Centre for Central Banking Studies

Deriving option-implied probability densities for foreign exchange markets

Andrew Blake and Garreth Rule

BANK OF ENGLAND

Deriving option-implied probability densities for foreign exchange markets

Andrew P Blake and Garreth Rule

Andrew.Blake@bankofengland.co.uk
Garreth.Rule@bankofengland.co.uk

This Handbook describes an important part of the *Empirical Finance for Monetary Policy* course run at the CCBS for the past few years, the use of options prices to estimate risk-neutral forecast densities for financial market prices. Central banks around the world make routine use of this information to assess market sentiment for a variety of asset classes. These for example include equities, oil futures and the focus of this Handbook, foreign exchange. With them, policymakers can assess risk-neutral predicted densities not just the most likely values. This provides a rich forecast information set, so that not just the most likely outcome can be ascertained, but also the market assessment of, say, any potential skewness in the density of possible outcomes.

This whole area might seem impenetrable to the casual reader, needing as it does option pricing, density estimation and some mastery of a suitable computer language, but it turns out to be remarkably self-contained and easy to implement. With a relatively simple set of tools, all described here, it is a straightforward matter to estimate densities. We give a complete how-to guide, from finding appropriate data to giving the complete MATLAB code needed to get up and running. We focus on foreign exchange options for reasons described in the text, but our general approach to density estimation follows a similar methodology to that found in textbooks such as Taylor (2005), who implements similar techniques in Excel for equities.

The views expressed in this *Technical Handbook* are those of the authors, and are not necessarily of the Bank of England.

Contents

Introduction

Options are financial instruments that give the purchaser the *right* but *not the obligation* to purchase or sell a defined asset at a specified price on a future date. Options, like futures contracts[1] initially came about from the need for insurance. For example, a farmer planting a crop may look to purchase an options contract to ensure a minimum price at which he can sell his goods when it comes time to harvest them. A futures contract, where the farmer agrees a minimum price at which he must sell, insures against the downside risk for the purchaser but also eliminates potential upside gains. By contrast, options contracts can be left to expire unused should it be beneficial to the holder of the option to so do. In the case of the farmer this means he can exercise his option and sell at a pre-agreed price if the prevailing market price falls below the agreed level. But equally the farmer can pass up the options contract and sell at the prevailing market price if this is higher than the option contract guarantees.

While options contracts began as bilateral transactions between buyers and sellers, many transactions now take place on centralised exchanges. The move to exchange trading has permitted the standardisation of contracts and opened the markets to a wider range of participants. Today options contracts are traded daily on a vast array of underlying assets ranging from agricultural produce through financial instruments such as exchange rates and interest rates to more esoteric items such as the weather and economic events.

Central banks may purchase options as part of their regular interactions with financial markets for policy purposes. For example, a central bank may use options as a way of hedging against risk in its management of foreign exchange reserves. But options are also of interest to analytical areas of the central bank, because options reveal the beliefs of markets participants. Options contracts allow us to infer what market participants believe about the likelihood of future events. Participants in the options markets are after all insuring themselves against a particular market outcome; and the price that they are willing to pay will reflect the perceived likelihood of such outcomes occurring.

The purpose of this technical handbook is to provide the reader with the intuition and the MATLAB®[2] code required to draw risk-neutral densities[3] for expectations of future asset price movements. In order to make the different methods transparent we focus on options written on one asset class, foreign exchange. Exchange rates play an important role in determining domestic inflation and output in an economy. For practical purposes, a good reason for looking at exchange rates is to take advantage of their bilateral nature. While many developing world countries lack deep and liquid option markets, such markets exist in major financial centres. Options written on

---

[1] As opposed to options, the holder of the futures contract has the obligation to purchase or sell the specified asset at the price agreed within the contract.

[2] MATLAB® is a registered trademark of [The MathWorks Inc](#).

[3] Under *risk-neutral* pricing, true subjective probabilities of an event happening are replaced by risk-neutral probabilities. They are required in asset pricing as they allow us to represent an asset's price as the discounted expected value of the asset's payoff without having to adjust for compensation for risk. The risk-neutral probability of any state of the world can be thought of the product of the true probability of that state occurring and the utility derived by the agent in that state. Hence risk-neutral probabilities correct for bearing risk. Risk-neutral can be a misleading terminology, as crucially risk-neutral probabilities place *higher* event-probabilities on states with *higher* marginal utility or disutility. Cochrane (2005) uses the example of people who report high probabilities of plane crashes, suggesting that instead of being irrational they are in fact reporting their risk-neutral probabilities; it is just that the personal disutility gained from the possibility of dying in a plane crash is extremely high. The conceptual conversion from real-world probabilities to risk-neutral probabilities overweights the probability of bad states reflecting that agents are risk averse and thus require compensation for risk.

developing world currencies are likely to be traded against the US dollar in US markets and so options data are readily available for many countries that themselves do not have domestic options markets. It is hoped that by focussing on this particular asset class that these methods will prove accessible to many researchers in central banks and elsewhere.

The Handbook is organised as follows. Section 1 gives a brief introduction to options, introducing both terminology and the structure of options contracts. Section 2 discusses options pricing and the key concept of implied volatility. Section 3 describes the Breeden-Litzenberger result which is the key insight into the derivation of risk-neutral densities from option pricing. Section 4 discusses the unique features of options contracts written on foreign exchange and discusses some of the data sources for these prices. Sections 5 and 6 discusses and implements two different methods for deriving option-implied probability densities from options written on foreign exchange markets both proposed by Alan Malz (1997, 2014). The first, covered in Section 5, which we will refer to as the parametric approach, involves solving simultaneously for the so-called delta of an option and its implied volatility using just three options contracts to parameterise a functional form for the volatility smile. The second, covered in Section 6 and referred to as the non-parametric approach, imposes no such restrictions on the shape of the volatility smile; instead it uses clamped cubic-spline functions to interpolate between observed volatility points. Finally Section 7 discusses how to interpret the densities derived.

1.   What are options?

Structure of an options contract

The purchaser of an option is often referred to as the **holder** and is said to be **long** the option, while the seller of an option is referred to as the **writer** and is said to be **short** the option. Ahead of expiry the writer of an option risks unknown losses, while the holder of the contract risks only the price they paid for the option.[4,5] A contract which gives the purchaser of the option the right to buy the asset at some future date is referred to as a **call option** and a contract that gives the purchaser of the option the right to sell the asset at some future date is referred to as a **put option**.

Within each option contract the following will be clearly stated:

- An **expiry date**; for *European options* the contract can only be exercised on this date, while for *American options* the contract can be exercised at any time up to this date.[6,7] In what follows we concentrate on European options.
- A **strike price**; this is the price which is agreed in advance that the underlying can be transacted at once the contract is exercised.
- The **underlying asset;** this is the specified asset upon which the contract is based. For exchange traded option these are often standardised quantities.

---

[4] No rational holder of an option would consider exercising the option if it would lead to a loss.

[5] For exchange-traded options there may be fees to pay which may increase the costs.

[6] However, it can be shown that it is always optimal not to exercise an American call option ahead of expiry, see Taylor (2005).

[7] The nomenclature for European and American options can be traced to Paul Samuelson, who as an act of revenge against European options traders in New York who he thought had slighted him, decided to christen the simpler form of option European and the more complex form American. http://www.afajof.org/details/video/2870851/Paul-Samuelson-Interview.html

Options are said to be **in-the-money** when the difference between the price of the underlying and the strike price means it would be beneficial for the holder to exercise the options. Equivalently, options where the difference between the underlying and strike prices means the holder should not exercise the option are said to be **out-the-money**. When the prices are identical the option is said to be **at-the-money**.

Payoffs

As we will see, unlike a futures contract, an option contract has a value *ahead* of expiry. Therefore the purchaser of an option will pay a premium to hold the option. That premium is a function of the *payoff*. Consider a European call option which gives the right to *buy* at the price and date specified in the contract. The option expires at a specified time T, is bought at some prior time t for the premium $C(t,T)$. Notice the cost of the option is a function of time to expiry, $T - t$. At the expiry date the payoff to the holder is

$$Payoff_T = \max[S_T - X_T, 0]$$

where S_T is the price of the underlying asset at the time of expiry and X_T is the strike price. Potential profits for the holder of the call option are unlimited; if the prevailing market price is above the strike price then the holder of the option should exercise the option and buy the underlying at the strike price before immediately selling the underlying back to the market at the prevailing price, making a profit of the payoff less the premium. Should the prevailing market price be below the strike price, the holder should let the option expire and receive no payoff. Losses to the *holder* are therefore limited to the initial premium. Losses to the *writer* of the option are potentially limitless as they must sell the asset at the agreed strike price, no matter what the prevailing market price has become. The potential profit for the writer is limited to the premium.

Similarly the payoff to the holder of a European put option (the right to sell) is as follows:

$$Payoff_T = \max[X_T - S_T, 0]$$

For a put option profits for the holder are limited, as the price of the underlying can never be lower than zero. If the prevailing market price is below the strike price then the holder can realise profits by buying the underlying in the spot market at the prevailing price and exercising the option to sell at the agreed strike price. If the prevailing price is above the strike price then the holder should let the option expire and receive no payoff. Losses to the writer of the option are also limited by the zero bound to asset prices (but could still be very large). Sample payoffs (and profits) for the holder and writer of call and put options are summarised graphically in **Table 1**.

Table 1. Option payoffs

Call payoff for holder	Put payoff for holder

Call payoff for writer	Put payoff for writer
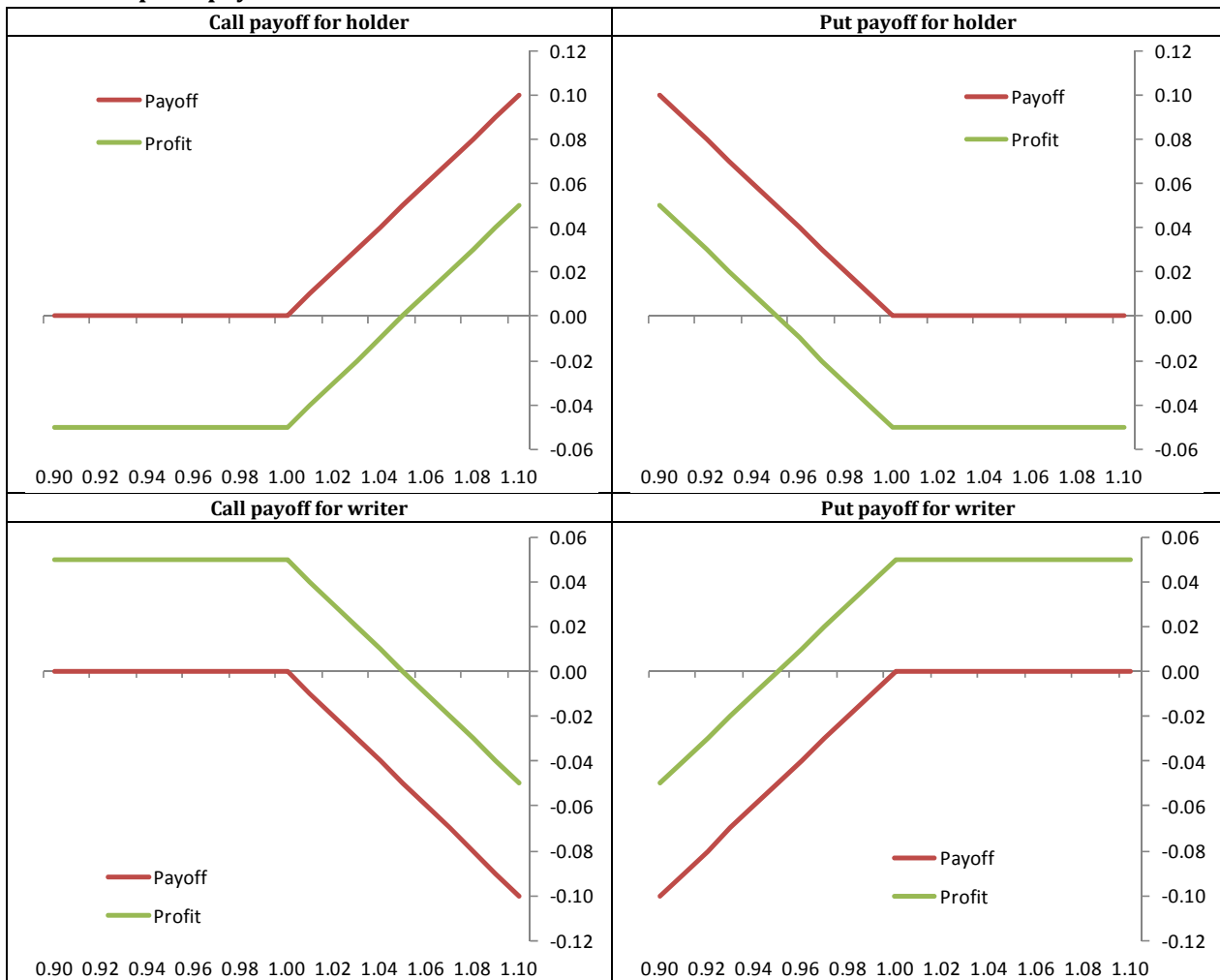
For both calls and puts there is no possibility to the holder of the option of a negative payoff. However, there is nearly always the *possibility* of a positive payoff. Even if the option is significantly out-the-money the price of the underlying can change during the time to expiry (only if the variance of the underlying is exactly zero would this not be possible). Hence if an options contract did not have some value it would lead to arbitrage opportunities.[8] Unlike standard cash flows the value of the option increases with time to maturity, hence *ceterus paribus* an option with a longer maturity will be worth more than one with a shorter maturity. The logic behind this result is that there is a greater chance the option will move further into the money because there is more time to yet elapse.

2.   Option pricing

Black-Scholes pricing formula

As discussed above options (unlike, say, futures) have value ahead of expiry as there is always the possibility of a positive payoff and no possibility of a negative one. Hence to avoid riskless profits an option must always have value and thus carry a positive price. A number of methods have been proposed for obtaining such a price, the most famous and widely used is the one proposed by Black and Scholes (1973). They proposed that the fair price of a European call option written on an underlying equity should be

---

[8] People could acquire costless out-the-money options in the hope that the price of the underlying asset in one or more cases could move such that the option would give a positive payoff.

$$C_{BS}(S, \tau, X, r, q, \sigma) = Se^{-q\tau}N(d_1) - Xe^{-r\tau}N(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{S}{X}\right) + (r - q - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}$$

and

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

with $N(\bullet)$ the cumulative probability from the standard normal distribution.

This classic equation is composed of six parameters: the spot price of the underlying (S), the time to maturity of the option ($\tau = T - t$), the embedded strike price (X), the risk-free interest rate (r), the dividend yield (q) and the volatility of the underlying asset (σ).

The crucial underlying assumptions for the Black-Scholes pricing formula to hold are that the price of the underlying asset follows a geometric Brownian motion and that markets are complete and free of imperfections[9]. The original approach used by Black and Scholes (1973) was based upon pricing the option using combinations of the underlying security and a risk-free bond to mimic an option's payoff structure. A derivation using a probabilistic approach, as presented by amongst others Baz and Chacko (2009), is outlined in **Appendix 1**, while Cochrane (2005) shows that the Black-Scholes formula can also be derived by solving for the appropriate stochastic discount factor.

An implication of the pricing formula is that we need only concern ourselves with either call or put options, as the price of a put option is related to the price of a call option through the put-call parity relationship. Put-Call parity relates the prices of call and put options, and arises from the possibility of arbitrage should the two not be connected, i.e. the possibility of making a profit from writing a call and purchasing a put at the same price or vice versa. It requires the price of a put option to be:

$$P_{BS} = C_{BS} - Se^{-qt} + Xe^{-rt}$$

This implies

$$P_{BS}(S, \tau, X, r, q, \sigma) = Xe^{-r\tau}N(-d_2) - Se^{-q\tau}N(-d_1)$$

where the properties of a standard Normal imply that $N(-d) = 1 - N(d)$.

Crucially for what follows the Black-Scholes pricing formulae assume risk-neutral probabilities[10]. This should be apparent as the risk preferences of agents do not enter into the pricing equations. The implications of this for the interpretation of the derived density functions will be discussed below.

---

[9] That is there are no taxes, short-selling bans, transactions costs *etc*. In addition trading is continuous and investors have access to unlimited borrowing and lending opportunities.

[10] Under risk-neutral pricing, true subjective probabilities of an event happening are replaced by *risk-neutral* probabilities. The risk-neutral probability of any state of the world can be thought of the product of the true probability of that state occurring and the utility derived by the agent in that state. Hence risk-neutral probabilities correct for bearing risk. Risk-neutral can be a misleading terminology, as crucially risk-neutral probabilities place higher event-probabilities on states with higher marginal utility or disutility. Cochrane (2005) uses the example of people who report high probabilities of plane crashes, suggesting that instead of being irrational they are in fact reporting their risk-neutral probabilities, it is just that the anticipated disutility from dying in a plane crash is extremely high for the individual concerned.

Pricing currency options

The Black-Scholes pricing formula was initially designed to price options written on equities; it can easily be adapted to price options on other underlying assets. The main alteration needed to the formula is related to the inclusion of the dividend yield. Often this is a matter of interpretation. The role of the dividend yield in the Black-Scholes formula can be related to the means of its derivation. The payoff of an option can be approximated by the payoff of a replicating portfolio of risk-free bonds and the underlying asset. The payoff to the risk-free bond is the risk-free rate and in this case the payoffs to the underlying asset, an equity or equity index, are the dividends.

To price options on other forms of assets the dividend should be replaced by the payoff from holding such assets. In the case of foreign exchange, instead of the underlying being an equity, the underlying is a foreign currency and the payoff to that foreign currency will be the foreign interest rate.[11]

Thus the formula for the fair price of a European call option written on foreign exchange, was proposed by Garman and Kohlhagen (1983) as

$$C_{BS}(S, \tau, X, r, r^*, \sigma) = Se^{-r^*\tau}N(d_1) - Xe^{-r\tau}N(d_2)$$

where

$$d_1 = \frac{\log\left(\frac{S}{X}\right) + (r - r^* - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}$$

and

$$d_2 = d_1 - \sigma\sqrt{\tau}.$$

Malz (1997) proposed a further simplification to the Garman-Kohlhagen pricing formula applicable to currency options by making an additional substitution so

$$C_{BS}(Q_t, t, r^*\sigma) = N\left(-\frac{\log(Q_t) - \tau\left(\frac{\sigma^2}{2}\right)}{\sigma\sqrt{\tau}}\right) - Q_t N\left(-\frac{\log(Q_t) + \tau\left(\frac{\sigma^2}{2}\right)}{\sigma\sqrt{\tau}}\right)$$

where

$$Q_t = \frac{X_T}{F_{t,T}}$$

and noting that the forward and sport rate are related by

$$F_{t,T} = S_t e^{(r-r^*)\tau}.$$

Here $F_{t,T}$ is the forward rate at time t for settlement at time T in the future. The forward rate can be seen as the current spot price adjusted by the cost of carry over the period to the future date in question. In the case of currencies the cost of carry is defined as the difference between the domestic and foreign interest rates.

---

[11] It should be noted that for some underlying assets the payoff can be negative, for example options written on agricultural produce could include a negative payoff should the costs of storing the underlying through the life of the contract outweigh perceived convenience yields from owning the product.

Implied volatility

When it comes to observing the six parameters that enter into the call price formulas above, it is nearly always the case that while spot price, time to maturity, strike price, risk-free interest rate and dividend yield (or equivalent), are at least approximately observable, perceived volatility is not. What are observable are the market prices for options, the left hand side of the pricing equation. Thus it is possible to back out a value for volatility by finding the value that equates the price suggested by the pricing formula and the observed market price. In most cases, including the examples discussed above, it is impossible to invert the pricing formula to find a closed-form solution; instead numerical methods must be used to iteratively find a solution.

The value of volatility that equates the market price with the pricing formula price is referred to as the implied volatility. It is important to understand what this value means. **The implied volatility is a measure of uncertainty of the asset price at the expiry date.**

One of the most notable features of implied volatilities is the variation in such values for options written on the same underlying asset across strike prices at the same maturities (and across maturities). Such observations are commonly referred to as the volatility smile (or surface).

The volatility smile was first observed when traders plotted implied volatilities for a range of foreign exchange options written on the same underlying currency pair with the same maturity for a range of strike prices. It was apparent that implied volatilities for options at or near the money were relatively low, but rose as the strike price moved further into or out-the-money, and drawing a line through these points created a smile-like shape. Such a smile is often observed in currency options (**Table 2**). The rationale for the appearance of the smile in currency options is that the Black-Scholes pricing formula assumes that price changes are distributed log-normally while in reality exchange rate movements (like other asset price movements) are often distributed differently. If returns were distributed log-normally then there would be no volatility smile, with implied volatility constant across strike prices. The volatility smile for currency options is presented in 'delta' space. 'delta', defined as the first derivative of the option price with respect to the price of the underlying, is an important concept in currency options and is explained below.

Table 2: Common volatility 'smile' for currency options.
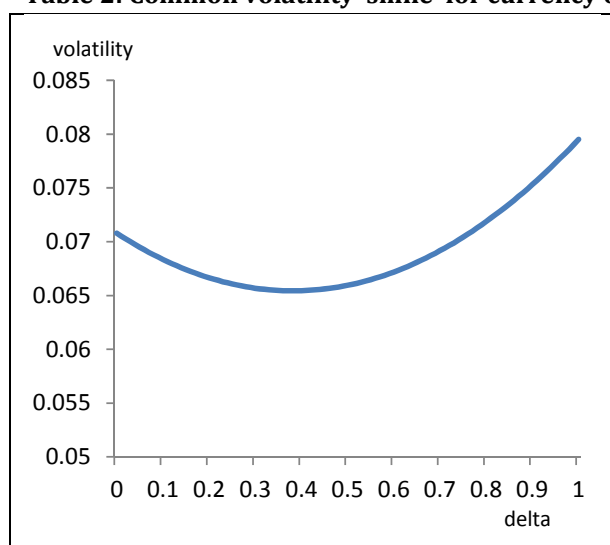
Source: Bloomberg

The smile shape observed for currency options is also seen in options written on other assets such as commodities. One notable exception is for options written on equities: for equity options the shape of the volatility is often described as a smirk where implied volatility decreases as the strike price increases.

This makes a lot of sense. One reason for the difference in shape is that in currency markets the demand for hedging options is relatively evenly split between those protecting against appreciation and deprecation, because different groups of economic agents (i.e. exporters and importers) will suffer in either cases. Similar cases can be made for options on other underlying assets, i.e. for agricultural commodities the market is balanced between producers and consumers. For equity options the majority of economic agents usually require hedging against a fall in equity prices. Therefore demand will be less evenly distributed between in and out-the-money options.[12]

Comparing the volatility smile across different strike prices and different maturities for the same underlying asset is referred to as the *volatility surface*. Hull (2006) notes that typically over time volatility will be mean-reverting. As such it will be expected to rise if near-term volatilities are low relative to historical experience and be expected to fall if near-term volatilities are historically high. In addition the effect of the smile or smirk becomes less pronounced as maturity lengthens.

3.   Obtaining risk-neutral probability density functions

Making use of Breeden and Litzenberger (1978)
Taylor (2005) notes that the following three problems are essentially equivalent:

a) Specify a risk-neutral density ($f_Q(S)$) for all possible prices (S).
b) Specify call prices ($C(X)$) for all possible exercise prices (X).
c) Specify implied volatilities ($\sigma_{implied}(X)$) for all possible exercise prices (X).

If we are able to obtain b) or c) then we are able to obtain a). To understand the logic behind this statement, we need to see that information regarding both the call price and implied volatility across sequential incremental changes in the exercise price reflects the location and shape of the underlying risk-neutral density across the same exercise prices. We know that b) and c) are equivalent from our previous discussion. In other words, to be able to specify call prices at all potential strike prices we need to have a value for implied volatility at all potential strike prices. How then is it that we know f_Q?

By specifying call prices for options written on every available strike price we can determine the location and shape of the risk-neutral density. When comparing the price of two sequential equally-spaced deeply out-of-the-money call options, both the price and the difference in price between the two options will tend to zero as the strike prices move further out-the-money. This happens because the possibility of the option returning to the money tends further towards zero. As both options are perceived to have very little chance of being exercised, the market does not make much distinction in the price of two long shots. In this instance the second difference between prices of options also tends to zero.

---

[12] Hull (2006) also notes that the volatility smirk was not detectable in equity options before the stock market crash of 1987, suggest a 'crashaphobia' in investors.

At the opposite end of the scale, the prices of two sequential equally spaced options deeply in-the-money options will tend towards the profit that can be made from exercising the option (i.e. $S_t - X$) and the difference in prices will tend towards the incremental difference in price between the two. This happens because the possibility of the option *not* being exercised decreases as the option becomes further and further in-the-money. Any difference in price between two options deep in-the-money reflects *only* the difference in profit that will be realised on expiry. Here too the second difference between the option prices tends to zero.

At both of these extremes the difference in the probability that the sequential options will be exercised is very low; deep out-of-the-money both will expire unexercised, deep in-the-money both will be exercised. That is, the underlying will almost certainly be above the strikes of the in-the-money calls but below the strike prices on the out-of-the-money call. In between these two extremes (at or around the money) both the *price* of the two sequential equally spaced options and the *difference in the prices* of these options differs at each interval. This reflects the differing probabilities of whether the underlying will end at the specified level and whether the option will expire in-the-money and hence be exercised. This area represents the central part of the distribution. In fact the mean of the underlying risk-neutral density can also be ascertained simply from specifying all possible call prices. This is the point where the incremental difference between the prices of two equally spaced sequential options will start to decline compared with two equally spaced options closer to the money. The rates at which the extremes of the call prices tend towards their limits also contains information about the standard deviation of the risk-neutral density, while the symmetry above and below the mean of the distribution gives information as to the skew. Finally, kurtosis can also be determined by how quickly the tails of the distribution tend to their extremes. Hence from specifying call prices for all possible strike prices we can ascertain information regarding the underlying risk-neutral density.

However, as discussed above, to be able to specify a call price for every potential strike we need a measure of volatility; else we can't use Black-Scholes. So specifying an implied volatility for every possible strike price means we are fundamentally using the same information. By specifying implied volatilities across a range of strike prices we are ascertaining the volatility smile. The shape and location of the volatility smile simply reflects the shape and location of the underlying risk-neutral density.

If returns on an asset are log-normally distributed (as the Black-Scholes pricing formula assumes) then implied volatility would be constant across all strike prices and there would be no smile effect in the implied volatility. If this was the case the risk-neutral density would also be a log-normal. It is the very presence and shape of a volatility smile that gives us information regarding the shape of the equivalent risk-neutral density. The skew of the smile will also reflect the skew of the risk-neutral density, while how pronounced the smile is will reflect in the kurtosis of the distribution. For currency options, where recall there is often a detectable smile with implied volatility for at or around the money options lower than for both in- and out-of-the money options, then the distribution will likely be symmetric. More pronounced smiles will imply a more leptokurtic distribution. For equity options or any with a pronounced smirk (where implied volatilities fall as the strike price increases) this gives a downward skew to the distribution.

Breeden and Litzenberger (1978) formally obtained the results reflecting the intuition discussed above, i.e. deriving the underlying risk-neutral density from the volatility smile and call price schedule. They show the exact risk-neutral density implied from options priced using the Black-Scholes formula. As discussed above, under risk neutrality the Black-Scholes formula implies that the fair price of a call option is ultimately the discounted expectation of the options payoff.

This can be obtained from the following. The call price function, as with any asset price, is ultimately just the discounted value of the expected return, which is

$$C_{BS}(X_T) = e^{-r\tau}E^Q[(S_T - X_T)].$$

This can be written

$$C_{BS}(X_T) = e^{-rT}\int_0^\infty \max(x - X_T, 0)\, f_Q(x)dx$$

or

$$C_{BS}(X_T) = e^{-rT}\int_X^\infty (x - X_T)f_Q(x)dx$$

where f_Q is the risk-neutral density. In other words the fair value for the call price is the weighted sum of all possible outcomes multiplied by their probability of occurring.

From this, it must be the case that if we differentiate the call price function with respect to the strike price we obtain

$$\frac{\partial C_{BS}(X_T)}{\partial X_T} = e^{-r\tau}\int_X^\infty f_Q(x)dx.$$

Differentiating a second time gives

$$\frac{\partial^2 C_{BS}(X_T)}{\partial X_T^2} = e^{-r\tau}f_Q(X_T)$$

which implies

$$f_Q(X_T) = e^{r\tau}\frac{\partial^2 C_{BS}(X_T)}{\partial X_T^2}.$$

Thus the risk-neutral density can be obtained from the discounted second derivative of the call price function. In general, the second derivative of the call price function simply needs to be scaled to integrate to unity, as it is a density function.

For the Black-Scholes call price formula there is an analytic solution, which is

$$f_Q(X_T) = n(d_2)\left\{\frac{1}{\sigma X_T\sqrt{\tau}} + \frac{2d_1}{\sigma}\frac{\partial\sigma}{\partial X_T} + \frac{d_1 d_2 X_T\sqrt{\tau}}{\sigma}\left(\frac{\partial\sigma}{\partial X_T}\right)^2 + X_T\sqrt{\tau}\frac{\partial^2\sigma}{\partial X_T^2}\right\}$$

where d_1 and d_2 are as defined before in the Black-Scholes pricing formula. The result can be applied straightforwardly to the Garman-Kohlhagen variation. Whilst closed form-solutions can be used, often just taking the second difference between call prices is used as a form of numerical differentiation.

4.   Currency options

The standard means of presentation for options contracts written on the majority of underlying asset classes has put and call contracts offered at a range of strike prices for a small number of fixed maturity dates. In these instances the option contract is nearly always quoted in currency units.

The presentation of options written on foreign exchange differs significantly. In FX markets it is common for options to be priced in terms of implied volatility instead of currency units. In addition delta, defined as the rate of change in the value of the option with respect to changes in the price of the underlying, becomes an important trading measure. Because of this, delta is an attractive way to price currency options, as for a given volatility and exercise price it gives a measure that depends on the asset price alone. This means it provides a pricing schedule for changes in the underlying asset price holding exercise price and volatility constant. In particular this means that the strike prices of in- and out-of-the-money options remain a fixed distance away from the at-the-money-volatility even as other elements vary.

Under the Black-Scholes assumptions, delta is an approximate measure of the 'moneyness' of an option. An at-the-money option will have a delta approximately equal to 0.5, representing a 50:50 chance that it will expire in-the-money. Exercise prices for FX options too are often expressed in terms of delta, so an at-the-money option would be referred to as 50-delta call or put.[13] In addition, round numbers in percentage points such as 25-delta and 75-delta, representing close in-the-money and out-of-the-money calls respectively (or close out-of- and in-the-money puts respectively) are particularly common.

Formally delta is defined as

$$\delta = \frac{\partial C_{BS}(S_t)}{\partial S_t}$$

which is obviously the first derivative of the call price function. Thus delta is

$$\delta = \frac{\partial (Se^{-r^*\tau}N(d_1) - Xe^{-r\tau}N(d_2))}{\partial S_t}$$

which is

$$\delta = e^{-r^*\tau}N(d_1)$$

where, as before,

$$d_1 = \frac{\log\left(\frac{S}{X}\right) + (r - r^* - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}.$$

Therefore we can formally express delta as a function of a range of known parameters and one unknown parameter, implied volatility (σ). Thus for given values of the other parameters, delta maps to a particular value of implied volatility. We also know that delta can at most vary between zero and unity.

---

[13] Malz notes that 'put-call parity implies that puts and calls with the same exercise price have identical implied volatilities, so the volatility of an *x*-delta put equals that of a (1-*x*)-delta call'.

It turns out to be useful to define delta in terms of the modified Garman-Kohlhagen formula. Malz (1997) suggests using:

$$\delta = e^{-r^*\tau}\left(-\frac{\log(Q_t) - (\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right)$$

A further wrinkle to the presentation of currency options is it often common to see quoted the price of a combination of options as opposed to a single option, for example the presentation of currency options on Bloomberg is discussed in **Appendix 2**. In particular it is common to see three combinations of options quoted these are *straddles*, *strangles* and *risk reversals* instead of individual calls and puts.

A *long straddle* involves purchasing both an at-the-money call and an at-the-money put with the *same* expiry date. The profit to the holder of a straddle is shown in **Table 3**, which may be negative taking into account the premium. The holder receives a positive payoff from the strategy as long as the price of the underlying moves away from the strike price; it does not matter if the price rises or falls. A straddle is therefore a hedge against volatility; the greater the potential move away from the strike price the greater the profit to the holder of a long straddle. In the case of currency options the price of a long straddle represents the at-the-money volatility (or the price of a 50-delta call or put).

A *long strangle* is similar to a straddle in that the holder profits from volatility in the price of the underlying. In a long straddle the holder purchases both a call and a put with the same expiry, however in a straddle both purchased options are initially 'out-the-money'. The profits from a long strangle are shown in **Table 2**. A greater degree of price movement is now required for the strategy to be profitable, but the initial cost of the component options is lower (they are less likely to be individually exercised). In the case of currency options the price of a strangle will be quoted as half of the volatility of *X*-delta call plus a (1-*X*)-delta call (or *X*-delta put), minus the volatility of an at-the-money, 50-delta call.[14]

**Table 3. Portfolio of options profits**
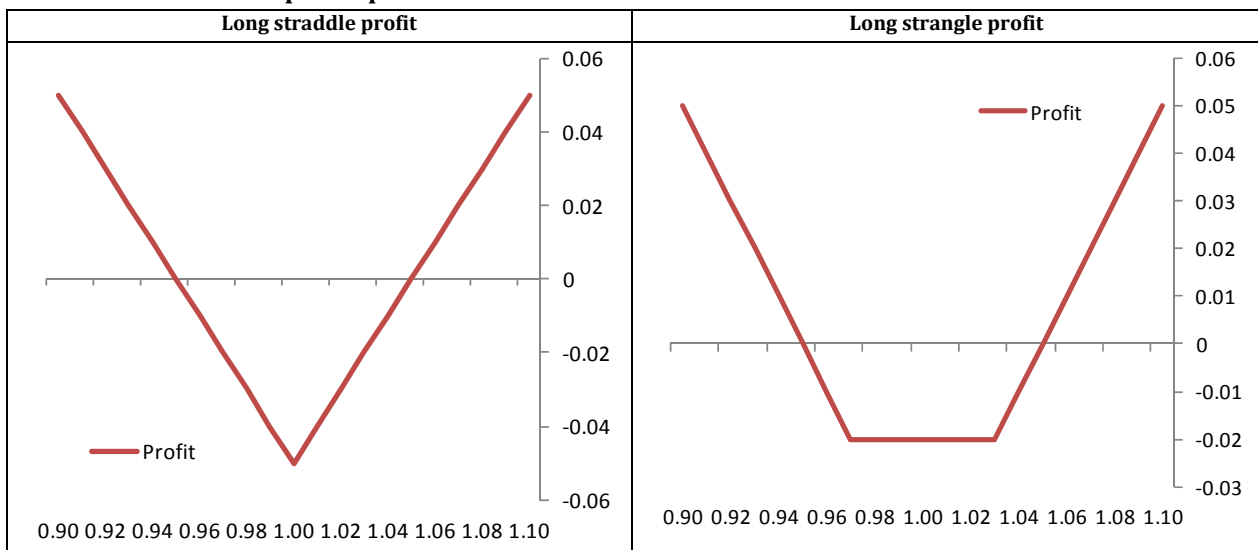
Finally, a risk reversal involves an investor selling an out-of-the-money put option and buying an out-of-the-money call option. The profits from a risk reversal are shown in **Table 4**. Unlike the two

---

[14] It is common to see strangles quoted with *X* = 10, 25 and 35.

previous strategies (whose payoff does not depend on whether the price of the underlying goes up or down) a positive payoff on a long risk reversal is dependent on the price of the underlying increasing. In the case of currency options the price of a *X*-delta risk reversal would be quoted as the difference between the volatility of a *X*-delta call and a (1-*X*)-delta call (equivalent to the *X*-delta put).[15]

Table 4. Long risk reversal profits
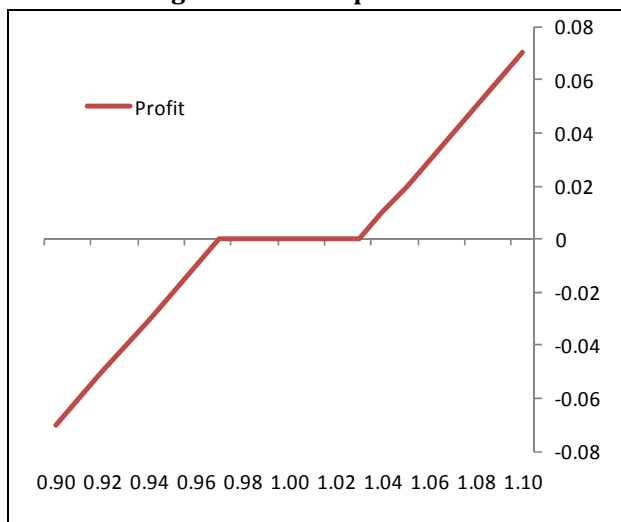
5. Parametric implied probability densities from currency options

As discussed above to ascertain a risk-neutral implied probability density from option prices we require the discounted second derivative of the call price function. To produce a call price function at all potential strike prices we require an estimate for implied volatility at all potential strike prices, Bahra (1997) discusses a number of such approaches. A common method for approaching this problem is to assume that implied volatility follows some underlying functional form and that the implied volatilities that we can calculate from actual published options prices are merely observations of this function. Therefore a first step in ascertaining a density is to estimate a functional form for the volatility smile.

One such approach for currency options was proposed by Malz (1997). Malz suggested the following form for the volatility smile

$$\sigma(\delta) = b_0 atm_t + b_1 rr_t(\delta - 0.5) + b_2 str_t(\delta - 0.5)^2.$$

This form utilises the combination of options discussed above. The at-the-money volatility gives the location of the smile, the risk reversal gives the skew, and the strangle kurtosis. In Malz's approach he utilises 25-delta risk reversals and strangles. The assumed functional form is a quadratic function which allows for the smile shape often observed in the volatility smile for currencies. Malz showed that for the parameterisation of the implied volatility function to hold, the b_i values must be[16]

$$\sigma(\delta) = atm_t + 2rr_t(\delta - 0.5) + 16str_t(\delta - 0.5)^2$$

---

[15] Again it is common to see risk reversals quoted with *X* = 10, 25 and 35.
[16] To do this he assumed that the observed values of these contracts must satisfy this functional form. For example in the case where observed volatility is at-the-money so delta = 0.5, the values of the risk reversal and strangle terms will be zero, therefore for the equation to hold the coefficient on the at-the-money volatility must be 1.

What this gives is an equation for implied volatility in terms of delta; however, we initially don't know either implied volatility or delta. However, from the definition of delta noted above we have an equation for delta in terms of implied volatility. This provides us with a pair of simultaneous equations, with the two unknowns implied volatility and delta:

$$\sigma(\delta) = atm_t + 2rr_t(\delta - 0.5) + 16str_t(\delta - 0.5)^2$$

$$\delta = e^{-r^*\tau}\left(-\frac{\log(Q_t) - (\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right)$$

Solving these simultaneous equations at each potential strike price (defined in currency units) will provide values for implied volatility and delta at those strike prices. This is easily done in MATLAB and the relevant code can be found in the **Application** below.

Once values for implied volatilities are determined for each strike price these can be put into the modified call price function for FX options, determining the price at every potential strike. Subsequently, numerical methods need to be used to determine the second derivative of the call price function and hence the implied probability density.

Application: Drawing parametric densities for currency options in MATLAB

The technique for drawing densities using the parametric approach to specifying the volatility smile in MATLAB is relatively straightforward. It involves setting up a loop to recursively solve for sigma and delta at each potential strike price, followed by the calculation and twice differentiation of the call price function to obtain the density function.

There are a number of techniques that can be used to solve simultaneous equations in MATLAB©, one of the simplest is to use the built-in `fsolve` function. However, using `fsolve` can be a slow process so a simple iteration which achieves the same result much more quickly can be coded as follows:

```
function [delta, sigma] = SolveV(delta, sigma, atm, rr, str, rstar, t, q)

lq   = log(q);
ert  = exp(-rstar*t);
st   = sqrt(t);
Vold = delta+sigma+1;
while abs(delta+sigma-Vold)>0.0000001
    Vold  = delta+sigma;
    sigma = atm-2*rr*(delta-0.5)+16*str*(delta-0.5)^2;
    d     = -(lq-sigma*sigma*t/2)/(sigma*st);
    delta = ert*0.5*erfc(-d/sqrt(2));
end
```

One notable feature of this code is the use of the complementary error function (`erfc`) instead of the cumulative normal density function (`normcdf`) to evaluate the cumulative normal of the calculated value of d, referred to in the text as $N(d)$. The relationship between the two is a simple rescaling, with

```
normcdf(d) = 0.5*erfc(-dx/sqrt(2))
```

We use this for two reasons, one the minor improvement that it is usually quicker, but more importantly that this then avoids the need to use the statistics toolbox which some all users of MATLAB may not have as it is only available at extra cost.

Using this solve function, which can be saved somewhere as a routine to call upon, it is straightforward to set up code to solve the desired equations. Given the constant horizon of the options contracts provided on currencies it is possible to not only to draw a density function for a single day but a series of density functions across time. To illustrate this code we use data for the UK pound/US dollar exchange rate and related 3-month constant maturity options for the period during November 2014.

Sample code to do this is as follows.

```matlab
% Program for calculating option implied pdfs using 3Month GBP / USD options.
% Data for November 2014
% Follows parametric technique outlined by Malz(1997)

%% Clean up
close all; clear all; clc;


%% Data

D = [1.599  1.600 -0.785  0.220  6.130  0.008  735906;
     1.599  1.601 -0.760  0.215  6.060  0.030  735907;
     1.598  1.600 -0.835  0.220  6.445  0.025  735908;
     1.584  1.586 -0.865  0.238  6.623  0.028  735909;
     1.587  1.589 -0.920  0.230  6.465  0.028  735910;
     1.586  1.587 -0.895  0.225  6.350  0.020  735913;
     1.589  1.590 -0.905  0.230  6.448  0.020  735914;
     1.580  1.581 -0.933  0.230  6.543  0.018  735915;
     1.571  1.572 -0.950  0.235  6.658  0.013  735916;
     1.567  1.569 -1.038  0.238  6.838  0.013  735917;
     1.564  1.566 -1.033  0.233  6.963  0.015  735920;
     1.565  1.567 -1.038  0.235  6.888  0.023  735921;
     1.568  1.570 -1.018  0.238  6.578  0.010  735922;
     1.569  1.570 -1.033  0.235  6.613  0.010  735923;
     1.566  1.567 -1.045  0.235  6.693  0.008  735924;
     1.570  1.571 -1.040  0.230  6.828  0.013  735927;
     1.571  1.572 -1.040  0.235  6.673  0.028  735928;
     1.578  1.580 -1.040  0.230  6.580  0.030  735929;
     1.572  1.574 -1.030  0.228  6.505  0.030  735930;
     1.565  1.566 -1.055  0.238  6.643  0.015  735931];


spot   = timeseries(D(:,1),datestr(D(:,7)));
fwd3m  = timeseries(D(:,2),datestr(D(:,7)));
rr3m   = timeseries(D(:,3),datestr(D(:,7)));
st3m   = timeseries(D(:,4),datestr(D(:,7)));
vol3m  = timeseries(D(:,5),datestr(D(:,7)));
usd3m  = timeseries(D(:,6),datestr(D(:,7)));
gdates = D(:,7);


rstar = usd3m.Data/100;
fwd   = fwd3m.Data;
```

```matlab
atm    = vol3m.Data/100;
rr     = rr3m.Data/100;
str    = st3m.Data/100;

% Define time (i.e. 3 months)
tau    = 3/12;
st     = sqrt(tau);

% Define strike prices
inc = 0.01;
xax = (1.35:inc:1.8)';
xaxis = xax(3:end);

% Set up storage matrices
k      = usd3m.timeinfo.startdate;
dates = datestr(D(:,7));
pdf    = nan(length(xax)-2,length(dates));

%% Main program loop

for j = 1:length(dates)-1;

  % Calculate the strike price relative to the forward rate
  q      = xax/fwd(j);

  % Calculate deltas by solving simultaneous equation (delta is the ratio of
  % change of the price of the option with the price of the underlying asset).
  delta = nan(length(xax),1);
  s      = nan(length(xax),1);

  for k=1:length(xax)
    [delta(k), s(k)] = SolveV(0.5,1,atm(j),rr(j),str(j),rstar(j),tau,q(k));
  end

  % Calculate the call prices
  d1    = -(log(q)-s.^2*tau/2)./(s*st);
  call = 0.5*(erfc(-d1/sqrt(2))-q.*erfc(-(d1-s*st)/sqrt(2)));

  % Calculate change in the price of the call per difference quotient, i.e.
  % create distribution: discrete version of Breeden and Litzenberger (1978)
  pr        = diff(call,2);
  pdf(:,j) = pr/sum(pr);

end

%% Creating graph
figure
surface(gdates,xaxis,pdf)
dateaxis('x',1)
colormap jet
view([90-37.5,30])
```

The resulting chart suggests that there was little change in participants expectations over the period covered. The location of the distribution remains relatively constant over the time horizon. There does appear to be a slight increase in uncertainty with the more recent distributions appearing less peaked than the earlier distributions.
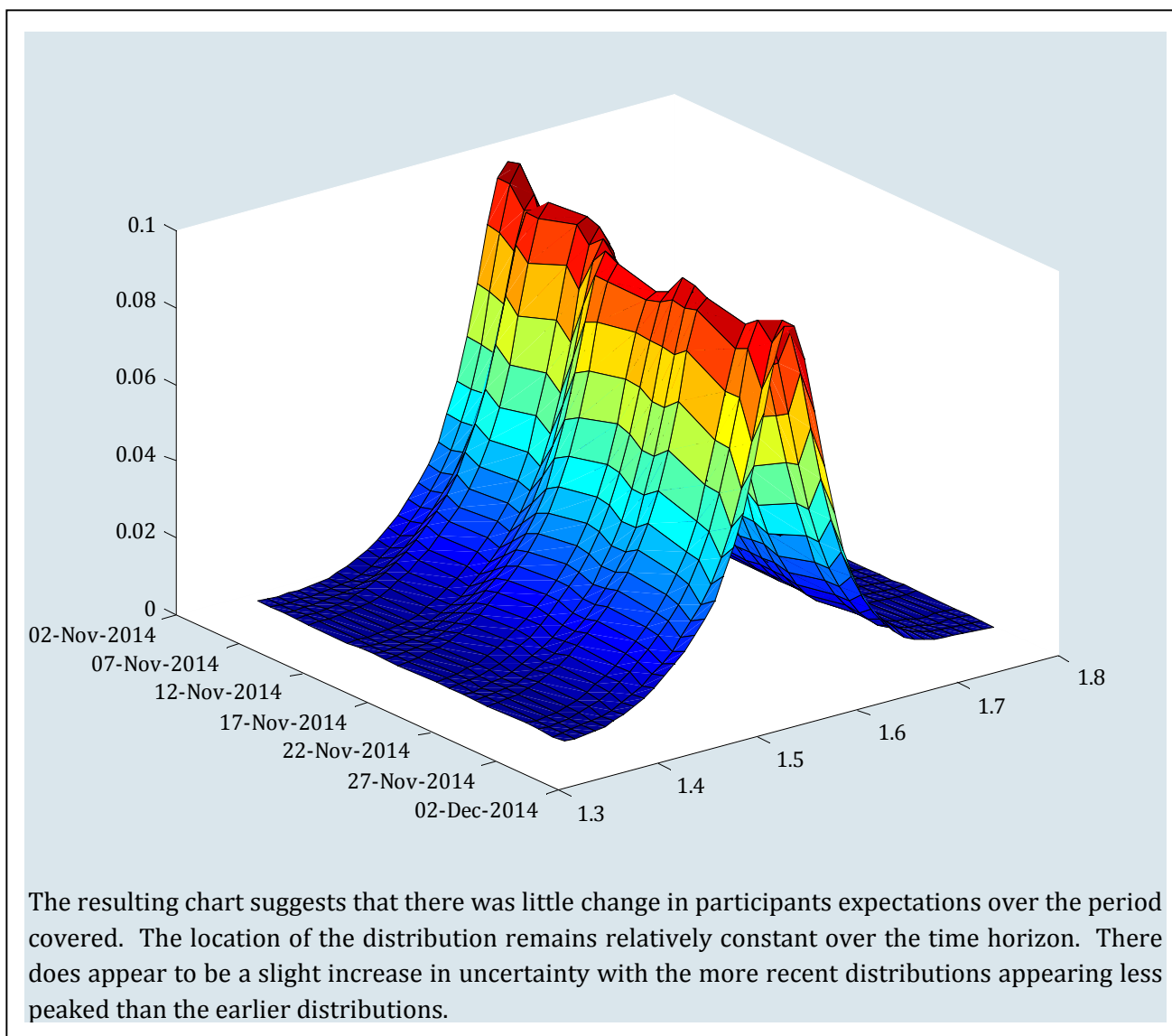
6.   Non-parametric implied probability densities from currency options

The approach detailed above imposes a functional form upon the volatility smile. A more recent approach proposed by Malz (2014) does not require assumptions regarding the shape of the volatility smile. Malz refers to this new approach as 'a simple and reliable' approach as the implementation is almost identical across a range of asset classes such as equities, foreign exchange and interest rates. The approach consists of three simple steps:

1. Interpolate and extrapolate the volatility smile using a clamped cubic spline (the intuition behind clamped splines is discussed in **Appendix 3** together with code) with flat ends.[17]
2. Use the call price function to return the interpolated volatilities into currency-unit options.
3. Make use of the Breeden-Litzenberger result and take the second differences of the call price function to obtain the density.

The application of this technique to currency options is relatively straightforward, but has a few small complications that need to be overcome.

---

[17] Malz utilises this form of cubic spline to ensure concavity in the volatility smile. Concavity is a necessity to ensure that positive values are returned for the second differences and hence all potential strike prices have strictly positive probability.

The first stage is to interpolate and extrapolate the volatility smile. To do this for currency options we require volatilities at various deltas in and out-the-money. As discussed previously, the majority of data providers present currency options as combinations priced in terms of volatility. This means that the options are priced in the space over which we want to draw the volatility smile, but we first need to manipulate the available data a little to find specific points. For most currency pairs Bloomberg provides data for strangles and risk reversals (see **Appendix 2**) at 10, 25 and 35 delta. From these prices we can calculate volatilities at 10, 25, 35, 65, 75 and 90 delta:

For $X = 10$, 25 and 35

$$Xdelta\ volatility = atm + (Xdelta\ str + Xdelta\ rr)/2$$

For $X = 65$, 75 and 90

$$Xdelta\ volatility = atm + (Xdelta\ str - Xdelta\ rr)/2$$

Here we face one of the minor complications. It turns out that the at-the-money volatility has a delta close to but not precisely 50. However the value of delta for the at-the-money volatility can be calculated using the initial formula for delta and setting the strike price equal to the current spot price.

Having done this we now have 7 data points over which to draw the volatility smile using a clamped cubic spline. This provides us with a value for volatility at all potential deltas. The main complexity for currency options is that to populate the call price function we need volatilities at all potential strike prices in currency units. As Malz notes, it is not possible to map directly between volatility/delta space to volatility/strike price space using the expression for delta. This is because the volatility argument in the equation for delta is not constant, but varies with delta itself. It is possible however to use numerical methods to find the value of volatility associated with each potential *strike price*. To do this we take advantage of the fact that once we have the appropriate volatility for each potential strike price, the value of delta calculated at that given strike price from the appropriate value of implied volatility should be identical to the value of delta calculated using the value of implied volatility calculated from the cubic spline function. We can numerically iterate to a value of implied volatility that achieves this. Formally we can express this as follows. Choose values such that

$$\sigma_{s+1} = \sigma_s$$

where

$$\hat{\delta} = e^{-r^*\tau} N\left(\frac{\log\left(\frac{S}{X}\right) + (r - r^* - \sigma_s^2/2)\tau}{\sigma_s \sqrt{\tau}}\right)$$

and

$$\sigma_{s+1} = a_k + b_k\left(\hat{\delta} - \delta_k\right) + c_k\left(\hat{\delta} - \delta_k\right)^2 + d_k\left(\hat{\delta} - \delta_k\right)^3$$

where a_k, b_k, c_k and d_k are the coefficients of the cubic spline evaluated at δ_k the observed volatilities, which we choose as the knot points of the interpolator, and $\hat{\delta}$ is the implied delta. So we are choosing the appropriate value of $\hat{\delta}$ to equate the values of σ. Once this is done we can calculate the option price at the strike price.

This can be repeated across the range of strike prices over which we are interested in drawing the density over. Note care must be taken to ensure that the correct values of the knot points are utilised once the calculated values of implied volatility move between the observed values of implied volatility. Once this is achieved we have values for sigma at all possible strike prices and can hence calculate call prices at all potential strike prices. Once we have call prices at all potential strike price we can take the numerical second difference to ascertain the risk-neutral density. Code for implementing this technique in Matlab is in the next **Application**.

Application – Drawing non-parametric densities for currency options in MATLAB

The technique for drawing densities using the non-parametric approach to specifying the volatility smile in MATLAB is again relatively straightforward. We describe the necessary sub-functions `cubic_spline` and `cubic_spline_eval` to calculate the coefficients of the cubic spline and to evaluate the spline over the desired interval in **Appendix 3**. The interpolation of the smile is an important issue but is not central to the non-parametric density itself. Interested readers should study this separately, including Malz's discussion of the appropriate form of spline for the problem in hand.

The code below calls upon two additional small sub-routines `c_bs` and `d_bs` which calculate the call price of a European call option and the delta of a European call option as follows:[18]

```
% European call option price
% Inputs:   s = underlying asset price
%        t = time to expiry
%        x = strike price
%        r = return on asset
%        q = dividend yield
%        v = volatility

function c = c_bs(s,t,x,r,q,v)
  vt = v*sqrt(t);
  d1 = (log(s./x)+t*((r-q)*0.5*v.^2))./vt;
  c  = 0.5*s*exp(-q*t).*erfc(-d1./sqrt(2)) ...
        - 0.5*x*exp(-r*t).*erfc(-(d1-vt)./sqrt(2));


% Black-Scholes formula for delta
% Inputs: S      asset price
%         tau    time to expiry
%         X      exercise price
%         r      return on asset
%         q      dividend yield
%         sigma volatility

function delta = d_bs(S,tau,X,r,q,sigma)
  v     = (log(S./X)+tau*(r-q+0.5*sigma.^2))./(sigma*sqrt(tau));
  delta = 0.5*exp(-q*tau).*erfc(-v/sqrt(2));
```

---

[18] The code could easily be modified to calculate these values within the main routine; however, calling on sub-functions makes the presented code both easier to follow and more portable to other applications.

Note that as before we use the complementary error function to calculate the cumulated normal density instead of a call to `normcdf`.

```matlab
% Program for calculating option implied pdfs using 3Month GBP/USD options.
% Data for November 2014
% Follows non-parametric technique proposed by Malz (2014)


%% Clean up
close all; clear all; clc;


% Data, entered a vector: [Spot, uk3m, us3m, atm, b10, b25, b35, r10, r25,
% r35, dates]


D = [1.599  0.448  0.008  6.130  0.665  0.220  0.075 -1.455 -0.785 -0.430   735906;
     1.599  0.447  0.030  6.060  0.645  0.215  0.075 -1.415 -0.760 -0.420   735907;
     1.598  0.447  0.025  6.445  0.645  0.220  0.075 -1.525 -0.835 -0.460   735908;
     1.584  0.453  0.028  6.623  0.680  0.238  0.085 -1.570 -0.865 -0.478   735909;
     1.587  0.437  0.028  6.465  0.670  0.230  0.080 -1.655 -0.920 -0.505   735910;
     1.586  0.445  0.020  6.350  0.660  0.225  0.080 -1.658 -0.895 -0.495   735913;
     1.589  0.445  0.020  6.448  0.675  0.230  0.080 -1.650 -0.905 -0.498   735914;
     1.580  0.444  0.018  6.543  0.680  0.230  0.080 -1.700 -0.933 -0.513   735915;
     1.571  0.444  0.013  6.658  0.673  0.235  0.080 -1.730 -0.950 -0.525   735916;
     1.567  0.427  0.013  6.838  0.680  0.238  0.085 -1.860 -1.038 -0.575   735917;
     1.564  0.453  0.015  6.963  0.670  0.233  0.080 -1.865 -1.033 -0.573   735920;
     1.565  0.460  0.023  6.888  0.680  0.235  0.080 -1.885 -1.038 -0.573   735921;
     1.568  0.459  0.010  6.578  0.685  0.238  0.083 -1.855 -1.018 -0.563   735922;
     1.569  0.458  0.010  6.613  0.673  0.235  0.080 -1.865 -1.033 -0.570   735923;
     1.566  0.450  0.008  6.693  0.690  0.235  0.080 -1.883 -1.045 -0.575   735924;
     1.570  0.456  0.013  6.828  0.680  0.230  0.080 -1.910 -1.040 -0.575   735927;
     1.571  0.460  0.028  6.673  0.675  0.235  0.080 -1.890 -1.040 -0.573   735928;
     1.578  0.459  0.030  6.580  0.665  0.230  0.080 -1.885 -1.040 -0.573   735929;
     1.572  0.458  0.030  6.505  0.665  0.228  0.080 -1.888 -1.030 -0.570   735930;
     1.565  0.450  0.015  6.643  0.685  0.238  0.083 -1.900 -1.055 -0.580   735931];


% Data Manipulation
SPOT = timeseries(D(:,1),datestr(D(:,11)));
RUK  = timeseries(D(:,2),datestr(D(:,11)));
RUS  = timeseries(D(:,3),datestr(D(:,11)));
ATM  = timeseries(D(:,4),datestr(D(:,11)));
B10  = timeseries(D(:,5),datestr(D(:,11)));
B25  = timeseries(D(:,6),datestr(D(:,11)));
B35  = timeseries(D(:,7),datestr(D(:,11)));
R10  = timeseries(D(:,8),datestr(D(:,11)));
R25  = timeseries(D(:,9),datestr(D(:,11)));
R35  = timeseries(D(:,10),datestr(D(:,11)));
gdates = D(:,11);


iv10 = (ATM.Data + B10.Data + R10.Data/2)/100;
iv90 = (ATM.Data + B10.Data - R10.Data/2)/100;
iv25 = (ATM.Data + B25.Data + R25.Data/2)/100;
iv75 = (ATM.Data + B25.Data - R25.Data/2)/100;
iv35 = (ATM.Data + B35.Data + R35.Data/2)/100;
iv65 = (ATM.Data + B35.Data - R35.Data/2)/100;


atm = ATM.Data/100;
r   = RUK.Data/100;
```

```matlab
q   = RUS.Data/100;
S   = SPOT.Data;


inc = 0.001;
xax = 1.35:inc:1.8;


tau = 3/12;
all_pdf = [];
all_cdf = [];


for j = 1:length(atm)

  % Set up volatility smile observed points.
  iv   = [iv10(j) iv25(j) iv35(j) atm(j) iv65(j) iv75(j) iv90(j)]';
  atmv = d_bs(S(j),tau,S(j),r(j),q(j),atm(j));
  delta = [.10 .25 .35 atmv .65 .75 .90]';

  % Calculate and draw cubic spline
  [a, b, c, d] = cubic_spline(delta,iv,0);
  sv   = 0*xax';
  iv_n = sv + atm(j);
  delta_n = d_bs(S(j),tau,xax',r(j),q(j),iv_n);

  % Convert to price space
  while max(abs(sv-iv_n)) > 1e-10
    sv = iv_n;
    iv_n  = cubic_spline_eval(delta_n',a,b,c,d,delta,iv);
    delta_n = d_bs(S(j),tau,xax',r(j),q(j),iv_n);
  end
  call = c_bs(S(j),tau,xax',r(j),q(j),iv_n);

  % Calculate pdf and cdf
  df   = (call(2:end)-call(1:end-1));
  pdf  = (df(2:end)-df(1:end-1));
  cdf  = cumsum(pdf)/sum(pdf);
  all_cdf = [all_cdf cdf];
  all_pdf = [all_pdf pdf];

end

%% Draw chart
figure
surface(gdates,xax(3:end-2),all_pdf(3:end,:))
dateaxis('x',1)
colormap jet
view([90-37.5,30])
```
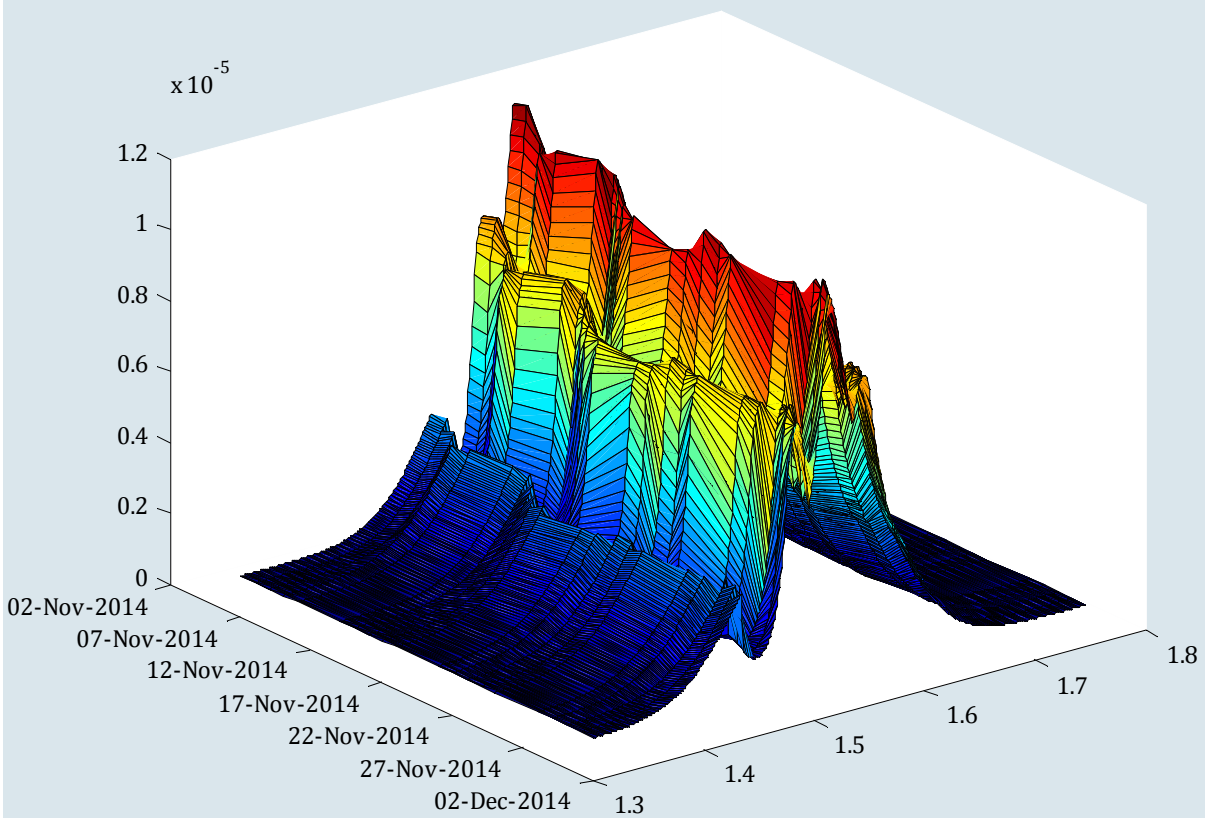
The resulting graph can be compared to the parametric approach we took earlier.

Similar to what we found with the parametric approach, there is little change in the location of the densities over the time period. More interestingly, the non-parametric approach reveals greater density in the left hand tail, suggesting investors were looking to insure against sterling depreciation (dollar appreciation). This is revealed by the ability to use more contracts either side of 50 delta when fitting the PDF.

The main difference between the densities produced by the parametric and non-parametric approach is the smoothness of the density. Those produced by the non-parametric approach will typically be less classically bell curve shaped than those produced by the parametric approach. This is a natural consequence of not imposing a shape on the volatility smile and in addition ensuring that all observed volatilities are hit exactly by the assumed function for volatility. While such an approach may be more accurate and the technique portable across asset classes it may be more sensitive to imperfections in market pricing related to illiquid contracts, i.e. because all observed volatility points are included then stale or illiquid market prices may have undue influence on the shape of the smile.

7.   Interpreting densities

Summary statistics

Having obtained a risk-neutral density the final piece of the puzzle is how to use them. Ultimately the derivation of option-implied probability densities is to shed some light on the future expectations of currencies. But some care is needed in interpreting what such densities are telling us. A significant amount can be taken from the main summary statistics of the densities produced. Each day on its website the Bank of England publishes data on the summary statistics of a number of densities calculated for UK equities and interest rates[19]. The Bank highlights six summary statistics that are of particular use in interpreting the data contained within the implied probability density functions (PDF) which are similarly relevant to the densities for exchange rates:

- *Mean:* the first moment of the implied PDF. It is a measure of central tendency or 'centre of gravity' for the implied PDF. Given the risk-neutral nature of the implied PDF, it is equal to the futures price of the underlying asset.
- *Standard deviation:* the square root of the second moment of the implied PDF. It provides a measure of the dispersion of the implied PDF. It is not annualised and is expressed in the same units as the price of the underlying asset.
- *Median:* the point of the implied distribution that has 50% probability above and below it. It is the 50th percentile and shows the level of the underlying asset that has a cumulative probability of occurring of 50%.
- *Skew:* the third central moment of the implied PDF standardised by the third power of the standard deviation. It provides a measure of asymmetry for the distribution. It measures the relative probabilities (weighted by cubic distances) above and below the mean outcome, that is, the futures price.
- *Kurtosis:* the fourth moment of the PDF divided by the fourth power of the standard deviation. It provides a measure of how peaked the distribution is or, equivalently, the concentration of probability in the upper and lower tails of the implied PDF. A frequently used benchmark for kurtosis is that of a normal distribution which has a kurtosis of 3. It is location invariant and unitless.
- X^{th} Percentile: the point of the distribution for which there is an x % probability for future values of the underlying being at/below this point (i.e. the cumulative probability of this asset price occurring).

By assessing these statistics we can infer a significant amount about market participants' expectations of future asset prices. Obviously such interpretations are subject to caveats regarding risk-neutral probabilities.
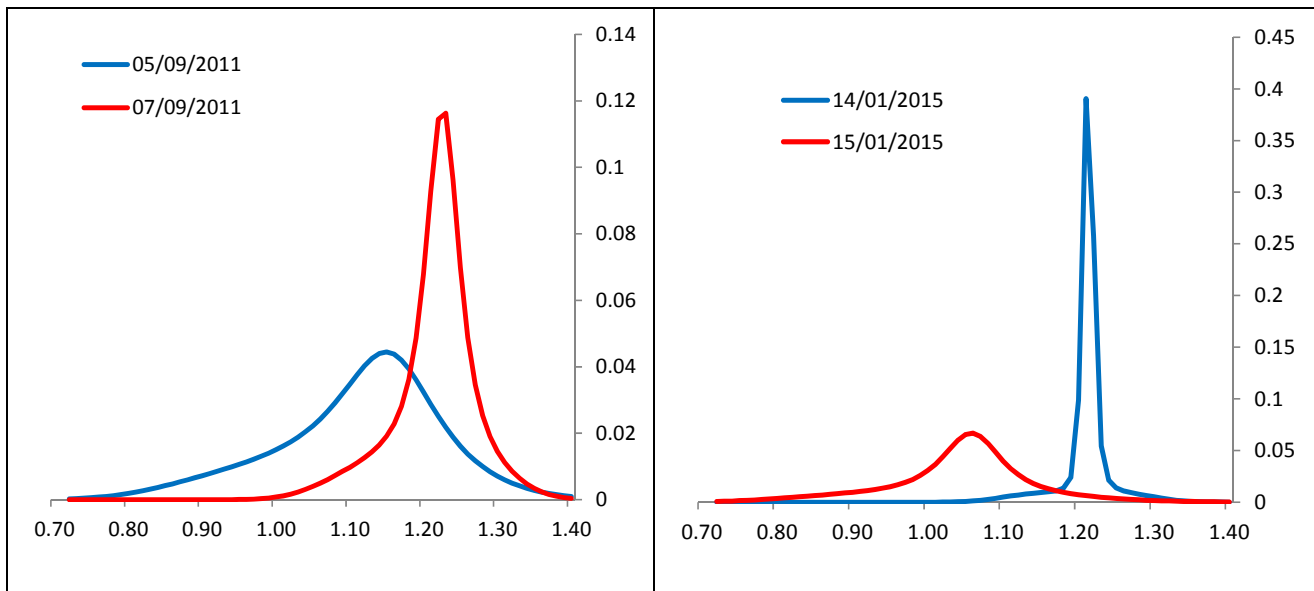
Changes in densities

In addition to analysing the information contained within individual densities, comparing the evolution of densities across dates relatively close together[20] can help to see how expectations of asset prices have changed. This technique can be used to analyse the impact of events or policy announcements. For example the significant impact of the Swiss National Bank's decision both to

---

[19] http://www.bankofengland.co.uk/statistics/Pages/impliedPDFs/default.aspx
[20] Assuming that the densities are adjusted to reflect a constant maturity horizon, which is true in the case of the Bloomberg currency options.

impose and remove a floor on the Swiss franc / euro exchange rate in September 2011 and January 2015 respectively can be clearly seen in expectations taken just before and just after the events (**Table 5**). In both cases both the location and shape of the PDF changes dramatically as uncertainty is removed with the imposition of the floor and increased with its removal.

Table 5. 3-month option-implied parametric probability densities for Swiss franc/euro exchange rates.

Source: Bloomberg

Comparing densities over extended time periods is often more difficult, particularly if the underlying exchange rate is non-stationary. In addition it is relatively difficult to compare PDFs across currencies as the shape and location of the densities is dependent on the values of the underlying exchange rate. A potential way to get around this problem is to compare implied log returns of different assets implied by the densities.[21] Another commonly used means of comparing assets is to compare at-the-money implied volatility which is often used to compare price uncertainty across different assets.

---

[21] Defined as the logarithmic difference between the current futures price and the spot price at the maturity date of the contract.

Appendix 1 – Deriving the Black-Scholes pricing equation

Baz and Chacko (2009) outline a probabilistic approach to deriving Black-Scholes that explicitly uses the assumption that the price of a security today is simply the expected value of the security in the future discounted by the risk-free interest rate. We sketch their proof here.

Begin by noting that a European call option written on an underlying stock S which does not pay a dividend and has expiry at time T and strike price X the price at the current time t can be written as:

$$C_t = \exp(-r\tau)\, E_t^Q[\max(0, S_T - X)]$$

Where $\tau = T - t$ or the time to expiry and E_t^Q is the risk-neutral expectations taken at time t. This can be trivially expressed in terms of log prices for the underlying asset:

$$C_t = \exp(-r\tau)\, E_t^Q[\max(0, \exp(\log(S_T))) - X)]$$

To put a value to this expectation we need some assumption about the future evolution of the underlying asset price. The most common one is that asset prices follow a geometric Brownian motion, which under risk-neutral pricing can be defined as:

$$\frac{dS_t}{S_t} = r\,d_t + \sigma W_t^Q$$

where W_t^Q is a risk-neutral Weiner process. Using Ito's Lemma and assuming a log-normal distribution we are able to say that the conditional density of the underlying at expiry given the current price of the underlying is as follows:

$$\log S_T \mid \log S_t \sim N\left[\log S_t + \left(r - \frac{\sigma^2}{2}\right)\tau\,,\,\sigma^2\tau\right]$$

Because of the Brownian motion assumption it depends on the initial value and the time interval, τ. Hence we express the conditional expectation of the future asset price in terms of an integral of the conditional density of that asset price:

$$C_t = \exp(-r\tau)\int_{\log X}^{\infty}[\exp(S_T) - X]\frac{1}{\sigma\sqrt{2\pi\tau}}e\left[-\frac{1}{2}\left(\frac{\log S_T - \log S_t - \left(r - \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)^2\right]d\log S_T$$

This can be simplified such that we are left with two integrals in the form of distribution functions:

$$C_t = \frac{1}{\sigma\sqrt{2\pi\tau}}\exp(-r\tau)\int_{\log X}^{\infty}\exp\left[\log S_T - \frac{1}{2}\left(\frac{\log S_T - \log S_t - (r - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right)^2\right]d\log S_T$$

$$-X\frac{1}{\sigma\sqrt{2\pi\tau}}\exp(-r\tau)\int_{\log X}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{\log S_T - \log S_t - \left(r - \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)^2\right]d\log S_T$$

These integrals can be simplified in turn. For the first term if we denote $Z = \log(S_T/S_t)$ then we can rewrite the first half of the expression as:

$$\frac{1}{\sigma\sqrt{2\pi\tau}}\exp(-r\tau)\,S_T\int_{\log\frac{X}{S_t}}^{\infty}\exp\left[Z-\frac{1}{2}\left(\frac{Z-(r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right)^2\right]dZ$$

After expanding and collecting the terms in the exponential we can further write the term as:

$$\frac{1}{\sigma\sqrt{2\pi\tau}}\exp(-r\tau)\,S_T\int_{\log\frac{X}{S_t}}^{\infty}\exp\left[-\frac{1}{2}\left(\frac{Z^2-2Z\left(r+\frac{\sigma^2}{2}\right)\tau+\left(r-\frac{\sigma^2}{2}\right)^2\tau^2}{\sigma^2\tau}\right)\right]dZ$$

By completing the square in the numerator of the exponential by adding and subtracting $2r\sigma^2\tau^2$ and then rearranging we can rewrite as:

$$\frac{1}{\sigma\sqrt{2\pi\tau}}\exp(-r\tau)\,S_T\int_{\log\frac{X}{S_t}}^{\infty}e\left[-\frac{1}{2}\left(\frac{Z-\left(r+\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)^2\right]dZ$$

Making a second transformation such that:

$$Y=\frac{Z-(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$$

We can simplify the expression such that:

$$S_T\int_{\frac{\log\frac{X}{S_t}-(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{Y^2}{2}\right]dY$$

Here the integral is simply the standard normal evaluated at $-\frac{\log\frac{X}{S_t}-(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$.

Hence we can finally write the first integral as:

$$S_t N\left(\frac{\log\frac{S_t}{X}+\left(r+\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)$$

where, as before, $N(x)$ represents the standard normal distribution evaluated at x and we use that $1-N(x)=N(-x)$.

Using similar techniques the second integral can be rewritten such that:

$$C_t=S_t N\left(\frac{\log\frac{S_t}{X}+\left(r+\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)-Xe^{-r\tau}N\left(\frac{\log\frac{S_t}{X}+\left(r-\frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}\right)$$

which is the familiar Black-Scholes equation for the price of a European call option.

Appendix 2 – Bloomberg currency options

Currency options data are available on Bloomberg for a wide array of currency pairs and constant maturities. Bloomberg provides the three contracts required to implement both the parametric and non-parametric approaches proposed by Malz: strangles, straddles and risk reversals. Some confusion can arise as the contracts on Bloomberg are referred to by different names to those in the original Malz paper. For the straddle the value we require is simply the at-the-money implied volatility. The risk reversal is still called that, however, the equivalent contract for strangles is referred to as a butterfly.

The naming convention is consistent across currency pairs with the first six letters of the code referring to the relevant currency pair (i.e. UK sterling to US dollar is GBPUSD). The following letters refer to the form of contract and the delta at which it is priced. For the at-the-money volatility a V is required, for butterflies B and risk reversals R. For the butterfly and risk reversal a value for delta is also required. For most currency pairs 10, 25 and 35 delta contracts are available. The final two letters of the mnemonic refer to the maturity where a wide range of constant maturities are available (i.e. three-month would be 3M). Therefore is one was looking to implement the parametric approach to ascertain a probability density function for the one-month Swiss franc to euro exchange rate then the following codes would be required:

Straddle - CHFEURV1M Currency.

Strangle – CHFEUR25B1M Currency.

Risk Reversal – CHFEUR25R1M Currency.

For the non-parametric approach additional risk reversals and strangles (butterflies) can be used. The 10 and 35 delta butterfly and risk reversal contracts can be found using the following codes:

Strangles – CHFEUR10B1M Currency and CHFEUR35B1M Currency.

Risk Reversals – CHFEUR10R1M Currency and CHFEUR35R1M Currency.

Appendix 3 – Natural and clamped splines

The implementation of the non-parametric approach to deriving option-implied probability densities requires us to interpolate between observed values for the volatility smile. For this task Malz uses a variation of a standard cubic spline known as a clamped spline. The clamped spline makes specific assumptions about the behaviour of the curve following the first and last knot points such that the curve becomes constant at chosen values.

A detailed explanation of the derivation of cubic spline coefficients can be found in Klugman *et al.* (2008). Briefly, a standard cubic spline is made up of a sequence of cubic functions defined over the intervals between the observed points. Such functions are subject to a series of restrictions that ensure they behave as desired on the interpolating interval x_j to x_{j+1}.

Cubic splines are defined by the following five properties:

$$\hat{y}_j = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3 = f_j(x) \text{ for } x_j \leq x \leq x_{j+1}$$

$$\hat{y}_j = f_j(x_j) = y_j$$

$$f_j(x_{j+1}) = f_{j+1}(x_{j+1})$$

$$f_j'(x_{j+1}) = f_{j+1}'(x_{j+1})$$

$$f_j''(x_{j+1}) = f_{j+1}''(x_{j+1})$$

- The first property defines the cubic function piecewise *between* the observed data points.
- The second ensures that at the observed data point the interpolated value is the same, ensuring that we hit all of the data points.
- The third ensures that the value at the end of an interval is identical to the start of the next one, and hence the interpolated line is *continuous*.
- The fourth and fifth ensure continuity of the first and second derivatives respectively at the points where the functions meet to ensure a *smooth* curve.

Therefore to produce the required spline we need to estimate values for *a*, *b*, *c* and *d* across the *n* intervals between observed *n+1* data points, so in total *4n* coefficients are required. Property 2 provides *n+1* conditions that must hold (the number of data points) while properties 3–5 each provide *n*-1 conditions (the number of intervals minus one because they are conditions that hold at both ends of an interval). Hence we have *n+1+3(n-1)* = *4n-2* conditions for the *4n* coefficients required. Two additional conditions are required to uniquely determine the values of the coefficients: typically these two conditions are used to determine the shape at the end-point of the interpolation. Choice of the end-points gives rise to different types of spline, which we discuss below.

If we let $m_j = f_j''(x_j)$ and number the observations from 0 to *n* we can write in matrix form:

$$Hm = u$$

where

$$H = \begin{bmatrix} 2(\Delta x_1 + \Delta x_2) & \Delta x_2 & 0 & 0 \\ \Delta x_2 & 2(\Delta x_2 + \Delta x_3) & \cdots & \vdots \\ 0 & & & 0 \\ \vdots & \vdots & \ddots & \Delta x_{n-1} \\ 0 & 0 & \Delta x_2 & 2(\Delta x_{n-1} + \Delta x_n) \end{bmatrix}$$

$$u = \begin{bmatrix} 6\left(\dfrac{\Delta y_2}{\Delta x_2} - \dfrac{\Delta y_1}{\Delta x_1}\right) - \Delta x_1 m_0 \\ 6\left(\dfrac{\Delta y_3}{\Delta x_3} - \dfrac{\Delta y_2}{\Delta x_2}\right) \\ \vdots \\ 6\left(\dfrac{\Delta y_n}{\Delta x_n} - \dfrac{\Delta y_{n-1}}{\Delta x_{n-1}}\right) - \Delta x_1 m_n \end{bmatrix} \quad \text{and} \quad m = \begin{bmatrix} m_1 \\ \vdots \\ m_{n-1} \end{bmatrix}$$

If we make some assumptions regarding the values of m_0 and m_n we know all of the values in H and u we can solve for m from:

$$m = H^{-1}u$$

We can subsequently solve for a, b, c and d using the following:

$$a_i = y_i$$

$$b_i = \frac{\Delta y_i}{\Delta x_i} - \frac{\Delta x_i}{6}(2m_{i-1} + m_i)$$

$$c_i = \frac{m_i}{2}$$

$$d_i = \frac{m_i - m_{i-1}}{6\Delta x_i}$$

The most common form of spline is the so called natural spline, where the second derivatives at the end points are set to zero, hence $m_0 = m_n = 0$. For the clamped spline we instead fix the first derivative of the end points to some fixed value (in our case this is zero so that the slope of the curve is flat at the end points, though it could be anything). By setting $f'_j(x_j) = f'_{j+1}(x_{j+1}) = 0$ this implies:

$$m_0 = 3\frac{\Delta y_1}{\Delta x_1^2} - \frac{m_1}{2}$$

$$m_n = -3\frac{\Delta y_n}{\Delta x_n^2} - \frac{m_{n-1}}{2}$$

With the imposition of these restrictions there are slight alterations to H and u, now:

$$H = \begin{bmatrix} 1.5\Delta x_1 + 2\Delta x_2 & \Delta x_2 & 0 & 0 \\ \Delta x_2 & 2(\Delta x_2 + \Delta x_3) & \cdots & \vdots \\ 0 & & & 0 \\ \vdots & \vdots & \ddots & \Delta x_{n-1} \\ 0 & 0 & \Delta x_2 & 2\Delta x_{n-1} + 1.5\Delta x_n \end{bmatrix}$$

and

$$
u = \begin{bmatrix}
6\left(\dfrac{\Delta y_2}{\Delta x_2} - \dfrac{\Delta y_1}{\Delta x_1}\right) - 3\dfrac{\Delta y_1}{\Delta x_1} \\[2mm]
6\left(\dfrac{\Delta y_3}{\Delta x_3} - \dfrac{\Delta y_2}{\Delta x_2}\right) \\[2mm]
\vdots \\[2mm]
6\left(\dfrac{\Delta y_n}{\Delta x_n} - \dfrac{\Delta y_{n-1}}{\Delta x_{n-1}}\right) + 3\dfrac{\Delta y_n}{\Delta x_n}
\end{bmatrix}
$$

We can write simple code in Matlab to create both natural and clamped splines. This is most easily done as a function we can call upon when we wish to calculate the coefficients.

Application – Interpolating using cubic splines in MATLAB

We now give some code to implement the spline methods and then use them to interpolate some representative data.

```matlab
% Code to calculate natural and clamped cubic splines.
% We write the function with three arguments, our X and Y values and
% a binary value for whether we want natural or clamped
function [a, b, c, d] = cubic_spline(x,y,nc)

% set nc = 0 for clamped at zero
%          = anything else for natural

% Define differences
dy = y(2:end)-y(1:end-1);
h  = x(2:end)-x(1:end-1);

% Cubic spline: Natural and clamped use H and u
H = diag(2*(h(1:end-1)+h(2:end)))+diag(h(2:end-1),1)+diag(h(2:end-1),-1);
r = 6./h;
z = zeros(length(r)-1,1);
R = [diag(r(1:end-1))  z z] + [z -diag(r(1:end-1)+r(2:end))  z] + [z z
diag(r(2:end))];
u = R*y;

if nc == 0 % Clamped flat  (zero 1st deriv at boundary)
    H(1,1)      = 1.5*h(1)   + 2*h(2);
    H(end,end)  = 1.5*h(end) + 2*h(end-1);
    u(1)        = u(1)   - 3*dy(1)/h(1);
    u(end)      = u(end) + 3*dy(end)/h(end);
    mc          = H\u;
    m           = [3*dy(1)/(h(1)^2)-mc(1)/2;
                    mc;
                    -3*dy(end)/(h(end)^2)-mc(end)/2];
else % Natural spline (zero 2nd deriv at boundary)
    m           = [0; H\u; 0];
end

% To make calculation easier shift m by 1.
me = m(2:end);
m1 = m(1:end-1);
```

```
% Calculate coefficients
a = y(1:end-1);
b = dy./h - h.*(2*m1 + me)/6;
c = m1/2;
d = (me-m1)./(6*h);
```

The above code permits us to calculate the coefficients. We now need code to evaluate the splines using these coefficients:

```
% Function for evaluating cubic spline using coefficient calculated
% using cubic_spline function. Function needs following entries, range
% over which spline plotted, a,b,c,d the calculated coefficient, and x and
% y data.

function f_interp = cubic_spline_eval(xaxis,a,b,c,d,x,y)

j = 0;
f_interp = zeros(max(size(xaxis)),1);
for xi = xaxis   %
   j = j+1;
   seg = sum(xi>=x);
   if seg > length(a)
      if c(1) == 0
          seg = length(a);
      else
          f_interp(j) = y(end);
          continue
      end
   elseif seg < 1
      if c(1) == 0  % 2nd deriv at start = 0
         seg = 1;
      else
         f_interp(j) = y(1);
         continue
      end
   end
   f_interp(j) = a(seg) + b(seg)*(xi-x(seg)) + c(seg)*(xi-x(seg))^2 +
d(seg)*(xi-x(seg))^3;
end
```
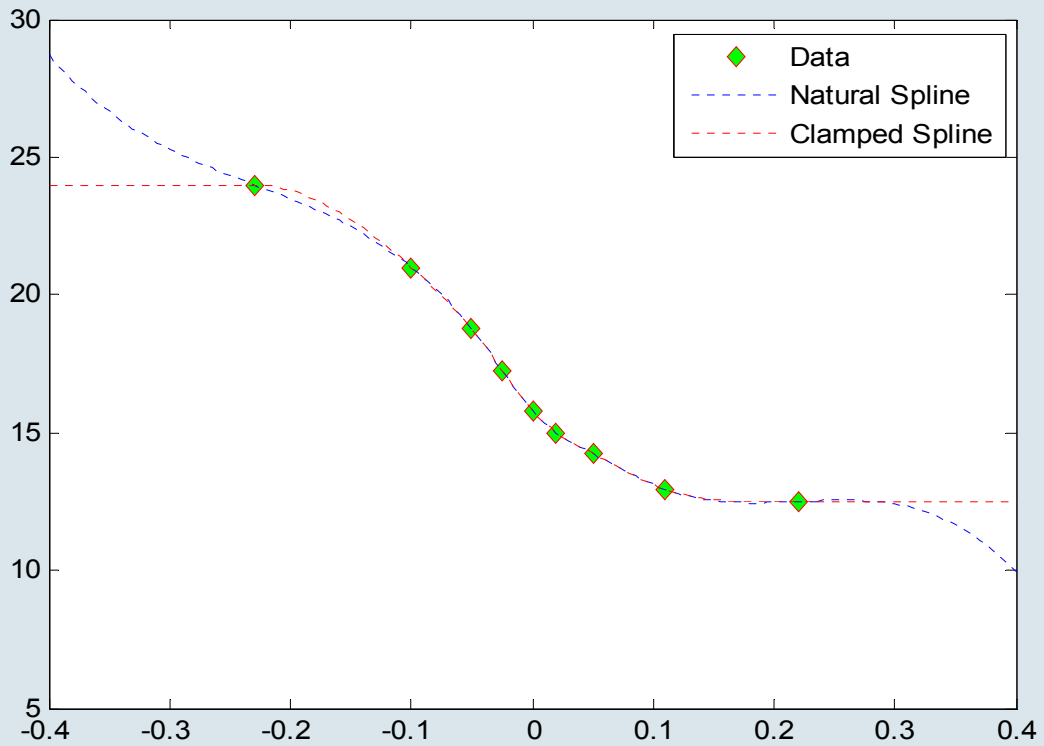
To illustrate assume we have the following pairs of data points:

```
X = [-0.23; -0.1; -0.05; -0.025; 0.001; 0.02;  0.05;  0.11; 0.22];


Y = [   24;   21; 18.75;  17.25; 15.75;   15; 14.25; 12.95; 12.5];
```

We can use the above codes to compute both natural and clamped cubic splines, which we do in the graph.

It is immediately obvious that the natural and clamped splines imply very different behaviour past the ends of the available sample. Malz (2014) discusses in detail why the clamped spline is to be preferred. The additional code needed to draw the precise graph shown is left as an exercise for the interested reader.

References

Bahra, B. (1997) 'Implied risk-neutral probability density functions from options prices: Theory and application', *Bank of England Working Paper* No. 66

Bank of England (n.d.) 'Implied PDF terminology and concepts: Notes on the Bank of England option-implied PDFs', available at
www.bankofengland.co.uk/statistics/Documents/impliedpdfs/boepdf_background_note.pdf

Baz, J. and Chacko, G. (2009) *Financial Derivatives: Pricing, Applications and Mathematics*, Cambridge University Press

Black, F. and Scholes, M. (1973) 'The pricing of options and corporate liabilities', *Journal of Political Economy* 81:637-659

Breeden, D.T. and Litzenberger, R.H. (1978) 'Prices of state-contingent claims implicit in options prices', *Journal of Business* 20:419-438

Cochrane, J.H. (2005) *Asset Pricing*, Princeton University Press, revised edition

Garman, M.B. and Kohlagen, S.W. (1983) 'Foreign currency option values', *Journal of International Money and Finance* 2:231-237

Hull, J. (2006) *Options, Futures and Other Derivatives, Prentice Hall,* sixth edition

Lewis, M. (2012) 'Market perceptions of exchange rate risk', Reserve Bank of New Zealand, Analytical Notes 2012/12

Malz, A. M. (1997) 'Option-implied probability distributions and currency excess returns', *FRBNY Staff Reports* No. 32

Malz, A. M. (2014) 'A Simple and Reliable Way to Compute Option-Based Risk-neutral Distributions', *FRBNY Staff Reports* No. 677

Taylor, S. J. (2005) *Asset Price Dynamics, Volatility and Prediction*, Princeton University Press