Umeå University
Department of Mathematics and Mathematical Statistics

B.Sc. Mathematics thesis

Swaption pricing and isolating volatility exposure

Author:
Tomas Forsberg, M.Sc. Engineering Physics

Supervisor:
Lars-Daniel Öhman, Ph.D.
Department of Mathematics and Mathematical Statistics,
Umeå University

Examinator:
Thomas Önskog, Ph.D.
Department of Mathematics and Mathematical Statistics,
Umeå University

May 2011

Umeå, Sweden

0.1 Abstract

Starting from basic financial mathematics, we cover the mathematics of pricing *swaptions*, options on interest rate swaps. We then continue to the topic of obtaining an approximately pure volatility exposure. This exposure to volatility, which in practice enables us to *trade volatility* according to our perceptions of the market, is obtained by buying or selling swaptions and appropriate amounts of the underlying interest rate swap contract. Taking offsetting positions in the underlying contract is called *hedging* and is covered in depth. We note that hedging can primarily be done in two ways, and discuss the advantages and disadvantages of each of them. After deriving the value formulas for such a swaption strategy aimed at isolating volatility exposure we end with a discussion on the transition from theory to practice.

We find that this way of trading volatility is conceptually simple, but that pre-trade profitability analysis is difficult due to the sometimes poor availability of the sophisticated data needed to simulate such a swaption strategy. Despite the possible limitations in the data necessary to translate this theory into an experimental setup, this thesis serves as a good basis for further research on the profitability of a volatility trading strategy using interest rate swaptions.

0.2. Sammanfattning

Med utgångspunkt i grundläggande finansiell matematik går vi igenom matematiken bakom prissättningen av *swaptions*, optioner på ränteswappar. Vi fortsätter sedan till hur man uppnår en approximativt ren exponering mot volatilitet. Denna exponering mot volatilitet, som i praktiken gör det möjligt för oss att *handla volatilitet* enligt vår syn på marknaden, uppnås genom att köpa eller sälja swaptions och lämpliga mängder av det underliggande ränteswap-kontraktet. Att ta kvittande positioner i det underliggande kontraktet kallas att göra en *hedge*, vilket behandlas i detalj. Vi noterar att hedgning framför allt kan göras på två sätt, och diskuterar fördelarna och nackdelarna med vart och ett av dem. Efter att ha härlett formlerna för en sådan swaption-strategi med syfte att isolera volatilitetsexponering avslutar vi med en diskussion om övergången från teori till praktik.

Vi finner att detta sätta att handla volatilitet är konceptuellt enkelt, men att förhandsanalys av den förväntade förtjänsten är svår på grund av den ibland dåliga tillgängligheten på det sofistikerade data som krävs för att simulera en sådan swaptionstrategi. Trots de eventuella begränsningarna i underlaget som behövs för att översätta denna teori till en experimentell uppställning, så tjänar denna uppsats som en god grund för fortsatt forskning kring avkastningsmöjligheterna hos en volatilitetstrading-strategi som använder sig av swaptions.

Contents

Chapter 1

Introduction

1.1 Background

The pricing of derivatives is one of the more advanced practices in mathematical finance. Starting with vanilla options and proceeding to exotic and American contracts in several asset classes and markets, the models and methods vary enormously. Much of the research in this area is done in-house by the financial industry where ground breaking results are company secrets, or at universities where there is a high threshold of knowledge to fully grasp the advanced mathematical content in the articles published. This makes it difficult for students at, say, graduate level to educate themselves in an efficient way. There is no shortage of texts on basic option topics such as basic European options pricing, explanations on what binomial trees and Monte Carlo methods are and how they relate to the practice of option pricing. Neither is it hard to find articles regarding advanced computational methods or mathematical concepts pushing the frontiers of financial engineering research forward. However, in between these two levels there is a shortage of easily accessible material, where one is guided from the basic concepts of finance to advanced applications in a coherent and pedagogical way. This thesis contributes with a guide to swaption pricing and volatility trading, accessible for students of mathematical finance at approximately graduate level.

Apart from the level of mathematical knowledge mentioned above, we assume that the reader is familiar with basic financial terminology. It is particularly suitable for the reader to have some prior experience with interest rate theory, as this is where we start off from in Chapter 2.

1.2 Outline

In Chapter 2 the reader is taken through a derivation of how to price swaptions, with separate sections explaining in detail each of the ingredients necessary to arrive at the swaption pricing formula.

In Chapter 3 we go deeper into the concept of volatility and hedging and explain some of the techniques necessary to trade volatility using swaptions.

Finally, Chapter 4 extends Chapter 3 in that it explicitly treats some practical issues of trading volatility using the techniques explained earlier. In the end of Chapter 4 we conclude with a discussion on the application of this type of strategy and the transition from theory to an experimental setting.

Symbols and notations used throughout the chapters can be found in Appendix A. Much of the material in this thesis, particularly Chapter 2, is general theory and mainly not the work of a single contributor. Chapter 3 on delta-hedging and swaptions, however, is largely inspired by literature by Emanuel Derman and others on obtaining volatility exposure. All references used throughout the text can be found in the Bibliography section.

When using monetary entities we will consistently use U.S. dollars, denoted *dollars* or just \$. Of course the logic and methods are the same regardless of which currency we use.

Chapter 2

Mathematics of pricing swaptions

To understand the logic behind the pricing of a swaption contract one has to understand the properties and mathematics of the different entities affecting the swaption value. This chapter takes you through this theory, explaining interest rates, bonds, swaps and options, arriving at the formula by which the swaption price is calculated.[1]

2.1 Time value of money

The concept of interest rate is based on the fundamental assumption that a dollar is worth more today than it is in the future.[2] When we talk about these things it is useful to define the *discount factor*, defined as

$$B(t,T) = \text{value at time } t \text{ of a dollar received at time } T. \qquad (2.1)$$

The $B(t,T)$ notation will be used interchangeably for both discount factors and bonds. The reason for this is that bond prices by definition *is* the discount factor, the price you pay today for future money (normalized so that the price refers to receiving *one dollar* at maturity), explained further in Section 2.4. Here, the notation $B(t,T)$ will always refer to a contract that pays one dollar at maturity. Thus,

$$t < T \longrightarrow B(t,T) < 1 \qquad (2.2)$$

$$t = T \longrightarrow B(t,T) = 1. \qquad (2.3)$$

[1]As with most advanced option contracts there are several ways of valuing and pricing swaptions, depending on which underlying model we assume. In this thesis we use the Black model, an extension of the Black-Scholes model, described further in Section 2.8.

[2]Excluding a deflationary situation.

The *yield* is defined as the unique constant interest rate r_y that has the same effect as $B(t, T)$ under continuous compounding,

$$e^{-r_y \cdot (T-t)} = B(t, T) \tag{2.4}$$

Let us also denote the rate between time t_1 and t_2 (the rate at which we can borrow money between time t_1 and t_2) by $r(t_1, t_2)$. By convention we express rates as annualized rates. So if, for example, $t_1 = 0$ and $t_2 = 2$ (years) a dollar invested at t_1 would have grown to $(1 + r(t_1, t_2))^2$ dollars at t_2. We also define the *present value* $PV(t_1, t_2)$ as the value at time t_1 of money received at t_2. Sometimes we will refer to the present value at t of money received at several future points in time up until contract maturity T. For simplicity this will be denoted $PV(t, T)$, using only present time and maturity.

2.2 The yield curve

The yield curve is a curve that plots interest rates (yields of bonds) of different maturities, $r(t, t_1)$, $r(t, t_2)$, $\ldots$, on the y-axis and maturities t_1, t_2, $\ldots$ on the x-axis. The maturity is the only difference between the bonds, so we can use the yield curve to draw conclusions about future interest rates as perceived by the bond market today. Normally we expect longer-term rates to be higher than shorter term rates since we expect a larger premium for lending money over a longer period of time, resulting in a positive slope at all points of the yield curve.

2.3 Forward rates

From the yield curve we can deduce current market rates depending on for how long we want to lend or borrow money.[3] However, at time t ($t < t_1 < t_2$) we only know the rates $r(t, t_1)$, $r(t, t_2)$, and so on. We do *not* know the rate $r(t_1, t_2)$ which we will be able to borrow money at for some future time period (t_1,t_2). We can make a "best guess" though, and this best guess is called the *forward rate* $fr(t_1, t_2)$. This rate is derived by noting that investing in two risk-free strategies over the same period of time should yield the same returns. So, investing one dollar in a one-year zero-coupon bond, $B(t, t_1)$, only to instantly reinvest the money received at time t_1 in a new

[3]In fact, there is not a single universal interest rate for a specific maturity or period, but countless different yields as provided by bonds issued by governments and corporations. The market naturally demands more yield from a bond issuer that is percieved as "risky" compared to someone with a high credit rating. When talking about the yield curve we are most often only referring to bonds that are practically risk-free, such as those issued by the U.S. government, or a common reference rate such as the *London Interbank Offered Rate* (LIBOR).

one-year zero-coupon bond, $B(t_1, t_2)$ at rate $fr(t_1, t_2)$, should be equivalent to investing one dollar in a two-year zero-coupon bond $B(t, t_2)$ today and holding it for two years. Thus, the value at t_2 of these two alternatives should be equal, as expressed in the following equation;

$$(1 + r(t, t_1)) \cdot (1 + fr(t_1, t_2)) = (1 + r(t, t_2))^2 \tag{2.5}$$

which leads to

$$fr(t_1, t_2) = \frac{(1 + r(t, t_2))^2}{(1 + r(t, t_1))} - 1. \tag{2.6}$$

Analogously, we can deduce all forward rates $fr(t_i, t_j)$ for intermediate periods ($t < t_i < t_j$) as long as we know $r(t, t_i)$ and $r(t, t_j)$.

2.4 Bonds

A bond is a debt security, like a loan, where the issuer borrows money from the bond holder. As we will see there are more complicated cases than just receiving a certain amount of money at maturity, which means we need a broader definition of *yield* than was provided in Section 2.1. We broaden the definition by saying that the *yield to maturity* of a bond is the fixed interest rate implied by *the payment structure in relation to the price.* The payment structure consists of intermediate payments (called *coupons*) as well as the money received at maturity.

It is worth pointing out that the *price* of a bond is the *market's opinion* of the present value of the payment structure, not a universal truth. This is why analysts can look at bond prices to draw conclusions on what "the market" thinks, and make judgements on *undervalued* or *overvalued* bonds.

2.4.1. Zero-coupon bond

The simplest of bonds is the zero-coupon bond, where there are only two cash flows; the buyer pays the issuer at the beginning of the period, and receives the face value (also called the *principal*) at maturity. There are no intermediate payments (coupons), hence the name zero-coupon. In simple terms, we are lending a certain amount of money (N) to someone in order to receive back a larger amount at a future time. The price $P(t, T)$ at time t of a zero-coupon bond with face value N and maturity T is

$$P(t, T) = B(t, T) \cdot N, \tag{2.7}$$

that is, the present value of the N dollars to be paid at maturity. Using this equation we can calculate $B(t, T)$ if we know the price and face value of an option.

This is very similar to deducing market rates from the yield curve and, indeed, there is an intimate relationship between forward rates, discount factors and bond prices. The derivation of the forward rate (2.6) can be done using bond or discount factor notation. We argued that a dollar invested without risk today should always be worth the same in two years regardless of the strategy. Equivalently, the present value of a dollar received two years from now, that is at t_2, should be the same regardless of the strategy. Say we have two alternative strategies to choose from;

1. invest in a two-year bond paying 1 dollar at t_2, or
2. invest in a one-year bond today to reinvest the principal at t_1 in another one-year bond paying 1 dollar at t_2.

What is the cost (present value) of these two strategies? The cost today of strategy 1 is simply the cost of the two-year bond, $B(t, t_2)$. In strategy 2 we see that the value of the final t_2-dollar at t_1 is the discounted value of a dollar, $B(t_1, t_2)$. Thus, the present value (cost today) of this strategy is the discounted value of $B(t_1, t_2)$ at t_1, i.e. $B(t, t_1) \cdot B(t_1, t_2)$. The present values should be equal, so

$$B(t, t_1) \cdot B(t_1, t_2) = B(t, t_2). \tag{2.8}$$

In the same manner as before, we can now deduce a "best guess" for the forward discount factor $B(t_1, t_2)$ by observing $B(t, t_1)$ and $B(t, t_2)$ in the market and using the above expression to find

$$B(t_1, t_2) = \frac{B(t, t_2)}{B(t, t_1)}. \tag{2.9}$$

These results are valid for all $t < t_1 < t_2$, provided that we have market prices for bonds of suitable maturities.

2.4.2. Coupons

Some bonds pay money at predefined points in time until maturity. These payments, as mentioned briefly earlier in this section, are called coupons. If the bond contract includes coupon payments, the price of the bond is the present value of the payment structure, i.e. the present value of the principal paid at maturity plus the present values of the individual coupon payments. The present value of the principal payment is the same as the price of a zero-coupon equivalent, and the present value of each coupon can be found using discount factors. We can express this as

$$PV(t, T) = B(t, T) \cdot N + \sum_{i=1}^{n} B(t, t_i) \cdot c_i \tag{2.10}$$

where c_i is the i:th coupon payment, which occurs at time t_i. The n:th (last) payment thus occurs at $t_n = T$. The size of the coupon payments can be fixed or floating (based on a reference interest rate). When we say that the coupon size depends on a certain rate, we mean that the coupon c_i is calculated by multiplying the face value N by a certain percentage which is dependent on the rate r. We will use the subscript f to refer to the *floating rate* and the subscript s to refer to the *fixed rate.*[4] Then, for a fixed rate coupon bond we have the present value

$$PV_s(t,T) = B(t,T)\cdot N + \sum_i^n B(t,t_i)\cdot \frac{c}{m} \tag{2.11}$$

$$= B(t,T)\cdot N + \sum_i^n B(t,t_i)\cdot \frac{r_s\cdot N}{m} \tag{2.12}$$

$$= B(t,T)\cdot N + \frac{r_s\cdot N}{m}\sum_i^n B(t,t_i) \tag{2.13}$$

where m denotes the number of periods (coupon payments) per year and n is m times the number of years until maturity, i.e. the total number of periods. The reason for dividing by m is that we are dealing with annualized rates and need to account for the periods being shorter than a year (such as semi-annually if $m = 2$).

For a floating rate bond things look a bit more complicated at first glance. Floating coupon rates are usually based on beginning-of-the-period rates and payed at the end of the period. So the first payment, for example, will be

$$\text{coupon payment at } t_1 = \frac{r_f(1)\cdot N}{m} = \frac{r(t_0,t_1)\cdot N}{m}. \tag{2.14}$$

Intuitively, we would write the present value analogous to the present value for the fixed rate,

$$PV_f(t,T) = B(t,T)\cdot N + \sum_i B(t,t_i)\cdot \frac{r_f(i)\cdot N}{m}, \tag{2.15}$$

but let us investigate the floating rate case further. At maturity $T = t_n$, additional to the last coupon payment we receive the payment of the face value. The value of this face value payment, at maturity when it is about to take place, is obviously N. Since we are at maturity there is no discounting. To deduce the total value of the payments at maturity we need to take the coupon payment into account. At t_{n-1}, the starting point of the n:th and last period, we set the floating rate to $r_f(n) = r(t_{n-1}, t_n)$ as determined by current market rates. Since this is the rate which the last coupon payment will be calculated from, at the end of this last period (which is $1/m$ years) we receive a payment of $r_f(n)\cdot N/m$, giving us a total value at time t_n of

[4] We use "s" as in "strike" since the fixed rate is the *strike price* of a swaption, covered further in Section 2.9.

$N + r_f(n) \cdot N/m$. The value of the floating rate bond at time t_{n-1} is the discounted value of these two payments,

$$PV_f(t_{n-1}, T) = \frac{N + r_f(n) \cdot N/m}{1 + r(t_{n-1}, t_n)/m} = N\frac{1 + r_f(n)/m}{1 + r_f(n)/m} = N. \qquad (2.16)$$

This procedure can be repeated all the way back to present time, proving that

$$PV_f(t, T) = N. \qquad (2.17)$$

2.5 Financial derivatives

Before moving on to the next section on *hedging*, it is necessary to give a quick explanation of the concept of *derivatives*.

In financial mathematics the word *derivative* could mean two things: a first- or higher-order derivative in the common mathematical sense, or a *derivative contract*. The latter refers to a contract which value is dependent on another asset or entity. It is called a *derivative* since its' value is *derived* from something else. Two common types of derivatives are *futures* and *options*.

Futures is an agreement on a transaction at a future point in time. This is very similar to a purchase today since it is a binding contract and all details are agreed upon (price, quantity, asset), but delivery and payment take place some time in the future. A common time span of a futures contract is three months, but there are other time spans as well. There are futures contracts based on both commodities, stocks and other assets, and they are traded on markets similar to stock markets.

Options are similar to futures in that payment and delivery of the underlying asset takes place at some predefined future point in time, but the difference is that the holder of an options contract is *not obligated* to go through with the transaction. The issues of such a contract is extending an offer to the holder, giving him or her the option to exercise the offer or not. For this option the holder, or buyer, of the options contract pays a *premium* to the seller. Options are covered further in Section 2.8.

2.6 Basic concept of hedging

To proceed smoothly to the topic of interest rate swaps (which is covered in the next section), we will now introduce and explain the concept of *hedging*.

Hedging in finance is the act of reducing some sort of risk or exposure. A classical example is a manufacturer with aluminium as one of the company's production factors. If the company receives a large order to manufacture components over the next year, it is exposed to changes in aluminium prices over this period (since it will need aluminium to produce its components over the next year). The company can now hedge away some or all of its exposure to swings in aluminium prices by, for example, buying aluminium futures with maturity dates distributed over the next year. The key in hedging is to take an offsetting position, so that what is lost in one position (increased production costs due to increased aluminium prices) is offset by earnings in the other position (gains in a long[5] aluminium futures position due to the increase in aluminium spot price). Since we are most often charged with some sort of transaction costs there is usually a small cost associated with taking the hedge position. This makes it very comparable to buying an insurance; we give up a small amount of money to guarantee avoidance of larger losses.

In the financial markets we can find hedge (offsetting) positions in many contexts. A very common and simple hedge is to take offsetting positions in a derivative and its underlying instrument, such as a stock (underlying instrument) and futures or options (derivatives) on this stock. Hedging with futures or options is beneficial in terms of costs since you are not buying or selling huge quantities in the underlying instrument causing large commission fees, you are only buying or selling *contracts* for possible deliverance at a future date. These contracts are much cheaper than the underlying instrument and so is the trading cost.

So far we have only talked about buying or selling something to hedge an existing position or situation, but there is another very applicable way of hedging; *swaps*. A swap contract is an agreement between two parties to exchange something in a structued manner. The contract can be as simple or as exotic as the parties wish, as long as they are in agreement. For example, let us assume that the manufacturing company mentioned above buys its' aluminium in a currency different from the currency on the market on which they sell their components. Assuming they buy aluminium in dollars and sell components in euro, they are exposed to changes in the USD/EUR relationship. If the dollar is strengthened in relation to the euro, they wont be able to buy as much aluminium for their sales income as they used to. This is where swaps come in handy. If the company receive an order where they will sell components (in euro) a year from now, they can enter into a currency swap which offsets possible gains or losses resulting from changes in the USD/EUR rate.

Another situation is where you are subjected to interest rate risk, by paying interest based on some reference rate for example. In this case you might want to hedge away the interest rate risk to find yourself in the situation of paying a fixed, known, interest rate instead. This is easily arranged

[5]A *long* position is usually an ownership of an asset, where we will benefit from an increase in the asset price. A *short* position is the opposite, commonly achieved by an action called *short selling* (not covered here). Another type of long/short position is the case of a position in a contract which benefits from a price increase/decline in the underlying asset.

by entering into an interest rate swap contract, which is the subject of the following section.

Due to the varying nature of swap contracts they are considered OTC (over the counter) instruments. This means that most swaps are not traded on exchanges in the same way as stocks or futures, but rather tailored individually by a broker on request by its clients.

2.7 Interest rate swaps

As noted in the previous section, a *swap* is an exchange of one payment stream for another. Thus, an *interest rate swap* is an exchange of interest rate payments. In this chapter we will look closer at the mathematics of an interest rate swap.

The simplest type of interest rate swaps is the *plain vanilla interest rate swap* where a floating rate is exchanged for a fixed one. The floating rate is pegged to a reference rate such as LIBOR.[6] The fixed rate is (usually) the *par rate*, which, based on forward rates extracted from the yield curve, makes the present value of the swap evaluate to zero. By convention, we call the counterparty paying the fixed rate *payer* and the counterparty paying the floating rate *receiver*. At each payment date there is a single cashflow from one party to the other, of the net value of both payments. The fixed rate is constant and the floating rate to be used is observed at the start of each period. Note that this is very similar to the floating coupon payments explained in Section 2.4. The i:th net payment for a payer swap, executed at the *end* of the period, is then

$$\frac{N}{m} \cdot (r_f(i) - r_s) \tag{2.18}$$

with a present value at t_0 of

$$B(t_0, t_i) \cdot \frac{N}{m} \cdot (r_f(i) - r_s). \tag{2.19}$$

Note the distinction $r_f(i) = r(t_{i-1}, t_i)$, as seen at t_{i-1} but relating to *period* i. The *coupon frequency* $\omega = 1/m$ is commonly used but we will continue to use m. The present value at t_0 of such a payment schedule starting at t_0 and ending at $T = t_n$ is

$$\sum_{i=1}^{n} B(t_0, t_i) \cdot \frac{N}{m} \cdot (r_f(i) - r_s) \tag{2.20}$$

$$= \frac{N}{m} \cdot \sum_{i=1}^{n} B(t_0, t_i) \cdot (r_f(i) - r_s) \tag{2.21}$$

[6]LIBOR is an abbrevation of the London Interbank Offered Rate.

where $n = mT$. If we don't have constant period lengths we need to multiply every payment with $t_i - t_{i-1}$ instead of dividing by m and keep this term within the sum expression. If the present value of the swap differs from zero by some amount at the inception of the swap, a payment of that amount is made from one party to the other at t_0 to adjust for this. If r_s has not yet been determined we adjust the fixed rate to make the present value equal to zero. This fixed rate is called the *par* rate.

Note that an interest rate swap position is equivalent to a portfolio consisting of a pair of bonds, one short/long fixed rate coupon bond and one long/short floating rate coupon bond, with the same face value. At maturity the face value payments cancel out, leaving the difference in coupon payments as the only cash flows (just like an interest rate swap). Hence, the present value $PV_{\mathrm{P}}(t_0, T)$ of a payer swap is

$$PV_{\mathrm{P}}(t_0, T) = PV_f(t_0, T) - PV_s(t_0, T) \tag{2.22}$$

Using (2.17) and (2.13), we can write the present value as

$$PV_{\mathrm{P}}(t_0, T) = N - \left(B(t, T) \cdot N + \frac{r_s \cdot N}{m} \cdot \sum_i B(0, t_i) \right). \tag{2.23}$$

The par rate is determined, as mentioned above, by requiring that the present value of the swap at inception should be zero. By setting $PV_{\mathrm{P}}(t_0, T) = 0$, (2.23) simplifies to

$$r_s = m \cdot \frac{1 - B(t_0, T)}{\sum_i B(0, t_i)}. \tag{2.24}$$

2.8 Options

An option is the right, but not the obligation, to buy (call option) or sell (put option) a specific asset, on specified terms. The terms vary widely, but the most important distinction is between *European* and *American* options. In the European case you are only allowed to exercise the option at maturity, while an American option can be exercised at any time including maturity. Another common type is the *Bermudan option* where you are allowed to exercise the option at a finite number of predefined intermediate points in time. For this option to buy or sell the underlying asset you pay a certain *premium*, which is the option price.

Before getting into the pricing of swaptions, which is a rather advanced options contract, we will look at options theory in general. Fundamental understanding is best obtained by starting with simple examples.

The value of a European call (put) option at maturity is the difference between how much your option contract lets you buy (sell) it for, and how much you can buy (sell) it for in the spot market. Also, since we can always choose not to exercise our option, the value can never be less that zero. Denoting the call option value as V_c and the put option value as V_p, this can be expressed as

$$V_c = \max[0, S - K] \tag{2.25}$$

$$V_p = \max[0, K - S] \tag{2.26}$$

where K denotes the *strike price*, the price we are allowed to buy/sell the asset for at maturity, and S denotes the market price of the underlying asset. Since an option value can never be negative, we cannot lose more money than the paid premium. We can always choose simply not to exercise the option. This is a good thing, but on the other hand there is always the possibility that the option will expire with zero value, which causes us to lose *all* of our invested money. The fact that we only invest the premium which is in the magnitude of single dollars, but still value the option based on $|S - K|$, opens up possibilities of serious returns. As an example, consider an option with strike \$110 bought when S=\$100 at a price of \$1. Let's say the underlying asset moves up to \$120 at maturity, representing a 20% return on an investment in the underlying asset. The option holder, however, has gotten a return of \$10 on his \$1 which means a return of 900%.

2.8.1. Time value of options

Another fundamental quality of options is *time value* which is based on the fact that "anything can happen". Even if an option is deep out-of-the-money there is always a non-zero probability of the option ending up in-the-money at expiration.[7] As we get closer and closer to expiration, the time value goes to zero. Time value means that the option value at time t, V_t, will always be greater than zero for $t < T$, even though deep out-of-the-money options will have V_t close to zero.

2.8.2. Options pricing

Options are priced either by analytic solutions or by numerical methods and/or simulation. Since analytic solutions are rare, due to the stochastic nature of price processes and the complex nature of applied stochastic calculus, most options are priced with numerical methods (e.g. binomial/trinomial models, finite difference or finite elements methods) or sim-

[7]In-the-money, at-the-money and out-of-the-money tells us whether the current price of the underlying asset (the spot price) is at such a level that the option, if exercised today, has non-zero value. A \$100 call option would be in-the-money for all spot prices larger than \$100, since we would make a profit if we were to exercise the option right now. If the current price of the underlying is exactly \$100 our \$100 call option would be at-the-money, and for spot prices below \$100 it would be out-of-the-money since there is no profit to be made from an immediate exercise.

ulation (e.g. Monte Carlo simulation). The price should, however, always reflect the *discounted expected value at maturity.*

For some simple option contracts it is possible to find analytical solutions. In the case of European call and put options for example, we can find an analytical solution by assuming that the evolution of an asset price S is a stochastic process, most commonly assumed to follow a geometric Brownian motion (a log-normal process)

$$dS = \mu S dt + \sigma S dz \tag{2.27}$$

where z is a Wiener process.[8] The parameters μ and σ are called the *drift* and the *volatility*, respectively. To see why this is a log-normal process (log-normal means that the *logarithm* is normally distributed), divide through by S

$$\frac{dS}{S} = \mu dt + \sigma dz \tag{2.28}$$

and note that $\frac{d}{dS}\ln[S] = \frac{1}{S}$, so we get

$$d\ \ln[S] = \mu dt + \sigma dz. \tag{2.29}$$

The properties of the Wiener process z imply that the right hand side follows a normal distribution, which in turn means that S follows a log-normal distribution.

By using Itō's lemma[9], we can show that the value of $V(S,t)$, a derivative instrument of S, satisfies the PDE

$$dV = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dz. \tag{2.31}$$

In a continous time framework using a replicating portfolio we can show that, based on the above formulas and assuming arbitrage-free pricing, the derivative price (which equals its value) satisfies the *Black-Scholes PDE*[10]

$$rV = rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \tag{2.32}$$

[8] An even more realistic assumption is that the price process parameters μ and σ are time dependent and not constant over the entire time span, but we will not discuss that further here.

[9] Itō's lemma states that for a drift-diffusion process $dS = \mu S dt + \sigma S dz$ and $C^2(S,t)$ (twice differentiable) function $V(S,t)$ we have

$$dV(S,t) = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dz \tag{2.30}$$

[10] See [4] for the original paper by Fischer Black and Myron Scholes.

where r is the risk-free interest rate. For a european call or put option of strike K, using boundary conditions $C(0,t) = 0$ and $C(S,T) = \max(S - K, 0)$, we can solve the Black-Scholes PDE to obtain

$$C(S,t) = S \cdot \Phi(d_1) - K \cdot e^{-r\cdot(T-t)} \cdot \Phi(d_2) \tag{2.33}$$

$$P(S,t) = K \cdot e^{-r\cdot(T-t)} \cdot \Phi(-d_2) - S \cdot \Phi(-d_1) \tag{2.34}$$

where

$$d_1 = \frac{\ln[S/K] + (r + \frac{\sigma^2}{2}) \cdot (T-t)}{\sigma \cdot \sqrt{T-t}}$$
$$d_2 = d_1 - \sigma \cdot \sqrt{T-t}$$

and $\Phi(\cdot)$ is the cumulative normal distribution funtion. Note that the boundary conditions only mention the call option value. This is because we can calculate $P(S,t)$ by first calculating $C(S,t)$ and then take a shortcut to $P(S,t)$ by plugging $C(S,t)$ into the *put-call parity* condition

$$P(S,t) = C(S,t) - S(t) + K \cdot e^{-r\cdot(T-t)}, \tag{2.35}$$

and solve for $P(S,t)$. The resulting expressions for $C(S,t)$ and $P(S,t)$ are called the Black-Merton-Scholes formulas and can be used to calculate the correct price of a European option contract, given the parameters μ and σ.

When pricing options on futures contracts (and contracts similar to futures) we can use the *Black model*, which is based on the Black-Scholes model we used to obtain formulas for ordinary European options.[11] To do this, we redefine the price of the underlying asset S as the forward price F, which is the undiscounted expected future value. Where we previously assumed that the underlying asset price S followed a log-normal process, we now assume that the forward price F at maturity is log-normally distributed. This leads us to the following formulas for pricing european call and put options on a futures contract of strike K;

$$C(F,t) = e^{-r\cdot(T-t)} \cdot [F \cdot \Phi(d_1) - K \cdot \Phi(d_2)] \tag{2.36}$$

$$P(F,t) = e^{-r\cdot(T-t)} \cdot [K \cdot \Phi(-d_2) - F \cdot \Phi(-d_1)] \tag{2.37}$$

where

$$d_1 = \frac{\ln[F/K] + \frac{\sigma^2}{2} \cdot (T-t)}{\sigma \cdot \sqrt{T-t}}$$
$$d_2 = d_1 - \sigma \cdot \sqrt{T-t}.$$

[11] See [3] for the original paper by Fischer Black.

2.9 Swaptions

A swaption is an option on a swap. This means that the holder of a (European) swaption has the option, but not the obligation, to enter into a swap agreement at some point in time (the expiration date of the swaption). By convention, and in line with swaps, the option to enter into a swap agreement where you pay the fixed interest rate is called a *payer swaption* and is the analogue of a call option. Conversely, the option of paying the floating rate and receiving the fixed rate is called a *receiver swaption* and is comparable to a put option. This convention of focusing on the fixed rate is logical since a call swaption then benefits from a rise in fixed swap rates during its maturity (since its predefined rate then becomes more valuable in comparison), just as in the case of an equity option which benefits from a rise in spot price.

2.9.1. Pricing swaptions

Since the underlying instrument of an interest rate swaption is an interest rate swap contract, the source of uncertainty here is not a stock price but the forward swap rate (the fixed leg of the swap contract) at maturity. The value of this swap rate depends on market expectations of the reference interest rate at the relevant future points in time (the yield curve and forward rates). Recall from Section 2.5 on derivatives that the difference between a futures contract and the purchase of a stock is that the actual transactions take place at a future point in time. This is very similar to an interest rate swap, where we agree on a fixed rate which will determine future payments. Hence, a swaption should be treated as an option on a *futures contract* with certain characteristics, and not as an option on a *stock*. We start however, by looking at the swap contract once again.

We define three points in time, $t_0 < T < T'$, where

- t_0 is the inception date of the swaption,
- T is the maturity of the swaption *and* the inception date of the underlying swap contract,
- T' is the maturity of the swap contract.

If we adjust (2.23) with these notations for a payer swap starting at time T at the fixed rate r_s, we get

$$PV_T = N - \left(B(T,T') \cdot N + r_s \cdot \frac{N}{m} \cdot \sum_i B(T,t_i) \right). \tag{2.38}$$

This will be priced with a fixed rate r_s such that $PV_T = 0$, but note that r_s is *not known* at time t_0. After dividing by N, setting $PV_T = 0$ and some basic algebraic operations, (2.38) simplifies to

$$r_s \cdot \frac{1}{m} \sum_i B(T, t_i) = 1 - B(T, T'). \tag{2.39}$$

Now, let us step back to time t_0. We have $t_0 < T < T'$, so we do not know r_s, $B(T, t_i)$ or $B(T, T')$ yet. We can, however give it our best guess using forward rates. We recall from Section 2.4 that

$$B(t_1, t_2) = \frac{B(t_0, t_2)}{B(t_0, t_1)} \tag{2.40}$$

and note that $1 = \frac{B(t_0,T)}{B(t_0,T)}$. Using this we can rewrite (2.39) as

$$r_s \cdot \frac{1}{m} \sum_i \frac{B(t_0, t_i)}{B(t_0, T)} = \frac{B(t_0, T)}{B(t_0, T)} - \frac{B(t_0, T')}{B(t_0, T)}. \tag{2.41}$$

This equation relates r_s, the future time T value of S (fixed swap rate for swap contracts written at T) to entities observable today, so we are in fact talking about a *forward* swap rate F_T (rate at time T as seen at time t_0). After multiplying all terms by $B(t_0, T)$ we have

$$F_T \cdot \frac{1}{m} \cdot \sum_i B(t_0, t_i) = B(t_0, T) - B(t_0, T'). \tag{2.42}$$

This holds for any $t < T$,

$$F_T \cdot \frac{1}{m} \cdot \sum_i B(t, t_i) = B(t, T) - B(t, T'). \tag{2.43}$$

Note that $F_T \to S_T = r_s$ as $t \to T$. Thus, a payer swaption bought today at strike K has value $PV_{\mathrm{P}}^{K}(T, T') - PV_{\mathrm{P}}^{S_T}(T, T')$ at maturity, where $PV_{\mathrm{P}}^{r}(T, T')$ denotes the present value at T of a payer swap contract with strike rate r spanning from T to T'. This evaluates to

$$(F_T - K)^+ \cdot \frac{N}{m} \cdot \sum_i B(T, t_i). \tag{2.44}$$

This can be seen as $\frac{N}{m} \cdot \sum_i B(T, t_i)$ swaptions worth $(F_T - K)^+$ each at maturity.[12] The price of a payer swaption F with payoff $(F_T - K)^+$ at maturity is, according to the Black model with the assumption of a lognormal distribution of the stochastic process determining the forward rate,

$$\begin{aligned} C(F, t) &= e^{-r\cdot(T-t)} \left[F_T \cdot \Phi(d_1) - K \cdot \Phi(d_2)\right] && (2.45) \\ &= B(t, T) \cdot \left[F_T \cdot \Phi(d_1) - K \cdot \Phi(d_2)\right]. && (2.46) \end{aligned}$$

[12] $(x - y)^+$ is shorthand for $\max[x - y, 0]$ and is commonly used in option mathematics, since options never have negative values as we can always choose not to exercise them.

So, for $\frac{N}{m} \cdot \sum_i B(T, t_i)$ options we have a price of

$$\frac{N}{m} \cdot \sum_i B(T, t_i) \cdot B(t, T) \cdot [F_T \cdot \Phi(d_1) - K \cdot \Phi(d_2)] \tag{2.47}$$

$$= \frac{N}{m} \cdot \sum_i B(t, t_i) \cdot [F_T \cdot \Phi(d_1) - K \cdot \Phi(d_2)] . \tag{2.48}$$

We can make this expression more general by removing the assumption of constant period lengths and replacing $1/m$ with $t_i - t_{i-1}$. We could also use the interest rate r instead of discount factors by replacing $B(t, t_i)$ with $e^{-r(t,t_i)\cdot(t_i-t)}$, and rewrite the swaption price as

$$NA\,[F_T\Phi(d_1) - K\Phi(d_2)] \tag{2.49}$$

where $A = \sum_i (t_i - t_{i-1}) \cdot e^{-r(t,t_i)\cdot(t_i-t)}$. Since

$$d_1 = \frac{\ln[F_T/K] + \frac{\sigma^2}{2} \cdot (T - t)}{\sigma \cdot \sqrt{T - t}}$$
$$d_2 = d_1 - \sigma \cdot \sqrt{T - t}.$$

all entities except volatility are observable and we can calculate this at any time $t < T$, provided we find the correct value of σ to use. This is a crucial point which is common to most, if not all, options; we can agree on most things except volatility. This is why trading volatility is intimately linked with options such as swaptions; *options share many parameters, but are priced with individual values of volatility.*

Now, say that we observe two similar swaptions with a significantly different price. They are obviously priced by the issuer using different volatilities, one which we agree with and one which we believe to be incorrect. How can we profit from this situation? The situation is not as clear-cut as when talking about stocks, when the benefit from buying an undervalued share and the mechanics of making profit out of it is obvious. The methods of catching exposure to volatility in search of profits is the subject of the next chapter.

Chapter 3

Trading volatility using swaptions

Volatility is most often denoted with lower case sigma, σ, and represents the dispersion characteristics of a stochastic process. The reasons for sharing notation with the *standard deviation* will become apparent later in this chapter. Applied to asset prices volatility represents the price fluctuation magnitudes and therefore it is often quoted as *the risk*. Just as the market quotes its view of fair prices it also quotes its view of *volatility* over different time periods, even if it is not as easily accessible as the price quotes. If we are confident that the market perception of future volatility is incorrect, we would want to exploit this. An intuitive idea is to "take a position in volatility" that corresponds to our beliefs about current volatility quotations. The problem is that volatility is not as investable as foreign exchange rates or asset prices. In this section we take a closer look at volatility and how to invest in it. We recall that swaptions are options, so even if we will talk about *options* to keep the discussions general, these topics are applicable to swaptions as well.

3.1 Realized and implied volatility

There are a few types of volatility that are closely related but different. As we will see in the next sections, *realized* volatility (interchangeable called *actual* volatility) is defined with respect to the price process of the underlying asset. *Historical* volatility is a hindsight measure of realized volatility. *Implied* volatility on the other hand, is derived from the prices of *derivatives* of the underlying asset (along with other known parameters used in pricing the derivatives), such as options, and not directly from the asset itself. Since σ is a general notation of volatility it might be unclear which type we are referring to. Here we will only refer to historical volatility in Section 3.1.2 below, and then focus on realized volatility, so for clarity we will use σ_h

when referring to *historical volatility* and let σ denote *realized volatility.*

3.1.1. Realized volatility

As mentioned above and in Section 2.8, realized volatility σ is the dispersion parameter of the stochastic process $dS = \mu S dt + \sigma S dz$ assumed to govern the evolution of the asset price. Realized volatility is to volatility what spot prices are to asset prices, so it is easy to see that it will play an important role in trading volatility.

A major issue is how to calculate or approximate realized volatility, since it is a hidden characteristic of the price process and not something defined by market quotes. Many different models and techniques can be used, one of which is to use stochastic calculus. If we define $dS = \mu S\ dt + \sigma S\ dz$ as usual and look at $(dS)^2$ we get

$$(dS)^2 = (\mu S dt + \sigma S dz)^2 \tag{3.1}$$
$$= \mu^2 S^2 dt^2 + \mu\sigma S^2 dt dz + \sigma^2 S^2 dz^2. \tag{3.2}$$

Since dt is assumed to be small, we neglect the dt^2 term. $dtdz$ is also much smaller than dt so we neglect that as well. Since $dz^2 \to dt$ as $dt \to 0$ we make the substitution $dz^2 \approx dt$ to get[1]

$$(dS)^2 \approx \sigma^2 S^2 dt. \tag{3.3}$$

This means that theoretically, for small dt, we can use

$$\sigma \approx \sqrt{\frac{(dS)^2}{S^2 dt}} = \frac{dS}{S\sqrt{dt}} \tag{3.4}$$

as a measure of realized volatility.

3.1.2. Historical volatility

Historical volatility σ_h (commonly just referred to as *volatility*) is defined as the standard deviation of the logarithmic returns R on an annual basis,

$$\sigma_h = \sqrt{\lambda \cdot Var(R)} = \sqrt{\frac{\lambda}{n} \cdot \sum_{i=1}^{n} (R_i - \mu_R)^2}, \tag{3.5}$$

where n stands for the total number of observations, R_i is the one-period return $\left(S_i = S_{i-1} \cdot e^{R_i}\right)$ and λ represents the frequency of observations as observations per annum. Using daily data it is common to use $\lambda = 252$, the

[1]This approximation has more to do with order of magnitude than convergence to some specific number. For a derivation and justification of this result please refer to [2].

approximate number of business days in a year. The *logarithmic return* is defined as

$$R_i = \ln\left[\frac{P_{t+1}}{P_t}\right] \tag{3.6}$$

where P_t and P_{t+1} denotes the price or value at time t and $t+1$, respectively. From (3.5) one can deduce a handy property of the volatility; we can transform the time period in which it is defined by adding a suitable multiplication factor. This factor is the square root of the fractional change,

$$\text{multiplication factor} = \sqrt{\frac{\text{days in new period}}{\text{days in old period}}}. \tag{3.7}$$

Even though volatility is originally defined on an annual basis it could still make sense to talk about *semi-annual volatility* or *daily volatility*, though this is far less used. For example, if we wanted to transform daily volatility into yearly volatility we would multiply $\sigma_{\text{h,daily}}$ by $\sqrt{252/1} = \sqrt{252}$ (as we did above for $\lambda = 252$). If we, for some reason, would be interested in transforming annual volatility to semi-annual volatility we would instead multiply $\sigma_{\text{h,annum}}$ by $\sqrt{126/252} = \sqrt{1/2}$. Here, unless stated otherwise, volatility is always on an annual basis.

There is a couple of useful approximations available when calculating historical volatility. By noting that the expected value μ_R of the daily returns is very close to zero, we can make the approximation $\mu_R \approx 0$ to get the alternative formulation

$$\hat{\sigma} = \sqrt{\frac{\lambda}{n}\cdot\sum_{i=1}^{n} R_i^2}. \tag{3.8}$$

Also, to transform daily volatility to annual volatility we know that we should multiply by $\sqrt{\lambda} = \sqrt{252} \approx 15.9$. But, this is close to $\sqrt{256} = 16$, so to simplify the transformation from daily to annual volatility we can multiply by 16 instead to arrive at the approximation

$$\hat{\sigma}_{\text{annual}} = 16 \cdot \hat{\sigma}_{\text{daily}}. \tag{3.9}$$

3.1.3. Implied volatility

Implied volatility is the Σ that, when plugged into pricing formulas together with other (known) parameters, makes the calculated price coincide with quoted market prices V_M,

$$f(\Sigma,\cdot) = V_M \;\longrightarrow\; \Sigma = f^{-1}(V_M). \tag{3.10}$$

Here we will distinguish between realized and implied volatility by using lower case (σ) and upper case (Σ) sigma. To know the implied volatility Σ one has to know the pricing formula $f(\sigma,\cdot)$ for the asset of interest (no pun intended). We assume that everything but Σ is observable and can be regarded as constants in the formula. Usually the implied volatility can't be solved for analytically but needs to be solved for using numerical methods. Call options, as explained in Section 2.8, have a pricing formula $f(S, K, \sigma, \tau, r)$ where everything except σ can be objectively observed.[2]

As explained in the beginning of this chapter, the volatility in option pricing is the characteristic volatility of the underlying asset. In the Black-Scholes model we assume that this volatility is constant (a so-called homoskedastic process) and that historical volatility is a good estimate of future volatility, which seldom (if ever) is the case in practice. The volatility is certainly not constant over every period in time. The volatility parameter in the option pricing formula is instead used as an instrument to incorporate *expected volatility* of the underlying asset during the option's time to maturity into the option price.[3] Hence, we would expect all options with the same expiration date and underlying asset to share the same volatility, as implied by their quoted prices. However, this is generally not the case. A well-observed phenomenon is the so-called *volatility smile*, which refers to the common case where options with different strikes are quoted at different volatilities. This is of course against our basic theory that options with the same time to maturity should have the same implied volatility, but it also opens up new possibilities. By plotting not only $\sigma(K)$ (implied volatility as a function of the strike) but also $\sigma(K, \tau)$ (adding time to maturity) we can construct an *implied volatility surface*. Using this, we can extract *local volatilities*, the volatility analogue of the forward rates discussed in Section 2.3. These local volatilities, extracted from market prices of liquid vanilla options,[4] can then be used to price more complex options. Since these local volatilities are based on market data, the option prices calculated using them are entirely consistent with market expectations.[5] We will explore these concepts further in the next section.

[2] τ is the time left until maturity, $\tau = T - t$. Of course, we could replace τ in the pricing formula with t and T. Note that we often use $t = 0$, in which case τ and T is equivalent. If the underlying asset pays dividends, we would incorporate this as well in our pricing function, but it is still just another observable parameter.

[3] At least this is the idea. It is not unlikely that an issuer of options charges not only its true expectations of the volatility, but also an additional premium for bearing risk. This statement is based on the fact that option prices have a positive dependence on the volatility parameter, and the fact that options can be used to hedge away risk by loading it onto the issuer for a small premium (the option price).

[4] Vanilla options are common European options, the simplest of option contracts. A *liquid* asset means that it is widely traded, with a large pool of people willing to trade around the current price. When dealing in *illiquid* assets one could find it hard to sell or buy the desired number of shares at a reasonable price, simply because there aren't enough buyers or sellers out there.

[5] See for example [6] for more information on how local volatilities helps us derive option prices consistent with market expectations.

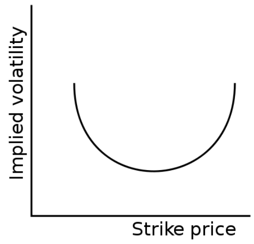

Figure 3.1: Schematic illustration of the volatility smile.

3.2 The volatility surface and local volatility

In the fundamental Black-Scholes world we assume volatility to be constant and to (partially) determine the price process of the underlying. This implies that all derivatives which share the same underlying instrument should be priced using *the same volatility*. Empirically this is not the case however. Options with strikes further away from the current spot price have a tendency to have higher implied volatilities than closer-to-the-money options, and out-of-the-money puts tend to have a higher implied volatility than equally out-of-the-money calls. This is the volatility smile mentioned above. When the differences aren't quite as obvious, we sometimes call it the *volatility smirk*.

Since options with longer maturities will be affected by market events that options with shorter maturities will not (since they will expire earlier), we would expect implied volatilities to vary between options of different maturities. We can combine this dependancy of the implied volatility on the maturity with the volatility smile by plotting the implied volatility on the z-axis with strike and maturity on the x- and y-axis. This surface plot is called the *volatility surface* and is analogous to the *yield curve* covered in Section 2.2, which tells us the yield on bonds of different maturities.

Just as bond traders interpret the yield curve, option/volatility traders interpret the volatility surface. Extending the analogy to interest rates, we recall that interest rates could be extracted at intermediate points in time from the yield curve (forward rates). Using the volatility surface we can do the same for volatilities. These volatility approximations at intermediate points or periods in time are called *local volatilities*. Just as the constant volatility in the Black-Scholes world (and in the world of stochastic processes) defines (partially) the price process over time, the local volatility corresponds to the volatility that (partially) defines the price process at that point or period in time, as stated by the implied volatilities extracted from current market prices.[6] Extracting local volatilities from market prices en-

[6]See [8] for a pedagogical introduction to the volatility surface and local volatility by

ables us to simulate the stock process more accurately and in line with market perception than before, for example using binomial or trinomial trees. This also means that we can price advanced exotic options accurately and with the confidence of taking market expectations into account, by using local volatilities from liquid vanilla options (simple and commonly traded options).[7] The mechanics of extracting local volatilities from option prices is, however, beyond the scope of this thesis.

3.3 Option hedging

Just as we can hedge a position in an underlying instrument by buying or selling appropriate amounts of certain derivatives, we can adjust the exposure of an existing derivative position by buying or selling the underlier. In fact, we can construct very specific exposures by starting with a swaption for example, and hedging away all the unwanted exposures. Here we will focus on option hedges and delta-hedges in particular. These will prove to be very practical in our efforts to obtain volatility exposure.

3.3.1. Delta-hedging

Delta, Δ, is the sensitivity of the price of a derivative V to first order price movements of the underlying instrument. It is defined as the first order derivative with respect to the price of the underlying instrument S, $\Delta = \frac{\partial V}{\partial S}$. A delta-hedged portfolio $\Pi(V(S), S)$ satisfies the condition $\frac{\partial \Pi}{\partial S} = 0$. Unfortunately, as soon as the price of the underlying changes, the initially delta-hedged portfolio is no longer completely insensitive to price changes and needs to be *re-hedged* which means adjusting the offsetting positions. This is due to the curvature of the $\frac{\partial V}{\partial S}$-curve, which means that the magnitude of the Δ is dependent on the price of the underlying S and will change as S changes. Delta-hedging is relatively simple in theory, but complicated in practice due to the trade-off between frequent re-hedging and trading costs. In practice we are also confined to a minimum size lot, which might be much bigger than what our models are advicing us to trade.

Even though there are practical limitations in delta-hedging, the strategy is the same in practice as in theory. If we own a single call option with delta (first order price sensitivity) $\hat{\Delta}$, our portfolio also has delta $\Delta_\Pi = \hat{\Delta}$. How can we make this portfolio delta-neutral? Note that the underlying asset itself has delta $\Delta = \frac{\partial S}{\partial S} = 1$. Therefore, if we take a short position of $\hat{\Delta}$ underlying contracts, the portfolio delta will be $\Delta_\Pi = \hat{\Delta} - \hat{\Delta} \cdot 1 = 0$.

Derman and Kani, the originators of the term *local volatility*.

[7] See [6], [7] and [9] for some excellent explanations of the benefits of local volatility and how to extract it from market quotes.

Generally, we delta-hedge by taking a short position in

$$\hat{\Delta}_S = \sum_{i=1}^{n} \Delta_i \tag{3.11}$$

shares in the underlying instrument, where n is the number of derivatives in our portfolio and Δ_i is the delta of derivative i. Note that Δ_i can be a negative number, depending on the nature of the derivative and the type of position. In this way, by taking an offsetting position in the underlying instrument, we can achieve a temporarily delta-neutral portfolio.

A delta-hedged portfolio with a delta close to zero even if S moves is said to have a *stable* delta-hedge, whereas a portfolio that needs frequent re-hedging is said to have an *unstable* delta-hedge. A stable delta-hedge is desirable since it reduces costs and effort.

3.3.2. Delta-Gamma-hedge

Gamma is defined as $\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 V}{\partial S^2}$ and represents the curvature of a derivative's price dependance. Since this curvature is the reason why a delta-hedged position is only insensitive to *small* price movements, Γ is a measure of the stability of a delta-hedge. One way to improve the stability of a delta-hedged portfolio is to *delta-gamma-hedge* it. This means adding a new condition $\Gamma_\Pi = \frac{\partial^2 \Pi}{\partial S^2} = 0$ to our previously stated $\Delta_\Pi = \frac{\partial \Pi}{\partial S} = 0$ condition, and constructing a portfolio that satisfies both conditions. To offset the Γ we can't trade underlying shares, since they have $\Gamma_S = \frac{\partial^2 S}{\partial S^2} = \frac{\partial}{\partial S}(1) = 0$. Instead we have to use suitable amounts and positions in other options. However, we will not cover this in dept but rather restrict our discussions to delta-hedging.

3.4 Isolating volatility exposure

Exposing ourselves to volatility by taking naked option positions is not ideal since we get exposed to several other factors as well. What we seek is to *isolate volatility exposure* to invest in pure volatility. We now explain two ways of gaining pure or approximately pure exposure to volatility.

3.4.1. Structured volatility contracts

Although beyond the scope of this thesis, it is worth mentioning that there are OTC (over the counter) products that capture pure volatility exposure. Examples include volatility swaps, variance swaps and volatility gadgets.[8] These are available only through brokers and investment banks, who then do all the work of obtaining the pure exposure for you.

[8] See for example [5] for more information on volatility contracts and [8] for more information on volatility gadgets.

3.4.2. Delta-hedging

By delta-hedging at-the-money option contracts, as we will see in the next section, one can obtain an approximately pure volatility exposure. The purity is unstable though, and needs to be reset (re-hedged) frequently. As we will see in the next section, this is a kind of volatility exposure we can manufacture for ourselves. However, in Chapter 4, we will see that there are other subtle (but important) issues also involved in catching volatility through delta-hedged portfolios.

3.5 Volatility exposure of a delta-hedged portfolio

So far we have simply stated that we can obtain exposure to volatility through delta-hedging swaptions and other options, but what type of exposure are we talking about here? This section looks to answe those questions by deriving the change in value $d\Pi$ of a delta-hedged portfolio Π, to see which parameters are affecting it.

Suppose we have a delta-hedged portfolio Π made up of an option and a delta-hedge amount of the underlying instrument, i.e. $\Pi = V + \Delta S$. We will now derive the infinitesimal change $d\Pi$ for such a portfolio, first using an ordinary option with a *stock* as the underlier and then using an option on an interest rate swap contract, i.e. a *swaption*. It might be beneficial to look through Sections 2.8 and 2.9 one more time, to get a clear view of the conceptual and mathematical differences between the two.

3.5.1. Definitions

A few definitions and equations are worth repeating since they are central in deriving $d\Pi$.

We assume that the price of the underlying asset is governed by a geometric Brownian motion

$$dS = \mu_S S dt + \sigma_S S dz, \tag{3.12}$$

and for a futures contract we assume that the futures price F at maturity follows a geometric Brownian motion

$$dF = \mu_F F dt + \sigma_F F dz. \tag{3.13}$$

We will simply use μ and σ to define the underlying processes, since we will look at them separately without risk of getting them mixed up.

We will use Itō's lemma which states that for the price of a derivative $V(x,t)$, for example an option, we have

$$dV = \left(\mu x \frac{\partial V}{\partial x} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V}{\partial x^2} \right) dt + \sigma x \frac{\partial V}{\partial x} dz. \tag{3.14}$$

We assume that the options used here are priced using the Black-Scholes model in the case of a stock option and the Black model in the case of a futures option. Central to the Black-Scholes model is the *Black-Scholes PDE* which states that, in the absence of arbitrage possibilities, the price of a derivative V of a stock S (or similar asset) subjected to the volatility σ of the underlying asset should satisfy the PDE

$$rV_\sigma = rS\frac{\partial V_\sigma}{\partial S} + \frac{\partial V_\sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_\sigma}{\partial S^2}. \tag{3.15}$$

As mentioned in Section 2.9.1, when pricing a swaption we should treat it as an option on a futures contract. For options on futures we use the Black model, an extension of the Black Scholes model, where the price of a derivative V of a futures contract F with volatility σ should satisfy the *Black PDE*

$$rV_\sigma = \frac{\partial V_\sigma}{\partial t} + \frac{1}{2}\sigma^2 F^2 \frac{\partial^2 V_\sigma}{\partial F^2}. \tag{3.16}$$

The difference between (3.15) and (3.16) is the $rS\frac{\partial V_\sigma}{\partial S}$ term. When deriving these equations we use the fact that if we are able to duplicate the effects of holding an option and exercising it at maturity using another investment strategy, the present cost and value of these two strategies should be the same. Duplicating a simple European stock option involves selling $\Delta = \frac{\partial V_\sigma}{\partial S}$ shares of the underlying asset. The money raised from this sale can be invested to yield interest. With a similar position in futures contracts however, the sale does not yield any immediate cash flow since transactions are made at maturity of the contract. Hence, the duplication of the option contract doesn't involve a term taking the sale of the underlying and the subsequent interest rates into account.

We will use the same notations as earlier;

- σ, the actual (realized) volatility which governs the price process. An option priced using the correct (but of course unknown) volatility σ will be denoted V_σ.
- Σ, implied volatility, which the option price is based on. The option price based on Σ will be denoted V_Σ.
- $\Delta = \frac{\partial V}{\partial S}$, the hedge ratio, i.e. the number of underlying contracts to hedge with.

The delta-hedged portfolios consist of

- A long position of *one* option contract (stock option or swaption).
- A short position of Δ underlying assets (stocks or interest rate swaps).

Thus, at any time, the portfolio value is $\Pi = V - \Delta S$ in the case of the stock option portfolio and $\Pi = V - \Delta F$ in the case of the swaption portfolio.

3.5.2. Using stock underlier S

We buy an option V_Σ at the implied volatility Σ and take a short position in $\Delta = \frac{\partial V_\Sigma}{\partial S}$ stocks, resulting in a portfolio

$$\Pi = V_\Sigma - \Delta S, \tag{3.17}$$

so

$$d\Pi = dV_\Sigma - d\Delta S = dV_\Sigma - \Delta dS \tag{3.18}$$

since, for each time period, Δ is just a constant. We know from (3.12) that

$$\Delta dS = \frac{\partial V_\Sigma}{\partial S} dS = \mu S \frac{\partial V_\Sigma}{\partial S} dt + \sigma S \frac{\partial V_\Sigma}{\partial S} dz. \tag{3.19}$$

and, from Itō's lemma,

$$dV_\Sigma = \left(\mu S \frac{\partial V_\Sigma}{\partial S} + \frac{\partial V_\Sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_\Sigma}{\partial S^2} \right) dt + \sigma S \frac{\partial V_\Sigma}{\partial S} dz \tag{3.20}$$

$$= \Delta dS + \left(\frac{\partial V_\Sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_\Sigma}{\partial S^2} \right) dt \tag{3.21}$$

where we made a substitution for ΔdS by matching the terms in (3.19). Notice that we have now incorporated the true volatility σ into the equations by using our assumptions on the price process and Itō's lemma. So, from (3.18) and substituting dV_Σ with (3.21) we now have

$$d\Pi = \left(\frac{\partial V_\Sigma}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_\Sigma}{\partial S^2} \right) dt. \tag{3.22}$$

However, the option V_Σ was priced under the assumption that it satisfies (3.15), which states that (after some re-arranging)

$$\frac{\partial V_\Sigma}{\partial t} = rV_\Sigma - rS\frac{\partial V_\Sigma}{\partial S} - \frac{1}{2}\Sigma^2 S^2 \frac{\partial^2 V_\Sigma}{\partial S^2}. \tag{3.23}$$

Replacing $\frac{\partial V_\Sigma}{\partial t}$ accordingly in (3.22) leaves us with

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \frac{\partial^2 V_\Sigma}{\partial S^2} dt + r\left(V_\Sigma - S\frac{\partial V_\Sigma}{\partial S} \right) dt. \tag{3.24}$$

If we continously delta-hedge then V_Σ and $\frac{\partial V_\Sigma}{\partial S} S$ should be equal and will cancel each other in the above expression. Thus, replacing $\frac{\partial^2 V_\Sigma}{\partial S^2} = \Gamma_\Sigma$, we get

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma dt. \tag{3.25}$$

This shows that through delta-hedging we have a change in portfolio value that is dependent on the difference between realized and implied volatility, multiplied by the square of the stock price and Γ. Note that neither S^2 nor Γ will change the sign of $d\Pi$, $d\Pi$ will always depend positively on $(\sigma^2 - \Sigma^2)$. It is apparent that S^2 will always be positive, and Γ will always be positive for a long option position such as the one we have here.[9] Thus, by delta-hedging, we have established an exposure to volatility exposure.

From (3.25) we can see that taking a long position in an option and delta-hedging it is a suitable strategy if we think that the implied volatility is lower than what the actual volatility will be. Since option prices have a positive dependancy on volatility, this is commonly referred to as *buying the option cheap*, referring to the low volatility we had to "pay" for the option.

If we view the option as overpriced, i.e. priced using $\Sigma > \sigma$, we should take a short position in the option and delta-hedge it by taking long positions in the underlying instrument.

3.5.3. Using futures underlier F

This derivation is completely analogous to the one for $V(S,t)$ except that we use the Black PDE instead of Black-Scholes, since we are dealing with a futures contract. We buy an option V_Σ at the implied volatility Σ and take a short position in $\Delta = \frac{\partial V_\Sigma}{\partial F}$ futures, resulting in a portfolio

$$\Pi = V_\Sigma - \Delta F, \tag{3.26}$$

so

$$d\Pi = dV_\Sigma - d\Delta F = dV_\Sigma - \Delta dF \tag{3.27}$$

since, for each time period, Δ is just a constant. We know by assumption that

$$\Delta dF = \frac{\partial V_\Sigma}{\partial F} dF = \mu F \frac{\partial V_\Sigma}{\partial F} dt + \sigma S \frac{\partial V_\Sigma}{\partial F} dz. \tag{3.28}$$

and, by Itō's lemma,

$$dV_\Sigma = \left(\mu F \frac{\partial V_\Sigma}{\partial F} + \frac{\partial V_\Sigma}{\partial t} + \frac{1}{2}\sigma^2 F^2 \frac{\partial^2 V_\Sigma}{\partial F^2} \right) dt + \sigma F \frac{\partial V_\Sigma}{\partial F} dz. \tag{3.29}$$

[9] To justify why Γ is positive for a long option position we can look at a long call option V on a stock S for example. At very low S far away from the strike K of the option, V will also be very low since it is highly unlikely that it will end up in-the-money. Since S is so far away from the strike, small movements in S wont affect the value of V since it is still approximately as unlikely that we will end up in-the-money. Hence, Δ is close to zero. If S instead were at very high levels, causing our option to be deep in-the-money, we would feel sure about the fact that S will end up above the strike making V worth $S - K$. A small movement in S hardly affects the fact that V will end up in-the-money, we are only concerned with the fact that the difference $S - K$ will be affected. Since the value of V increases by approximately \$1 for every \$1 S increases, we have Δ close to 1. Between these extreme cases we have increasing Δ with a positive dependence on S, i.e. a positive Γ. For a put option the dependence of the magnitude of Δ on S is reversed, but since the sign of Δ is negative we still have a postive Γ. Note that increasing positive Δ yields the same Γ as decreasing negative Δ. The reverse is true for short option positions.

So, from (3.27) we have

$$d\Pi = \left(\frac{\partial V_\Sigma}{\partial t} + \frac{1}{2}\sigma^2 F^2 \frac{\partial^2 V_\Sigma}{\partial F^2} \right) dt. \tag{3.30}$$

But, just as the stock option $V(S,t)$ was priced under the Black-Scholes model, the derivative V_Σ is priced under the definition that it satisfies (3.16), which states that (after re-arranging)

$$\frac{\partial V_\Sigma}{\partial t} = rV_\Sigma - \frac{1}{2}\Sigma^2 F^2 \frac{\partial^2 V_\Sigma}{\partial F^2}. \tag{3.31}$$

Replacing $\frac{\partial V_\Sigma}{\partial t}$ accordingly and writing $\frac{\partial^2 V_\Sigma}{\partial F^2} = \Gamma_\Sigma$ leaves us with

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2) F^2 \Gamma_\Sigma dt + rV_\Sigma dt. \tag{3.32}$$

The sign of the first term will be positive for the same reasons as discussed in the stock option case in the previous section. Also note that the interest rate r, the option value V_Σ, and the time step dt are all non-negative. So, just as in the case of a stock option, we have shown that by delta-hedging a swaption position we can obtain exposure to realized volatility, even though the magnitude of the PnL[10] is still dependent on F and Γ.

In delta-hedging a portfolio the delta Δ plays an important role, but so far we have only covered it conceptually and on a basic level. In the next section we will look closer at where we get Δ from and how we calculate it.

3.6 Delta-hedging in detail

The delta of a derivative V of S is defined as $\Delta = \frac{\partial V}{\partial S}$, but how do we determine Δ in practice? As it turns out, in the case of European option contracts, there are closed formulas for Δ;

$$\Delta_C = \Phi(d_1) \tag{3.33}$$
$$\Delta_P = \Phi(d_1) - 1. \tag{3.34}$$

These refer to a call and a put option, respectively, where $\Phi(\cdot)$ and d_1 are the same parameters as in Section 2.8. The rest of this section proves these results and addresses a few additional points on delta-hedging.

[10]PnL is an abbrevation of *profits and losses*. The expression is widely used, and even though the exact meaning can differ depending on the nature of the business and strategy, it is most commonly used to denote the monetary profits or losses over some specific period in time.

Proof: First of all, let us explicitly define the normal cumulative distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy. \tag{3.35}$$

We recall that d_1 is defined by

$$d_1 = \frac{\ln[S/K] + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} \tag{3.36}$$

where $\tau = T - t$ is time to maturity. We will start with the case of a call option. We know that the price of a European call is

$$C = S\Phi(d_1) - Ke^{-r\tau}\Phi(d_2) \tag{3.37}$$

where $d_2 = d_1 - \sigma\sqrt{\tau}$, so

$$\Delta = \frac{\partial C}{\partial S} = \Phi(d_1) + S\Phi'(d_1)\frac{\partial d_1}{\partial S} - Ke^{-r\tau}\Phi'(d_2)\frac{\partial d_2}{\partial S}. \tag{3.38}$$

So, to prove that $\Delta = \Phi(d_1)$ we need to prove that

$$S\Phi'(d_1)\frac{\partial d_1}{\partial S} = Ke^{-r\tau}\Phi'(d_2)\frac{\partial d_2}{\partial S}. \tag{3.39}$$

where

$$\Phi'(d_i) = \left(\frac{\partial}{\partial x}\Phi(x)\right)\bigg|_{x=d_i} \tag{3.40}$$

for $i = 1, 2$. Looking at the different components we identify and calculate

$$\Phi'(d_1) = \frac{1}{\sqrt{2\pi}} e^{-d_1^2/2} \tag{3.41}$$

$$\Phi'(d_2) = \frac{1}{\sqrt{2\pi}} e^{-d_2^2/2} \tag{3.42}$$

$$\frac{\partial d_1}{\partial S} = \frac{\partial d_2}{\partial S} = \frac{1}{\sigma\sqrt{\tau}S} \tag{3.43}$$

where (3.41) and (3.42) follows from the fundamental theorem of calculus which states that $\frac{\partial}{\partial x}\int_a^x f(t)\ dt = f(x)$, and (3.43) is obtained by direct calculation. By inserting these expressions into (3.39) and re-arranging we can formulate the problem as proving that the equality

$$S/K = e^{-r\tau} e^{\frac{1}{2}(d_1^2 - d_2^2)} \tag{3.44}$$

holds. We start by looking at the right hand side of the equation. If we take the logarithm of the right hand side of (3.44) we see that

$$\ln\left[e^{-r\tau} e^{\frac{1}{2}(d_1^2 - d_2^2)}\right] = \frac{1}{2}(d_1^2 - d_2^2) - r\tau \tag{3.45}$$

$$= \frac{1}{2}(d_1 + d_2)(d_1 - d_2) - r\tau. \tag{3.46}$$

Replacing $d_2 = d_1 - \sigma\sqrt{\tau}$ we see that

$$\frac{1}{2}(2d_1 - \sigma\sqrt{\tau})\sigma\sqrt{\tau} - r\tau = \left(d_1 - \frac{1}{2}\sigma\sqrt{\tau}\right)\sigma\sqrt{\tau} - r\tau. \tag{3.47}$$

At this point we insert the definition of d_1 from (3.36) to get

$$\left(\frac{\ln[S/K] + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}} - \frac{1}{2}\sigma\sqrt{\tau}\right)\sigma\sqrt{\tau} - r\tau \tag{3.48}$$

$$= \ln[S/K] + r\tau + \frac{1}{2}\tau\sigma^2 - \frac{1}{2}\sigma^2\tau - r\tau \tag{3.49}$$

$$= \ln[S/K] \tag{3.50}$$

and thus the right hand side of (3.44) is equal to S/K. Hence we have proved that, for a call option,

$$\Delta_C = \Phi(d_1). \tag{3.51}$$

The equivalent expression for a put option is $\Delta_P = \Phi(d_1) - 1$, which can be derived in the same way as above or using the *put-call parity*

$$P = C + Ke^{r\tau} - S \tag{3.52}$$

which implies

$$\Delta_P = \frac{\partial}{\partial S}P = \frac{\partial}{\partial S}C - 1 \tag{3.53}$$

$$= \Phi(d_1) - 1. \ \square \tag{3.54}$$

The cricial point of this is that Δ is computable using S, K, r, τ and σ. All these parameters except σ are easily observable, so in order to hedge with Δ in this way we need to decide on what volatility σ to use. We will now look at two natural choices; the implied volatility observed in the market, and the actual volatility we believe to be correct. We cannot truly know the actual volatility of course, but this analysis still provides us with some interesting results.

3.6.1. Hedging with actual volatility σ

The present value of the total profit can be shown to be[11]

$$PV(\text{total profit}) = V_\sigma - V_\Sigma \tag{3.55}$$

where V_σ is the price of V calculated using σ, and V_Σ is the price of V calculated using Σ. The change in portfolio value at any one time can likewise be shown to be

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2)S^2\Gamma_\Sigma\, dt + (\Delta_\Sigma - \Delta_\sigma)((\mu - r)S\, dt + \sigma S\, dz). \tag{3.56}$$

[11] See Appendix B.1 for a derivation of these formulas.

Looking at this expression for $d\Pi$ we see that there is a stochastic term dz involved. Remembering that the overall (present value) profit of $V_\sigma - V_\Sigma$ still holds, we interpret this as follows: We *do* secure a profit of $V_\sigma - V_\Sigma$, but the path of mark to market value as we get there is random.[12] We could illustrate this as a set of possible "profit paths" starting at the same point, moving up and down randomly and independent from each other, but converging to $V_\sigma - V_\Sigma$ as $t \to T$.

3.6.2. Hedging with implied volatility Σ

When hedging with implied volatility we get an expression for the instantaneous change in portfolio value of the form[13]

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2)S^2\Gamma_\Sigma \, dt \tag{3.57}$$

and a present value of the total profits

$$PV(\text{total profits}) = \frac{1}{2}(\sigma^2 - \Sigma^2)\int_{t_0}^{T} e^{-r(t-t_0)}S^2\Gamma_\Sigma \, dt. \tag{3.58}$$

Once again we ask ourselves if there are other things than the difference in volatility affecting the total profits. By looking at the components of the integral, we see that each component is strictly positive. This means that the incremental change in total profits will always be positive, assuming that $\sigma > \Sigma$, but highly path dependent due to the S^2 term.

[12] Mark to market value is the value of our portfolio at some point in time, calculated by observing the market price/value of each of our assets at that particular point in time.

[13] See Appendix B.2 for a derivation of these formulas.

Chapter 4

From theory to practice

In the light of the simple mechanics of delta-hedging, it is clear that exposure to volatility is within reach also for the individual investor. Before we start trading or testing a strategy, there are a few further issues to address.

4.1 The choice of option strategy

So far we have talked about owning a single option and delta-hedging it. Of course we could start with *two* options of the same type and delta-hedge them individually to still end up with a delta-hedged option portfolio. The number of underlying contracts would just be twice as large compared to the case of hedging a single option, negative or positive depending on whether we have a long or short position in the options. An unwanted effect of hedging two options instead of one is that we have to double the hedging amounts, which increases the costs of trading. Fortunately there is a solution to this. The results presented are valid for both call and put options since they both have positive Γ.[1] Also remember that calls always have positive Δ while puts always have negative Δ. If we combine these two insights, it is obvious that it would be beneficial to let our volatility trading delta-hedged option portfolio consist of *one call and one put*, since the hedging amounts to a large extent will *offset each other*, resulting in cheaper hedging. For a portfolio Π consisting of one call option V_c, one put option V_p, a short position in Δ_c stocks to hedge the call and a long position in Δ_p stocks to hedge the put, we have

$$\Pi = V_p + V_c + (\Delta_p - \Delta_c) \, , \tag{4.1}$$

[1] In theory they even have the *same* Γ.

and from (3.25) we know that

$$d\Pi = d\Pi_c + d\Pi_p \tag{4.2}$$
$$= \frac{1}{2}(\sigma^2 - \Sigma_c^2)S^2\Gamma_{\Sigma,c}dt + \frac{1}{2}(\sigma^2 - \Sigma_p^2)S^2\Gamma_{\Sigma,p}dt. \tag{4.3}$$

Such an option strategy, where we use one call and one put that are equal in their specifications and delta-hedge them both, is called a *delta hedged straddle*. It is common to add "long" or "short" to denote whether we are buying or selling the options. We use ATM (at-the-money) options since both ATM calls and puts have a delta of approximately 0.5, which makes the hedge amounts almost completely cancel each other out.[2]

4.2 Backtesting

Before trading a strategy in practice it is common practice to let it go through a trial phase, where we might not even trade with real money. A common practice is to simulate a track record by applying the trading algorithm on historical data. This is called *backtesting*. This is highly recommended since we are not restricted to calculating the final PnL. We can apply any number of statistical techniques and analyze the results to draw conclusions on the statistical properties of the strategy PnL. Even so, it is important to maintain a sober skepticism and remember that we are operating in a world governed largely by randomness. These are a few of the issues we need to address:

- **Our beliefs about the market** Are options currently overvalued or undervalued when it comes to volatility? As explained in Section 4.1, if we believe they are currently or systematically *undervalued* ($\sigma > \Sigma$) we should construct a *long* straddle, if we believe they are *overvalued* ($\sigma < \Sigma$) we should construct a *short* straddle.

- **Which volatility to hedge with** Do we want to hedge using σ or Σ? A strong reason for choosing Σ is that it is easily observable, while σ is not. The down side of hedging with Σ is the stochastic nature of final profit, as discussed in Section 3.6.2. By hedging with σ we have a theoretically guaranteed and precise profit locked in from the start, but we should probably take this benefit of σ-hedging lightly since it is based on assumptions of continous hedging and known σ, none of which are easily (if at all) transferred from theory to practice.

  We should also remember that the strictly theoretical parameters σ and Σ represents the dispersion of a stochastic process which is associated with the lognormal distribution. This distribution assumption

[2] ATM options have a strike equal to the current spot price or forward rate, i.e. "a strike equal to the price we observe for the underlying asset today".

should of course be scrutinized and taken into consideration when calculating *our* view of the correct option price (and thus our view of the correct volatility parameter value). Since the pricing formula itself, obtained using the Black-Scholes and Black model, is derived under the assumption of lognormal distribution, it is easy to see that an analysis of how the probability distribution affects the option price and volatility parameter quickly becomes very complex.

- **Tracking the strategy performance** Regardless of whether we are backtesting the strategy or tracking its' performance in the market we would like to have the ability to measure the *value* of our positions. There are different ways of doing this which brings us to another necessary decision: how closely and in how many ways must we be able to measure the performance of the strategy? Is end-of-day mark-to-market value enough or should we be able to derive intraday $d\Pi$ as well? The answers to these questions will give rise to different sets of obstacles that need to be addressed before we can handle the strategy in a professional manner. Analyzing intraday data can be a very messy business, as well as costly. Market data suppliers charge a substantial amount of money to aggregate and supply intraday data of high quality.

- **Tracking $d\Pi$ and realized volatility** The technique for tracking $d\Pi$ described in Section 3.5 is entirely based on the Black-Scholes and Black models and the assumption that we can track realized volatility, but one big question remains. How do we accurately track realized volatility? The method of extracting σ from $(dS)^2$ as described in Section 3.1.1 is only theoretically valid for small dt, so if we are using end-of-day data we will get very unstable results. We might be forced to use a combination of parametric and non-parametric models, probably requiring high-resolution data which might be hard to obtain. This should be taken into account before we start implementing the strategy, so that we can plan and design our PnL analysis properly.

4.3 Conclusions

Through chapters 2 and 3 we have provided a stringent coverage of the mathematics necessary to price options on interest rate swaps, called *swaptions*. In Chapter 3 we used options theory to prove that we could obtain volatility exposure by trading swaptions and the underlying interest rate swap. Here in Chapter 4 we discussed the transition from theory to practice. The difficulties lies in the availability of high quality data to test our strategies on, as well as in interpreting results arising from impure volatility exposures. Nevertheless, this thesis provides a good foundation for further research on the profitability of a delta-hedge swaption strategy. This would of course in essence be a study of whether swaptions are consistently under- or overvalued in respect to the volatility parameter. The mathematics are known and we wouldn't need more tools than a spreadsheet. All we need is the data. As usual, *information is the key.*

Appendix A

Symbols and notations

Table A.1: List of symbols and notations

B	Bond value/price
c	Coupon payment
C	Call option
d	Differential operator, or infinitesimal change
δ	Small change, e.g. $\delta t = t_{i+1} - t_i$
Δ	Option delta defined as $\partial V/\partial S$
f	Frequency of coupon payments defined by $f = 1/m$
F	Underlying futures contract
fr	Forward (interest) rate
Γ	Option gamma defined as $\partial^2 V/\partial S^2$
K	Strike price
m	Coupon payments per year
μ	Drift parameter in underlying process
P	Put option
Π	Portfolio value
r	Interest rate
S	Underlying stock or similar
σ	Realized/actual volatility
Σ	Implied volatility
t	Time variable
t_0	Present time or start of contract period
T	Maturity
τ	Time to maturity defined by $\tau = T - t$
Θ	Option theta defined as $\partial V/\partial t$
V	Option value/price
z	Standard Brownian motion (Wiener process)

Appendix B

Derivations

Here we derive the incremental (infinitesimal) changes in the value of a delta-hedged option portfolio, as well as the total profit made over some period (t_0, T), that we used in Section 3.6.

B.1. $d\Pi$ when delta-hedging using actual volatility σ

This derivation follows largely Section 3 of [1] by Ahmad and Wilmott.

We start with a portfolio Π consisting of an option V bought at implied volatility Σ and delta-hedged using actual volatility σ, and cash invested in an instrument or account yielding interest rate $r \cdot dt$ at each time step. Delta-hedging a long position means selling Δ_σ shares of the underlying.

$$\Pi = V_\Sigma - \Delta_\sigma S + (\Delta_\sigma S - V_\Sigma). \tag{B.1}$$

The change in value one time step dt forward, if the cash part $\Delta_\sigma S - V_\Sigma$ is invested yielding an interest rate r, is

$$d\Pi = dV_\Sigma - \Delta_\sigma dS + (\Delta_\sigma S - V_\Sigma) r \ dt. \tag{B.2}$$

If the option would have been correctly priced at t_0 so that $\Sigma = \sigma$, we would have $d\Pi = 0$ due to the arbitrage-free pricing. That is,

$$dV_\sigma - \Delta_\sigma dS + (\Delta_\sigma S - V_\sigma) r \ dt = 0. \tag{B.3}$$

Since (B.3) is equal to zero we can subtract it from (B.2) without violating the equality. (B.2)−(B.3) yields

$$\begin{aligned} d\Pi &= dV_\Sigma - dV_\sigma + (V_\sigma - V_\Sigma) r \ dt && \text{(B.4)} \\ &= dV_\Sigma - dV_\sigma - (V_\Sigma - V_\sigma) r \ dt && \text{(B.5)} \\ &= d(V_\Sigma - V_\sigma) - (V_\Sigma - V_\sigma) r \ dt. && \text{(B.6)} \end{aligned}$$

Before integrating to obtain the total change in value we can rewrite this in a clever way by noting that

$$
\begin{aligned}
d\left(e^{-rt}\cdot(V_\Sigma - V_\sigma)\right) &= e^{-rt}\cdot d\ (V_\Sigma - V_\sigma) + (V_\Sigma - V_\sigma)\cdot d\ e^{-rt} && \text{(B.7)}\\
&= e^{-rt}\cdot d\ (V_\Sigma - V_\sigma) - e^{-rt}\cdot (V_\Sigma - V_\sigma) r\ dt && \text{(B.8)}\\
&= e^{-rt}\cdot \Big(d\ (V_\Sigma - V_\sigma) - (V_\Sigma - V_\sigma) r\ dt\Big) && \text{(B.9)}\\
&= e^{-rt}\cdot d\Pi. && \text{(B.10)}
\end{aligned}
$$

So, we can write the infinitesimal value change as

$$d\Pi = e^{rt} d\left(e^{-rt}(V_\Sigma - V_\sigma)\right) \tag{B.11}$$

and the present value of this instantaneous value change as

$$
\begin{aligned}
PV(d\Pi) &= e^{-r(t-t_0)} e^{rt} d\left(e^{-rt}(V_\Sigma - V_\sigma)\right) && \text{(B.12)}\\
&= e^{rt_0} d\left(e^{-rt}(V_\Sigma - V_\sigma)\right). && \text{(B.13)}
\end{aligned}
$$

When calculating the present value of the total profit (the sum of all instantaneous profits) using an integral over (t_0, T), we see the benefit of rewriting $d\Pi$ as in (B.13) as we get

$$
\begin{aligned}
PV(\text{total profit}) &= e^{rt_0}\cdot \int_{t_0}^{T} d\left(e^{-rt}(V_\Sigma - V_\sigma)\right) && \text{(B.14)}\\
&= e^{rt_0}(V_\Sigma - V_\sigma)\cdot \int_{t_0}^{T} d\left(e^{-rt}\right) && \text{(B.15)}\\
&= e^{rt_0}(V_\Sigma - V_\sigma)\cdot \Big[e^{-rt}\Big]_{t=t_0}^{t=T} && \text{(B.16)}\\
&= e^{rt_0}(V_\Sigma - V_\sigma)\cdot \left(e^{-rT} - e^{-rt_0}\right) && \text{(B.17)}\\
&= (V_\Sigma - V_\sigma)\cdot \left(e^{rt_0 - rT} - e^{rt_0 - rt_0}\right) && \text{(B.18)}\\
&= (V_\Sigma - V_\sigma)\cdot \left(e^{rt_0 - rT} - 1\right) && \text{(B.19)}\\
&= (V_\sigma - V_\Sigma)\cdot \left(1 - e^{rt_0 - rT}\right) && \text{(B.20)}\\
&\approx V_\sigma - V_\Sigma. && \text{(B.21)}
\end{aligned}
$$

Where we have used the approximation $e^{rt_0 - rT} \approx 0$ assuming T is large with respect to t_0. This shows that, assuming that we are correct in our prediction of actual volatility, we lock in a guaranteed approximate profit of $V_\sigma - V_\Sigma$.

One thing left to find out is how the value Π of our position changes during the period, i.e. the formula for $d\Pi$ and how it depends on σ. To find the final dependancy of $d\Pi$ on σ we look recall (B.2) and expand dV_Σ using Itō's lemma. Just as we denote the first order derivative of the price of a derivative contract V with respect to the underlying instrument *delta*

$\Delta = \frac{\partial V}{\partial S}$, we usually denote the first order derivative with respect to *time* t or *time to maturity* τ by *theta* $\Theta = \frac{\partial V}{\partial t}$ or $\Theta = \frac{\partial V}{\partial \tau}$. Since $\tau = T - t$, they both express the time dependance of the change in option value, and it is a matter of taste which definition to use. For our purposes, since Θ is cancelled out in our formulas, it doesn't matter which one we choose.

$$\begin{aligned} d\Pi &= \Theta_\Sigma \ dt + \Delta_\Sigma \ dS + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma \ dt - \Delta_\sigma dS + (\Delta_\sigma S - V_\Sigma) r \ dt && \text{(B.22)} \\ &= \Theta_\Sigma \ dt + (\Delta_\Sigma - \Delta_\sigma) dS + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma \ dt + (\Delta_\sigma S - V_\Sigma) r \ dt. && \text{(B.23)} \end{aligned}$$

We remember that the price process of the underlying is assumed to be $dS = \mu S \ dt + \sigma S \ dz$ where dz is a standard Brownian motion (Wiener process). Expanding dS in this manner and re-arranging the terms yields

$$d\Pi = \left(\Theta_\Sigma + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma + (\Delta_\sigma S - V_\Sigma) r \right) dt + (\Delta_\Sigma - \Delta_\sigma)(\mu S \ dt + \sigma S \ dz). \tag{B.24}$$

But V is priced under the assumption of fulfilling the Black Scholes equation

$$r V_\Sigma = \Theta_\Sigma + r S \Delta_\Sigma + \frac{1}{2}\Sigma^2 S^2 \Gamma_\Sigma \tag{B.25}$$

which is equivalent to saying

$$\Theta_\Sigma - r V_\Sigma = -r S \Delta_\Sigma - \frac{1}{2}\Sigma^2 S^2 \Gamma_\Sigma. \tag{B.26}$$

Substituting this in our expression for $d\Pi$ and rearranging yields

$$\begin{aligned} d\Pi = &\left(\frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma + (\Delta_\sigma S - \Delta_\Sigma S) r \right) dt \\ &+ (\Delta_\Sigma - \Delta_\sigma)(\mu S \ dt + \sigma S \ dz) && \text{(B.27)} \\ = &\frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma \ dt - (\Delta_\Sigma - \Delta_\sigma) r S \ dt \\ &+ (\Delta_\Sigma - \Delta_\sigma)(\mu S \ dt + \sigma S \ dz) && \text{(B.28)} \\ = &\frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma \ dt + (\Delta_\Sigma - \Delta_\sigma)((\mu - r) S \ dt + \sigma S \ dz) && \text{(B.29)} \end{aligned}$$

where we see the expected dependancy on σ. For a short discussion on this dependancy, please refer to Section 3.6.

B.2. $d\Pi$ when delta-hedging using implied volatility Σ

This derivation follows largely Section 4 of [1] by Ahmad and Wilmott. We start of with the same portfolio as in the previous section but hedge with *implied volatility* Σ instead of actual volatility σ. The change in value of the portfolio is

$$d\Pi = dV_\Sigma - \Delta_\Sigma dS + (\Delta_\Sigma S - V_\Sigma) r \ dt. \tag{B.30}$$

From the assumptions on the price process of the underlying and Itō's lemma we have

$$\Delta_\Sigma dS = \Delta_\Sigma \mu S \ dt + \Delta_\Sigma \sigma S \ dz \tag{B.31}$$

$$dV_\Sigma = \left(\mu S \Delta_\Sigma + \Theta_\Sigma + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma\right) dt + \sigma S \Delta_\Sigma \ dz \tag{B.32}$$

respectively. Inserting these expressions into (B.30) yields

$$d\Pi = \left(\Theta_\Sigma + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma + (\Delta_\Sigma S - V_\Sigma) r\right) \ dt \tag{B.33}$$

$$= \left(\Theta_\Sigma - V_\Sigma r + \frac{1}{2}\sigma^2 S^2 \Gamma_\Sigma + \Delta_\Sigma S r\right) \ dt. \tag{B.34}$$

Now, V_Σ satisfies the Black-Scholes equation by definition, so

$$\Theta_\Sigma - rV_\Sigma = -rS\Delta_\Sigma - \frac{1}{2}\Sigma^2 S^2 \Gamma_\Sigma. \tag{B.35}$$

Replacing $\Theta_\Sigma - rV_\Sigma$ in (B.34) gives

$$d\Pi = \frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma \ dt. \tag{B.36}$$

The present value of this is

$$PV(d\Pi) = e^{-r(t-t_0)} \frac{1}{2}(\sigma^2 - \Sigma^2) S^2 \Gamma_\Sigma \ dt, \tag{B.37}$$

so the present value of the total profits is

$$PV(\text{total profit}) = \frac{1}{2}(\sigma^2 - \Sigma^2) \int_{t_0}^{T} e^{-r(t-t_0)} S^2 \Gamma_\Sigma \ dt. \tag{B.38}$$

Bibliography

[1] Riaz Ahmad and Paul Wilmott. *Which Free Lunch Would You Like Today, Sir?: Delta Hedging, Volatility Arbitrage and Optimal Portfolios.* Wilmott magazine (64-79), November 2005.

[2] Martin Baxter and Andrew Rennie. *Financial calculus: an introduction to derivative pricing*, chapter 3.3 Itō calculus, pages 57–59. Cambridge University Press, 1996.

[3] Fischer Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, 1976.

[4] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[5] Peter Carr and Dilip Madan. Towards a theory of volatility trading, January 2002.

[6] Neil Chriss, Emanuel Derman, and Iraj Kani. *Implied Trinomial Trees of the Volatility Smile.* Quantitative Strategies Research Notes. Goldman Sachs, February 1996.

[7] Emanuel Derman, Michael Kamal, and Iraj Kani. *Trading and Hedging Local Volatility.* Quantitative Strategies Research Notes. Goldman Sachs, August 1996.

[8] Emanuel Derman, Michael Kamal, Iraj Kani, John McClure, Cyrus Pirasteh, and Joseph Z. Zou. Investing in volatility. *Futures and Options World*, Special Supplement on the 25th Anniversary of the Publication of the Black-Scholes Model, 1998.

[9] Emanuel Derman, Iraj Kani, and Joseph Z. Zou. *The Local Volatility Surface - Unlocking the Information in Index Option Prices.* Quantitative Strategies Research Notes. Goldman Sachs, December 1995.