\partial P}{\partial K} \frac{\partial \Pi}{\partial P} \frac{e^{rT} d_2}{n(d_2)} \right)\end{aligned}$$

where $d_2 = \frac{\log(S/K) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}}$ and $n(\dots)$ is the standard normal density function.

Appendix B

ADI engine (C++)

We here present the C++ code used when pricing vanilla options under the stochastic volatility model. Model parameters and output data are passed through reading and writing text files. This particular implementation is based around a Microsoft Windows[®] system but would only need minor changes to run on other platforms.

```
#include<math.h>
#include<stdio.h>
#include<stdlib.h>
#include<iostream>
#include<fstream>
#include<iomanip>
#include"StochCorrADI_LogOU_Vanilla.h"
using namespace std;
int main(int argc, char* argv[])
{
    if(argc < 2+1)
    {
        cout << "Function takes two arguments. (DataInFile and DataOutFile)" << endl;
        system("pause");
        return 0;
    }
    ifstream file;           // Define file handler for input and output
    ofstream outFile;        // Define file handler for output
    char tmpStr[200];         // Temporary string buffer used when reading in-data
    file.open(argv[1], ifstream::in);
    if(!file)
    {
        cout << "Error opening in-data file." << endl;
        system("pause");
        return 0;
    }
    // -- Product, model and method specific values -----
    file.getline(tmpStr,200); double K = atof(tmpStr);
    file.getline(tmpStr,200); double S = atof(tmpStr);
    file.getline(tmpStr,200); double T = atof(tmpStr);
    file.getline(tmpStr,200); double r = atof(tmpStr);
    file.getline(tmpStr,200); double q = atof(tmpStr);
    file.getline(tmpStr,200); double Y0 = atof(tmpStr);
    file.getline(tmpStr,200); double YMean = atof(tmpStr);
    file.getline(tmpStr,200); double YRev = atof(tmpStr);
    file.getline(tmpStr,200); double YVol = atof(tmpStr);
    file.getline(tmpStr,200); double corr0 = atof(tmpStr);
    file.getline(tmpStr,200); double corrMean = atof(tmpStr);
    file.getline(tmpStr,200); double corrRev = atof(tmpStr);
    file.getline(tmpStr,200); double corrVol = atof(tmpStr);
    file.getline(tmpStr,200); double drw13corr = atof(tmpStr);
    file.getline(tmpStr,200); double lambda1 = atof(tmpStr);
    file.getline(tmpStr,200); double lambda2 = atof(tmpStr);
    file.getline(tmpStr,200); double lambda3 = atof(tmpStr);
    // -----
    file.getline(tmpStr,200); long N_t = atol(tmpStr);
    double min_t = 0;
    double max_t = T;
    long N_x[3];
    file.getline(tmpStr,200); N_x[0] = atol(tmpStr);
    file.getline(tmpStr,200); N_x[1] = atol(tmpStr);
    file.getline(tmpStr,200); N_x[2] = atol(tmpStr);
    double min_x[3];
```

```

file.getline(tmpStr,200); min_x[0] = atof(tmpStr);
file.getline(tmpStr,200); min_x[1] = atof(tmpStr);
file.getline(tmpStr,200); min_x[2] = atof(tmpStr);
double max_x[3];
file.getline(tmpStr,200); max_x[0] = atof(tmpStr);
file.getline(tmpStr,200); max_x[1] = atof(tmpStr);
file.getline(tmpStr,200); max_x[2] = atof(tmpStr);
file.getline(tmpStr,200); long startNode = atoi(tmpStr); // Initial node to write to out-data file
file.getline(tmpStr,200); long numNodes = atoi(tmpStr); // Number of nodes to write
file.getline(tmpStr,200); long stepNode = atoi(tmpStr); // Index step between nodes
file.close(); // Close in-data file
// -----
const long d = 3; // Number of dimensions (excluding time)
const long nCross = d*(d-1)/2; // number of cross derivatives
const double theta = 0.5; // Degree of implicitness
// -----
int i,j,k,l,m,n;
double dt = (max_t-min_t)/N_t; // Calculate time step
double *dx = new double[d]; // Calculate step size (space dimensions)
for(i=0; i<d; i++) dx[i] = (max_x[i]-min_x[i])/(N_x[i]-1);
long gridSize = 1;
for(i=0; i<d; i++) gridSize *= N_x[i]; // Calculate total number of grid points
// Allocate memory for grid points
double *u = new double[gridSize]; // Initial values in step
double *lhs = new double[gridSize]; // Left-hand side of linear equation
double *rhs = new double[gridSize]; // Right-hand side of linear equation
double *gam = new double[gridSize]; // Temporary use in solving tri diag system
// -----
// Create indexes for accessing the grid with different dimensions as base.
// 'd' vectors which all have the same size as the entire grid net [1 x 'gridSize']
// Indexing is done with 'main dimension' as least significant and then the other
// dimensions in ascending order with increasing significance.
// Create M-matrices
long (*pos)[d] = new long[gridSize][d]; // (Works just because d is a "const")
long (*indexVector)[d] = new long[d][d]; // Temporary indexes
for(i=0; i<d; i++) for(j=0; j<d; j++) indexVector[i][j] = 0; // Initialize to 0
double (*M_u)[d] = new double[gridSize][d]; // M-matrix upper diagonals
double (*M_m)[d] = new double[gridSize][d]; // M-matrix middle diagonals
double (*M_l)[d] = new double[gridSize][d]; // M-matrix lower diagonals
double (*M_cross)[nCross] = new double[gridSize][nCross]; // Cross deriv. coeffs.
double x[d]; // Variable values used then creating matrices
double a,b,c; // Coefficients for 2nd, 1st and 0th order differentials
// 'i' holds the unsorted node number
for(i=0; i<gridSize; i++)
{
    // ----- Generate M-vectors -----
    // (We use a modified M matrix for non-cross derivatives M = 1-theta*M_orig)
    // 'a' - 2nd order differential coefficients.
    // 'b' - 1st order differential coefficients.
    // 'c' - 0th order differential (proportional) coefficients)
    for(j=0; j<d; j++)
    {
        // Translate d-dimensional 'indexVector' to a 1-dimensional node number
        long posAdd = 1; // First dimension is always next to each other
        pos[i][j] = 0;
        for(k=0; k<d; k++) // All dimensions must be considered
        {
            pos[i][j] += posAdd * indexVector[j][k];
            posAdd *= N_x[k];
        }
        // Translate d-dimensional 'indexVector' to variable values
        for(k=0; k<d; k++)
        {
            x[k] = min_x[k] + dx[k]*indexVector[j][k];
        }
        // PDE coefficients for dimension 'j' (different for each dimension)
        // ('a' is 2nd order, 'b' is 1st order and 'c' is 0th order differential)
        double riskPrice; // Market price of risk (Either log-volatility or correlation.)
        switch(j)
        {
            case 0: // spot [S] dimension + proportional contribution
                // ----- Stochastic correlation (LogOU) -----
                a = exp(2*x[1])*x[0]*x[0]/2.0; // 1/2*exp(2*Y)*S^2
                b = (r-q)*x[0]; // (r-q)*S
                c = -r; // r
                break;
            case 1: // volatility [vol] dimension
                // ----- Stochastic correlation (LogOU) -----
                riskPrice = x[2]*lambda1 + sqrt(1-x[2]*x[2])*lambda2; // vol-risk
                a = YVol*YVol/2.0;
                b = YRev*(YMean-x[1])-YVol*riskPrice;
                c = 0;
                break;
            case 2: // correlation [rho] dimension
                // ----- Stochastic correlation (LogOU) -----
                riskPrice = drw13corr*lambda1 + sqrt(1-drw13corr*drw13corr)*lambda3; // corr-risk
                a = corrVol*corrVol*(1-x[2]*x[2])*(1-x[2]*x[2])/2.0;
                b = corrRev*(corrMean-x[2])-corrVol*(1-x[2]*x[2])*riskPrice;
                c = 0;
        }
        if(indexVector[j][j] == 0) // Lower boundary node?
        {
            switch(j)
            {
                case 0:
                case 1:
                case 2:
                    // Assume zero-gamma (e.g no 2nd order derivative)
            }
        }
    }
}

```

```

        // Use one-sided finite difference for 1st order differential
        M_l[i][j] = -theta*0; // Not used (outside grid)
        M_m[i][j] = 1-theta*(-b/dx[j] + c)*dt;
        M_u[i][j] = -theta*( b/dx[j] )*dt;
        break;
    }
}
else if(indexVector[j][i] == N_x[i]-1) // Upper boundary node?
{
    switch(j)
    {
        case 0:
        case 1:
        case 2:
            // Assume zero-gamma (e.g. no 2nd order derivative)
            // Use one-sided finite difference for 1st order differential
            M_l[i][j] = -theta*(-b/dx[j] )*dt;
            M_m[i][j] = 1-theta*( b/dx[j] + c)*dt; // Not used (outside grid)
            M_u[i][j] = -theta*0;
            break;
    }
}
else // Interior node?
{
    M_l[i][j] = -theta*( a/(dx[j]*dx[j]) -b/(2*dx[j]) )*dt;
    M_m[i][j] = 1-theta*(-2*a/(dx[j]*dx[j]) + c)*dt;
    M_u[i][j] = -theta*( a/(dx[j]*dx[j]) +b/(2*dx[j]) )*dt;
}
// -- Generate cross-derivative vector and initial option values --
if(j==0) // Choose the natural indexing case (first of 'd' possible)
{
    // Cross-derivatives ('Natural' ordering of nodes)
    // -- Stochastic correlation (LogOU) --
    M_cross[i][0] = ( x[2]*YVol*exp(x[1])*x[0] )*dt; // dSdY
    M_cross[i][1] = ( drw13corr*corrVol*exp(x[1])*(1-x[2]*x[2])*x[0] )*dt; // dSdCorr
    M_cross[i][2] = ( x[2]*drw13corr*YVol*corrVol*(1-x[2]*x[2]) )*dt; // dYdCorr
    // Initialise option values at expiry ('Natural' ordering of nodes)
    // Vanilla call option
    // (Special case when K close to node to avoid "quantisation error".)
    if( x[0]-dx[0]/2.0 < K && K < x[0]+dx[0]/2.0 )
        u[i] = (2*(x[0]-K)+dx[0])*(2*(x[0]-K)+dx[0])/(8*dx[0]);
    else
        u[i] = __max(x[0]-K,0);
}
}
// -- Update Indices --
// 'j' holds the dimension to be considered "main dimension"
// 'i' already holds the main node number
for(j=0; j<d; j++)
{
    // Update all indices
    indexVector[j][i]++; // Main dimension 'j' is the first index to change
    if(indexVector[j][i] == N_x[j]) // Time to roll main dimension over?
    {
        indexVector[j][i]=0;
        for(k=0; k<d; k++)
        {
            if(k==j) continue; // The main dimension j is not to be touched here
            indexVector[j][k]++;
            if(indexVector[j][k] == N_x[k])
                indexVector[j][k]=0; // The dimension k must also roll over
            else
                break; // All indices are fine
        }
    }
}
}
// -- Actual ADI engine --
double *tmpDataPnt; // Pointer to temporary data
double bet; // Temporary variable used in tri-diagonal solver
double crossDenom[d]; // Denominator to use with cross derivatives
long crossIdx[d]; // Index ('Natural') to use when calculating cross derivatives
long crossDisp[d]; // Node displacement for moves in each dimension
long crossDispUp[d]; // Index displacement for up movement
long crossDispDown[d]; // Index displacement for down movement
crossDisp[0]=1; // Fill in values for 'crossIdx'
for(i=1; i<d; i++) crossDisp[i] = crossDisp[i-1]*N_x[i-1];
// Loop over all time steps
for(n=0; n<N_t; n++)
{
    // Each step has two sweeps in order to time centre cross terms
    for(int sweepRound=1; sweepRound<=2; sweepRound++)
    {
        // Each ADI-sweep has d different intermediate steps
        for(i=0; i<d; i++)
        {
            // The first intermediate ADI-step has a special rhs-vector
            if(i==0)
            {
                // First dimension
                for(l=0; l<grdsize; l++)
                    rhs[i] = (d/theta + (1-1/theta)*M_m[i][0])*u[l];
                for(l=1; l<grdsize; l++)
                    rhs[i] += (1-1/theta)*M_l[i][0]*u[l-1];
                for(l=0; l<grdsize-1; l++)
                    rhs[i] += (1-1/theta)*M_u[i][0]*u[l+1];
                // Other dimensions
                for(k=1; k<d; k++)
            }
        }
    }
}

```

```

        for(i=0; i<gridsize; i++)
            rhs[pos[i][k]] -= 1/theta*M_m[i][k]*u[pos[i][k]];
        for(i=1; i<gridsize; i++)
            rhs[pos[i][k]] -= 1/theta*M_l[i][k]*u[pos[i-1][k]];
        for(i=0; i<gridsize-1; i++)
            rhs[pos[i][k]] -= 1/theta*M_u[i][k]*u[pos[i+1][k]];
    }
    // Cross derivatives
    if(sweepRound == 1) // First sweep (explicit crosses etc)
    {
        tmpDataPnt = u; // Simply use 'u' to calculate crosses
    }
    else // Second sweep (time-centred crosses etc)
    {
        for(i=0; i<gridsize; i++)
            lhs[i] = (lhs[i]+u[i])/2.0; // 'rhs' = average of 'rhs' and 'u'
        tmpDataPnt = lhs; // Use average of 'u' and 'lhs' for crosses
    }
    // Calculate CROSS DERIVATIVES' contribution
    for(i=0; i<d; i++) crossIdx[i]=0; // Clear indices
    for(i=0; i<gridsize; i++)
    {
        // Generate index displacements and denominators (for each dim.)
        for(k=0; k<d; k++)
        {
            if(crossIdx[k]==0) // Lower boundary? (dim. 1)
            {
                crossDispDown[k] = 0;
                crossDenom[k] = 0;
            }
            else
            {
                crossDispDown[k] = -crossDisp[k];
                crossDenom[k] = dx[k];
            }
            if(crossIdx[k]==N_x[k]-1) // Upper boundary? (dim. 2)
                crossDispUp[k] = 0;
            else
            {
                crossDispUp[k] = crossDisp[k];
                crossDenom[k] += dx[k];
            }
        }
        // Calculate cross derivatives (using indices generated above)
        m = 0;
        for(k=0; k<d-1; k++)
        {
            for(l=k+1; l<d; l++)
            {
                rhs[i] += M_cross[i][m] *
                ((tmpDataPnt[i+crossDispUp[k]+crossDispUp[l]]-tmpDataPnt[i+crossDispDown[k]+
                crossDispUp[l]])-(tmpDataPnt[i+crossDispUp[k]+crossDispDown[l]]-
                tmpDataPnt[i+crossDispDown[k]+crossDispDown[l]])/(crossDenom[k]*crossDenom[l]));
                m++;
            }
        }
        for(k=0; k<d; k++) // Increase indices
        {
            crossIdx[k]++;
            if(crossIdx[k]==N_x[k])
                crossIdx[k]=0;
            else
                break;
        }
    }
}
// All steps but first intermediate step are similar but
// they are dependant on dimension 'intermediate'.
else
{
    for(i=0; i<gridsize; i++)
        rhs[pos[i][j]] = lhs[pos[i][j]] + (M_m[i][j]-1)*u[pos[i][j]];
    for(i=1; i<gridsize; i++)
        rhs[pos[i][j]] += M_l[i][j] * u[pos[i-1][j]];
    for(i=0; i<gridsize-1; i++)
        rhs[pos[i][j]] += M_u[i][j] * u[pos[i+1][j]];
}
// Solve for implicit step (tridiagonal solver)
// (Using a tridiagonal solver based on section 2.4 in Numerical Recipes
// but modified to take dynamic indexing of nodes.)
if(M_m[0][j]==0) throw("Error #1 in tridiagonal solver.");
bet = M_m[0][j];
lhspos[0][j] = rhs[pos[0][j]]/bet;
for(i=1; i<gridsize; i++)
{
    gam[i] = M_u[i][j]/bet;
    bet = M_m[i][j]-M_l[i][j]*gam[i];
    if(bet==0) throw("Error #2 in tridiagonal solver.");
    lhs[pos[i][j]] = (rhs[pos[i][j]]-M_l[i][j]*lhs[pos[i-1][j]])/bet;
}
for(i=gridsize-2; i>1; i--)
    lhs[pos[i][j]] = gam[i+1]*lhs[pos[i+1][j]];
}
// Swap data pointers 'u' and 'rhs'. ('lhs' becomes the new 'u' in next step.)
tmpDataPnt = u;
u = lhs;
lhs = tmpDataPnt;
// Display loop number

```

```

    cout << N_t-n << "\n" << flush;
}
// ~ Data output ~~~~~
// Write out-data to file
if(startNode >= 0 && (startNode+(numNodes-1)*stepNode) < gridSize && stepNode > 0)
{
    outFile.open(argv[2], fstream::out);
    outFile << setprecision(16);
    for(i=0; i<numNodes; i++)
    {
        outFile << u[startNode+i*stepNode] << endl;
    }
    outFile.close();
}
else
{
    cout << "Syntax Error: Output node numbers (No out-file written)" << endl;
}
// Delete all allocated memory
delete dx;
delete [] u;
delete [] lhs;
delete [] rhs;
delete [] gam;
delete [] pos;
delete [] M_l;
delete [] M_m;
delete [] M_u;
delete [] M_cross;
delete [] indexVector;
return 0;
}

```

Bibliography

- [1] N. H. Bingham and Rüdiger Kiesel. *Risk-neutral valuation: pricing and hedging of financial derivatives*. Springer, London, 2nd ed edition, 2004.
- [2] Tomas Björk. *Arbitrage theory in continuous time*. Oxford Univ. Press, Oxford, 1998.
- [3] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.
- [4] Peter Carr and Liuren Wu. Stochastic skew in currency options. 2003.
- [5] Peter Carr and Liuren Wu. Stochastic skew models for fx options. 2004. Notes and lecture slides.
- [6] Les Clewlow and Chris Strickland. *Implementing derivatives models*. Wiley series in financial engineering. Wiley, Chichester; New York, 1998.
- [7] Rama Cont and José da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2:45–60, 2002.
- [8] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985.
- [9] Ian J.D. Craig and Alfred D. Sneyd. An alternating-direction implicit scheme for parabolic equations with mixed derivatives. *Computers and Mathematics with Applications*, 16(4):341–350, 1988.
- [10] Emanuel Derman and Iraj Kani. Riding on a smile. *Risk*, 7:32–39, 1994. February.
- [11] Daniel J. Duffy. *Financial instrument pricing using C++*. John Wiley, Hoboken, NJ, 2004.
- [12] Bruno Dupire. Pricing with a smile. *Risk*, 7:18–21, 1994. January.

BIBLIOGRAPHY

- [13] Jean-Pierre Fouque, K. Ronnie Sircar, and G. C. Papanicolaou. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, Cambridge, 2000.
- [14] Jürgen Hakala and Uwe Wystup. *Foreign Exchange Risk*. Risk Books, 2002.
- [15] Espen Gaarder Haug. *The complete guide to option pricing formulas*. McGraw-Hill, New York; London, 1997.
- [16] Michael T. Heath. *Scientific computing: an introductory survey*. McGraw-Hill computer science series,. McGraw-Hill, New York, 1997.
- [17] Steven L. Heston. Closed- form solution for options with stochastic volatility with applications to bond and currency options. *Review Of Financial Studies*, 6(2):327–43, 1993.
- [18] John C. Hull. *Options, futures, and other derivatives*. Prentice Hall International, Upper Saddle River, N.J., 4. edition, 2000.
- [19] John C. Hull and Alan White. Pricing of options on assets with stochastic volatilities. *Journal Of Finance*, 42:281–300, 1987. June.
- [20] Aapo Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.
- [21] Aapo Hyvärinen and Erkki Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000.
- [22] Arieh Iserles. *A first course in the numerical analysis of differential equations*. Cambridge texts in applied mathematics, [15]. Cambridge Univ. Press, Cambridge, 1996.
- [23] Peter Jäckel. *Monte Carlo methods in finance*. Wiley finance series. Wiley, New York; Chichester, 2002.
- [24] Peter Jäckel. Stochastic volatility models - past, present and the future. In *Quantitative Finance Review conference*, London, 2003.
- [25] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. Applications of mathematics, 23. Springer-Vlg, Berlin; New York, 1992.
- [26] Hayne E. Leland. Option pricing and replication with transaction costs. *The Journal of Finance*, 40:1283–1301, 1985.

BIBLIOGRAPHY

- [27] Alan L. Lewis. *Option valuation under stochastic volatility: with Mathematica code*. Finance Press, Newport Beach, CA, 2000.
- [28] Alexander Lipton. *Mathematical methods for foreign exchange: a financial engineer's approach*. World Scientific, Singapore; River Edge, N.J., 2001.
- [29] Alexander Lipton and William McGhee. Universal barriers. *Risk*, 15(2):81–85, 2002.
- [30] Cornelius Luca. *Trading in the global currency markets*. New York Institute of Finance, New York, 2nd ed edition, 2000.
- [31] R. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- [32] Bernt Øksendal. *Stochastic differential equations: an introduction with applications*. Universitext., Springer, Berlin, 5. edition, 1998. Bernt Øksendal.
- [33] Marcus Overhaus, Andrew Ferraris, Thomas Knudsen, Ross Milward, Laurent Nguyen-Ngoc, and Gero Schindlmayr. *Equity derivatives: theory and applications*. Wiley Finance, New York, 2002.
- [34] D. M. Pooley, K. R. Vetzal, and P. A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40, 2003.
- [35] William H. Press. *Numerical recipes in C: the art of scientific computing*. Cambridge Univ. Press, Cambridge, 1992.
- [36] Riccardo Rebonato. *Volatility and correlation in the pricing of equity, FX and interest-rate options*. Wiley series in financial engineering., Wiley, Chichester, 1999.
- [37] Riccardo Rebonato. *Volatility and correlation: the perfect hedger and the fox*. J. Wiley, Chichester, West Sussex, England; Hoboken, NJ, 2nd edition, 2004.
- [38] L. C. G. Rogers and David Williams. *Diffusions, Markov processes and Martingales*. Cambridge University Press, Cambridge, 2 edition, 2000. Vol. 1, Diffusions, Markov processes and martingales.
- [39] Mark Rubinstein. Implied binomial trees. *Journal of Finance*, 49:771–818, 1994. July.
- [40] Wim Schoutens. *Lévy processes in finance: pricing financial derivatives*. Wiley series in probability and statistics. Wiley, Chichester, 2003.

BIBLIOGRAPHY

- [41] Louis O. Scott. Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial and Quantitative Analysis*, 22:419–439, 1987.
- [42] Elias M. Stein and Jeremy C. Stein. Stock prices distribution with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4(4):727–752, 1991.
- [43] Nassim Taleb. *Dynamic hedging: managing vanilla and exotic options*. Wiley series in financial engineering. Wiley, New York; Chichester, 1997.
- [44] Domingo Tavella and Curt Randall. *Pricing financial instruments: the finite difference method*. Wiley series in financial engineering. John Wiley, New York; Chichester, 2000.
- [45] James B. Wiggins. Option values under stochastic volatility; theory and empirical estimates. *Journal Of Financial Economics*, 19:351–72, 1987. December.
- [46] Uwe Wystup. The market price of one-touch options in foreign exchange markets. *Derivatives Week*, XII(13), 2003.
- [47] R. Zvan, P.A. Forsyth, and K.R. Vetzal. Robust numerical methods for pde models of asian options. *Journal of Computational Finance*, 1(2), 1998.