OpenGamma Quantitative Research

Local Volatility

Richard White
Richard@opengamma.com

OpenGamma Quantitative Research n. 2 | First version: 27 January 2012; this version February 26, 2013

Abstract

We present details of computing a local volatility surface from market data, then numerically solving different PDE representations to reproduce the market prices, and compute greeks.

Contents

1 The Black-Scholes-Merton Backwards PDE **1**

2 Dupire Local Volatility **2**

3 Interpolation of Volatility Surfaces **4**
3.1 Time interpolation 5

4 Local Volatility **7**

5 PDE Solving and Results **9**
5.1 Strike Sensitivity 9
5.2 Bucketed Vega 11
5.3 Delta and Gamma 12
5.3.1 Gamma 13
5.4 Volatility Greeks 17
5.4.1 Vanna and Vomma 17

6 Conclusion **18**

1 The Black-Scholes-Merton Backwards PDE

Starting from log-normal dynamics for the spot price of some asset, S_t:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1}$$

where μ is the drift, σ is the constant volatility term and W_t is a Weiner process, the classic Black-Scholes-Merton equation for the price of an European option, $V(t, S; T, K)$[1], at time, t, on a underlying, S, struck at K and with an expiry of T, where the risk free (continuously compounded) rate is r, and cost of carry is q[2] , is given by:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - q)S\frac{\partial V}{\partial S} - rV = 0 \tag{2}$$

This equation can be solved numerically, backwards in time, starting from the final condition $V(T, S; T, K) = (S - K)^+$. For a call option, the Dirichlet boundary condition $V(0, t) = 0$ and the Neumann boundary condition $\frac{\partial V}{\partial S}|_{S=M} = e^{-q(T-t)}$ are imposed for some $M \gg k$. Of course, this can also be solved analytically by the Black-Scholes-Merton option pricing formula:

$$V(t, S; T, K) = \omega \left(e^{-q\tau} S_t N(\omega d_1) - e^{-q\tau} K N(\omega d_2) \right) \tag{3}$$

$$d_1 = \frac{\ln(S_t/K) + (r - q)\tau + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}} \qquad d_2 = d_1 - \sigma\sqrt{\tau} \qquad \tau = T - t$$

where ω is +1 for a call and -1 for a put. The model can be extended to term structures for r, q and σ with the time-averaged values used in the formula, i.e.

$$r \to \bar{r} = \frac{1}{T-t}\int_t^T r_s ds \quad r \to \bar{q} = \frac{1}{T-t}\int_t^T q_s ds \quad \sigma^2 \to \bar{\sigma}^2 = \frac{1}{T-t}\int_t^T \sigma_s^2 ds \tag{4}$$

While this will produce different implied volatility levels at different expiries (the implied volatility being simply the root-mean-squared (RMS) of the instantaneous volatility), it cannot produce different implied volatility levels across strikes (volatility smiles). However, if the instantaneous volatility is extended to be a function of both time and the level of the underlying, $\sigma \to \sigma(t, S)$, then the market prices of all options written on the underlying can be recovered by a suitable choice of $\sigma(t, S)$. Additionally, given a (hypothetical) continuum of prices across expiry and strike, there is a unique *Local Volatility* $\sigma_L(t, S)$. This is a result from stochastic processes due to Gyöngy [Gyo86].

Defining the forward option price as $\bar{V}_t = \exp(\int_t^T r_s ds)V_t$ and the forward value of the underlying as

$$F(t, T) = \exp\left(\int_t^T (r_s - q_s)ds\right) S_t \tag{5}$$

equation 2 can be written as

$$\frac{\partial \bar{V}}{\partial t} + \frac{1}{2}\sigma(t, F)^2 F^2 \frac{\partial^2 \bar{V}}{\partial F^2} = 0 \tag{6}$$

[1] here S and t are the state variables and K and T are parameters of the option.

[2] For Equity options q is the dividend yield, while for FX r is the domestic risk free rate r^d, and q is the foreign risk free rate r^f

The forward can be identified as

$$F(t,T) = \mathbb{E}^{\mathbb{T}}\left[S_T|\mathcal{F}_t\right] \tag{7}$$

where the expectation is under the T-forward measure - i.e. the numeraire is the zero coupon bond $P(t,T)$. The option price is $V_t = P(t,T)\bar{V}_t$, and both $F(t,T)$ and $P(t,T)$ are market observables (at time t for a range of T) in many markets. By this transformation, any reference to the (generally) unobserved r and q is removed.

2 Dupire Local Volatility

Using the Fokker-Planck result that for a SDE

$$dx_t = a(t,x_t)dt + b(t,x_t)dW_t \tag{8}$$

the transition probability, $p(T,x_T) \equiv p(T,x_T;t,x_t)$,[3] is governed by the PDE

$$\frac{\partial p(T,x)}{\partial t} = -\frac{\partial[a(T,x)p(T,x)]}{\partial x} + \frac{1}{2}\frac{\partial^2[b^2(T,x)p(T,x)]}{\partial x^2} \tag{9}$$

together with the Breeden-Litzenberger result that

$$p(T,x;t,s_t) = e^{r(T-t)}\frac{\partial^2 C(t,S_t;T,K)}{\partial K^2}\bigg|_{K=x} \tag{10}$$

where $C(\cdot)$ denotes a call option, leads to the forward PDE

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 C}{\partial K^2} - (r-q)K\frac{\partial C}{\partial K} - qC \tag{11}$$

In this equation, the state variables are expiry, T, and strike, K. Numerically solving the PDE forward in time, T, with the initial condition is $C(t,S_t;t,K) = (S_t - K)^+$, will give call prices for all expiries and strikes (within the chosen boundaries). Dupire [Dup94] rearranged this equation to:

$$\sigma(T,K) = \sqrt{\frac{2\frac{\partial C}{\partial T} + (r-q)K\frac{\partial C}{\partial K} + qC}{K^2\frac{\partial^2 C}{\partial K^2}}} \tag{12}$$

which, given a continuous, twice-differentiable in strike and once in time, surface of call options prices, will give a unique local volatility. A real market will only have a finite number of (liquid) option prices. Direct interpolation of market prices is difficult since calendar arbitrage (i.e. $\frac{\partial C}{\partial T} + (r-q)K\frac{\partial C}{\partial K} + qC < 0$) and strike arbitrage (i.e. $\frac{\partial^2 C}{\partial K^2} < 0$) must be avoided. Even if these conditions are met, finding local volatility this way is dangerous, not least because for OTM options, the derivatives will be with respect to very small prices (and thus in turn produce small, possibly inaccurate numbers), leading to a division of one very small number by another. Since it

[3] the probability of going from x_t at time t to x_T at time $T \geq t$

is a market standard to quote prices as implied volatility, $\sigma_{\text{imp}}(T,K)$, the local volatility formula can be rewritten in terms of implied volatility and its derivatives:

$$\sigma(T,K) = \sqrt{\frac{\sigma_{\text{imp}}^2 + 2\sigma_{\text{imp}} T(\frac{\partial \sigma_{\text{imp}}}{\partial T} + (r-q)K\frac{\partial \sigma_{\text{imp}}}{\partial K})}{1 + 2d_1 K\sqrt{T}\frac{\partial \sigma_{\text{imp}}}{\partial K} + K^2 T\left(d_1 d_2\left(\frac{\partial \sigma_{\text{imp}}}{\partial K}\right)^2 + \sigma_{\text{imp}}\frac{\partial^2 \sigma_{\text{imp}}}{\partial K^2}\right)}} \tag{13}$$

Interpolation of the implied volatility surface is discussed in section 3, but first we will reformulate equations 11 and 13 to eliminate r and q.

We define the moneyness, $x = K/F(t,T)$ and the fractional call price as

$$\hat{C}(T,x) \equiv \frac{C(T, xF(t,T))}{P(t,T)F(t,T)} \tag{14}$$

With these changes of variables, the forward PDE for the fractional call price is

$$\frac{\partial \hat{C}}{\partial T} = \frac{1}{2}\hat{\sigma}(T,x)^2 x^2 \frac{\partial^2 \hat{C}}{\partial x^2} \tag{15}$$

and the local volatility is now a function of expiry and moneyness (denoted by the hat symbol). The relationship between these two local volatility surfaces is trivially

$$\sigma(T,K) = \hat{\sigma}(T, K/F(t,T)) \qquad \hat{\sigma}(T,x) = \sigma(T, xF(t,T)) \tag{16}$$

Equation 15 must be solved numerically[4], with initial condition $\hat{C}(0,x) = (1-x)^+$, lower boundary condition $\hat{C}(T,0) = 1$ and upper boundary condition either $\hat{C}(0, x_{\max}) = 0$ or $\frac{\partial \hat{C}}{\partial x}|_{x=x_{\max}} = 0$, for some $x_{\max} \gg 1$. After reversing the change of variables, this gives the call price across all expiries and strikes (within the set boundaries).

Just as in equation 13, we can rearrange equation 15 to give local volatility in terms of implied volatility:

$$\hat{\sigma}(T,x) = \sqrt{\frac{\hat{\sigma}_{\text{imp}}^2 + 2\hat{\sigma}_{\text{imp}} T\frac{\partial \hat{\sigma}_{\text{imp}}}{\partial T}}{1 + 2d_1 x\sqrt{T}\frac{\partial \hat{\sigma}_{\text{imp}}}{\partial x} + x^2 T\left(d_1 d_2\left(\frac{\partial \hat{\sigma}_{\text{imp}}}{\partial x}\right)^2 + \hat{\sigma}_{\text{imp}}\frac{\partial^2 \hat{\sigma}_{\text{imp}}}{\partial x^2}\right)}} \tag{17}$$

The implied volatility, $\hat{\sigma}_{\text{imp}}(T,x)$, is a function of expiry and moneyness, and

$$d_1 = \frac{-\ln(x) + \frac{1}{2}\sigma^2\tau}{\sigma\sqrt{\tau}} \qquad d_2 = d_1 - \sigma\sqrt{\tau} \qquad \tau = T - t \tag{18}$$

Again, it is trivial to convert between an implied volatility surface parameterised by strike to one parameterised by moneyness.

Assuming we have obtained a smooth, interpolated, implied volatility surface from market prices of options on a single underlying, we can then numerically or analytically take derivatives to obtain the local volatility surface. Armed with this surface we numerically integrate (i.e. solve) equation 15 once, read off the prices and compare with the market prices. Alternatively, we can solve equation 6 once for each option. Either way the only discrepancy with the input market option prices should be due to numerical error.

[4]Except in the degenerate case where $\hat{\sigma}$ is a function of time only, in which case the RMS volatility can be plugged straight into the Black formula.

3 Interpolation of Volatility Surfaces

The first condition for an interpolated volatility surface is that it matches exactly the (liquid) market option prices[5]. To obtain a continuous local volatility surface, the implied volatility surface should be at least C^1 (once differentiable) in the T direction and C^2 in the strike/moneyness direction, and in general a (C^n_T, C^m_K) implied volatility surface, will produce a (C^{n-1}_T, C^{m-2}_K) local volatility surface.

The condition to avoid calendar arbitrage is $\hat{\sigma}^2_{\text{imp}} + 2\hat{\sigma}_{\text{imp}} T \frac{\partial \hat{\sigma}_{\text{imp}}}{\partial T} \geq 0$. Defining integrated implied variance as $\hat{\nu}_{\text{imp}}(T,x) = T\hat{\sigma}^2_{\text{imp}}(T,x)$, the condition can be rewritten as

$$\frac{\partial \hat{\nu}_{\text{imp}}}{\partial T} \geq 0 \tag{19}$$

In the strike direction, the arbitrage condition is more complex, but can be satisfied by using a smile model that does not admit arbitrage.

To make the discussion concrete we consider FX options with ten expiries from 1 week to 10 years, and 5 strikes per expiry (with deltas of 0.25, 0.35, 0.5, 0.75 and 0.85). Since these data are dense on an expiry/strike grid, we can fit a separate smile model at each expiry, then interpolate in the time direction between these fitted smiles.

We choose to use the SABR model[6], which has four parameters. As such it is not possible to guarantee an exact fit to all five option strikes. Since the CEV parameter, β, and the correlation, ρ, have similar effects on the smile, they tend to play off against each other when fitting the parameters, and it is common practice to fix β and fit for the other three parameters. We choose β by running a least-squares fit between model- and market-implied volatilities for all five strikes. With β fixed, we then make three different fits of SABR to the three sets of three consecutive points. These fits should be exact (a failure would indicate bad data). Clearly, the fits will agree on the implied volatility at the market strikes that they share (which in turn will be the market-implied volatilities), but not at points in between. For the points in between, we take a weighted average. For a point x, between x_i and x_{i+1}, the weighted average is

$$f(x) = w\left(\frac{x_{i+1}-x}{x_{i+1}-x_i}\right) f_i(x) + \left[1 - w\left(\frac{x_{i+1}-x}{x_{i+1}-x_i}\right)\right] f_{i+1}(x) \tag{20}$$

where $f_i()$ is the fit centred on the point x_i and the weight function has the property $w(0) = 0$ and $w(1) = 1$. An obvious choice is $w(y) = y$, but we must consider how the derivatives of $f()$ will behave.

The first and second derivatives are;

$$\begin{aligned} f'(x) = \frac{1}{\Delta x_i} w'\left(\frac{x_{i+1}-x}{\Delta x_i}\right)[-f_i(x) + f_{i+1}(x)] \\ + w f'_i(x) + (1-w) f'_{i+1}(x) \end{aligned} \tag{21}$$

$$\begin{aligned} f''(x) = \frac{1}{\Delta x_i^2} w''\left(\frac{x_{i+1}-x}{\Delta x_i}\right)[f_i(x) - f_{i+1}(x)] \\ - \frac{2}{\Delta x_i} w'\left(\frac{x_{i+1}-x}{\Delta x_i}\right)\left[f'_i(x) - f'_{i+1}(x)\right] \\ + w f''_i(x) + (1-w) f''_{i+1}(x) \end{aligned} \tag{22}$$

[5] It is not strictly interpolation if this is not met, however there are situations where a little mispricing is tolerable to obtain a smooth, arbitrage-free surface.

[6] SABR is ubiquitous despite its shortcomings, not least of which is that it can admit arbitrage.

This means the first derivative is continuous for any choice of the weight function. However the second derivative can be discontinuous regardless of how smooth the basis function f_i is:

$$\begin{aligned}\lim_{\epsilon\to 0} f''(x_i+\epsilon) - f''(x_i-\epsilon) = &\frac{2}{\Delta x_i} w'(1)\left[f'_{i+1}(x_i) - f'_i(x_i)\right] \\ &-\frac{2}{\Delta x_{i-1}} w'(0)\left[f'_i(x_i) - f'_{i-1}(x_i)\right]\end{aligned} \tag{23}$$

If the additional constraint $w'(0) = w'(1) = 0$ is applied, then the second derivative will be continuous. A candidate weight function is

$$w(y) = \frac{1}{2}\left(\sin\left[\pi\left(y - \frac{1}{2}\right)\right] + 1\right) \tag{24}$$

Extrapolation is handled by using either the SABR fit to the lowest three strikes (for strikes less than the lowest market strike) or the fit to the highest three strikes (for strikes greater than the highest market strike)[7].

3.1 Time interpolation

Clark [Cla11] points out that in FX markets there is no connection between the volatility levels at a particular strike across expiries, and therefore it makes little sense to interpolate between common strike levels from the smile fits in the time direction. Instead, time interpolation should be performed between common delta values. If the volatility at some (T, Δ) is required, the corresponding volatility at each fitted smile is found by root finding for the strike[8]. Once these volatilities are known, we can interpolate for the time T. An extra complication is that the PDEs we will solve require the volatility at (potentially arbitrary) expiry and strike (or moneyness) points, rather than expiry and delta, meaning we would have to iterate the above procedure to find the delta and volatility at the required point. As this would be done several thousand times to solve the PDE, it makes it impractical.

Clark [Cla11] suggests interpolating the ATM volatility, risk reversals and strangles[9] to the required expiry, then performing a smile fit from these interpolated values. This may require performing 50-100 separate smile fits, and be liable to numerical instability.

To avoid time consuming root finding, we define a proxy delta as

$$d = \frac{\ln\left(\frac{F}{K}\right)}{\sqrt{T}} \tag{25}$$

The volatility of common d values on four adjacent fitted smiles is then found, and the integrated variances computed. Note that the condition of equation 19 (i.e. integrated variance increases with time) need not hold for these four values as we are not moving along a line of constant moneyness. We find that a log-log natural cubic spline interpolation[10] works best for the data (however we note that this method can admit calendar arbitrage). Figure 1 shows the implied volatility surface using this method.

[7] With SABR there is a danger of getting arbitrage at very low strikes.

[8] Of course, if the fitted smile was parameterised by delta in the first place, this root finding would be unnecessary.

[9] The typical way FX volatilities are quoted.

[10] taking the log of the time values and the log of the integrated variance, and interpolate these values.

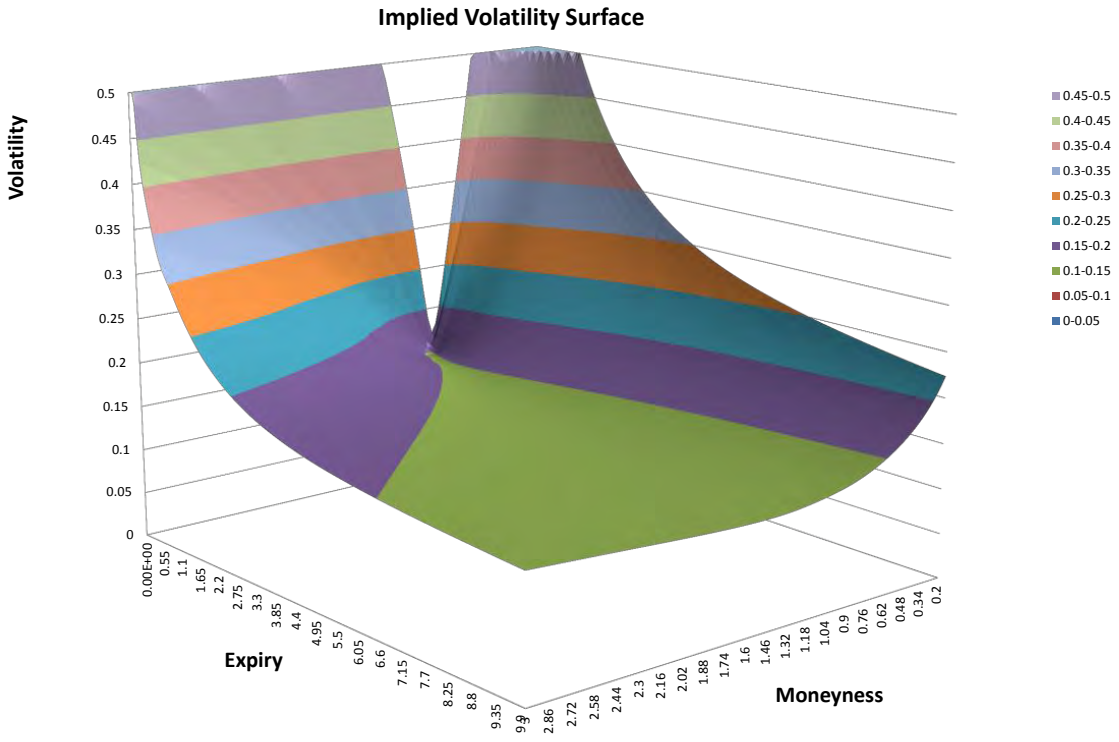


Figure 1: The implied volatility surface fitted to FX market data using the technique discussed in the main text.

4 Local Volatility

If the implied volatility surface is treated as a black box, i.e. you request the volatility at a given expiry and strike (or moneyness) and get a number back, then the only way to form the local volatility surface is by forming difference quotients as approximations to the partial derivatives. For each point in the local volatility surface, the value at five (close by) points on the implied volatility surface are needed. If we can analytically differentiate the smile model with respect to strike (which you can for SABR), then all the information is available to calculate the partial derivatives analytically, which will be faster and numerically more stable.

Figure 2 shows the local volatility surface derived from the implied volatility surface of figure 1. As expected, it is considerably less smooth.

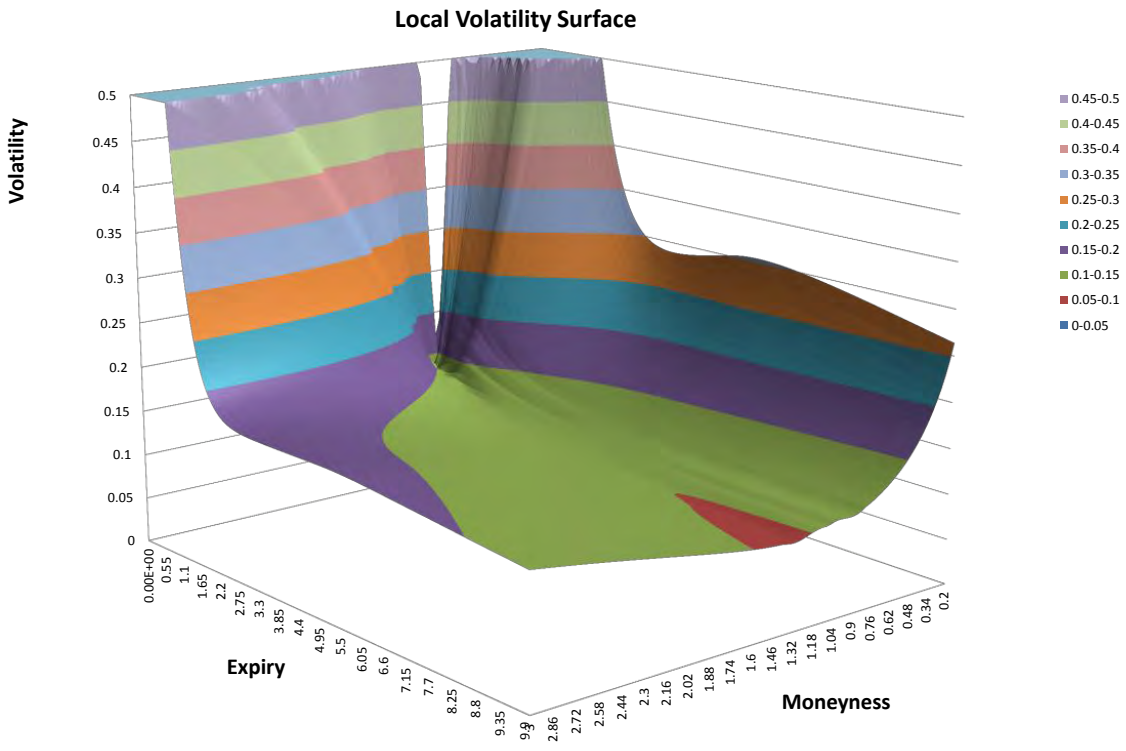


Figure 2: The local volatility surface derived from FX market data using the technique discussed in the main text.

Expiry-moneyness is not the best way to display the volatility surface, as there is little relevance in say the volatility (implied or local) of a 1W expiry, 0.2 moneyness option. Following the definition of proxy delta above, we show the same local volatility surface in expiry - proxy delta coordinates in figure 3.

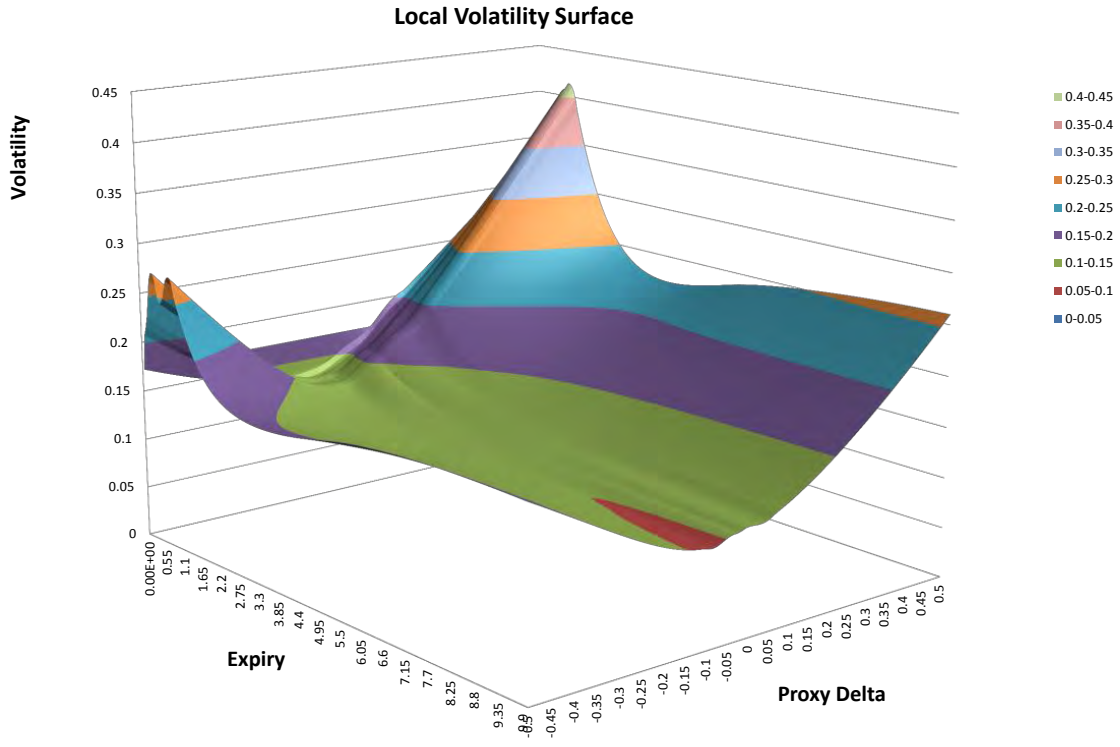


Figure 3: The local volatility surface derived from FX market data using the technique discussed in the main text and shown in expiry - proxy delta coordiantes.

5 PDE Solving and Results

We solve the PDE using standard finite difference techniques described in [Duf06], with a weighting between fully explicit and fully implicit time stepping schemes[11]. The grid is non-uniform, having a greater density of points near $t = 0$ and around the strike (or moneyness equals 1). The details of the PDE solver call be found in another note, [Whi13].

Having formed the local volatility surface (parameterised by moneyness), we solve equation 15 using 100 points in both time and moneyness directions (a total of 10000 points in the grid), from time (expiry) 0 to 10 years, and moneyness 0 to 3.5. The fractional call price is then converted to an implied volatility for each grid point[12], and interpolated. Figure 4 shows the market and model implied volatilities - the maximum error is 0.1% and the average absolute error is 1.5bps.

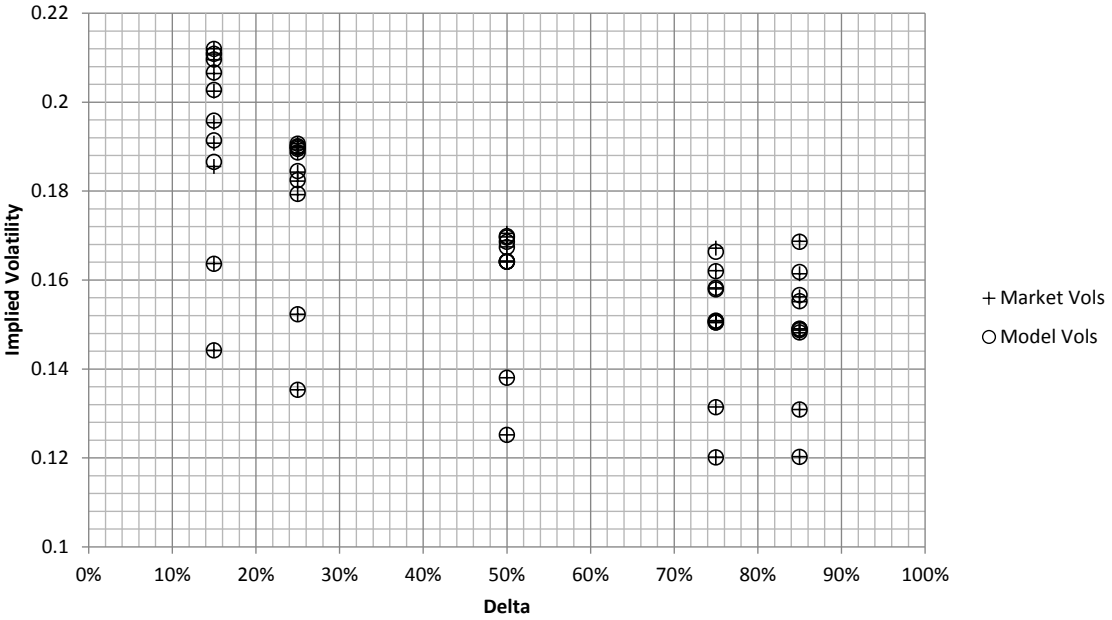


Figure 4: The market and model (i.e. fit implied volatility surface → derive the local volatility surface → run a PDE solver) implied volatilities.

5.1 Strike Sensitivity

As well as fitting the market well, the smiles should be smooth. Figure 5 shows the smiles at 1 week and 5 years over a relevant range of strikes. There is no numerical noise.

[11] An equal weighting is the Crank-Nicolson scheme, which has some numerical instability problems (see [Duf04].
[12] actually the moneyness is restricted to 0.3 to 3.0

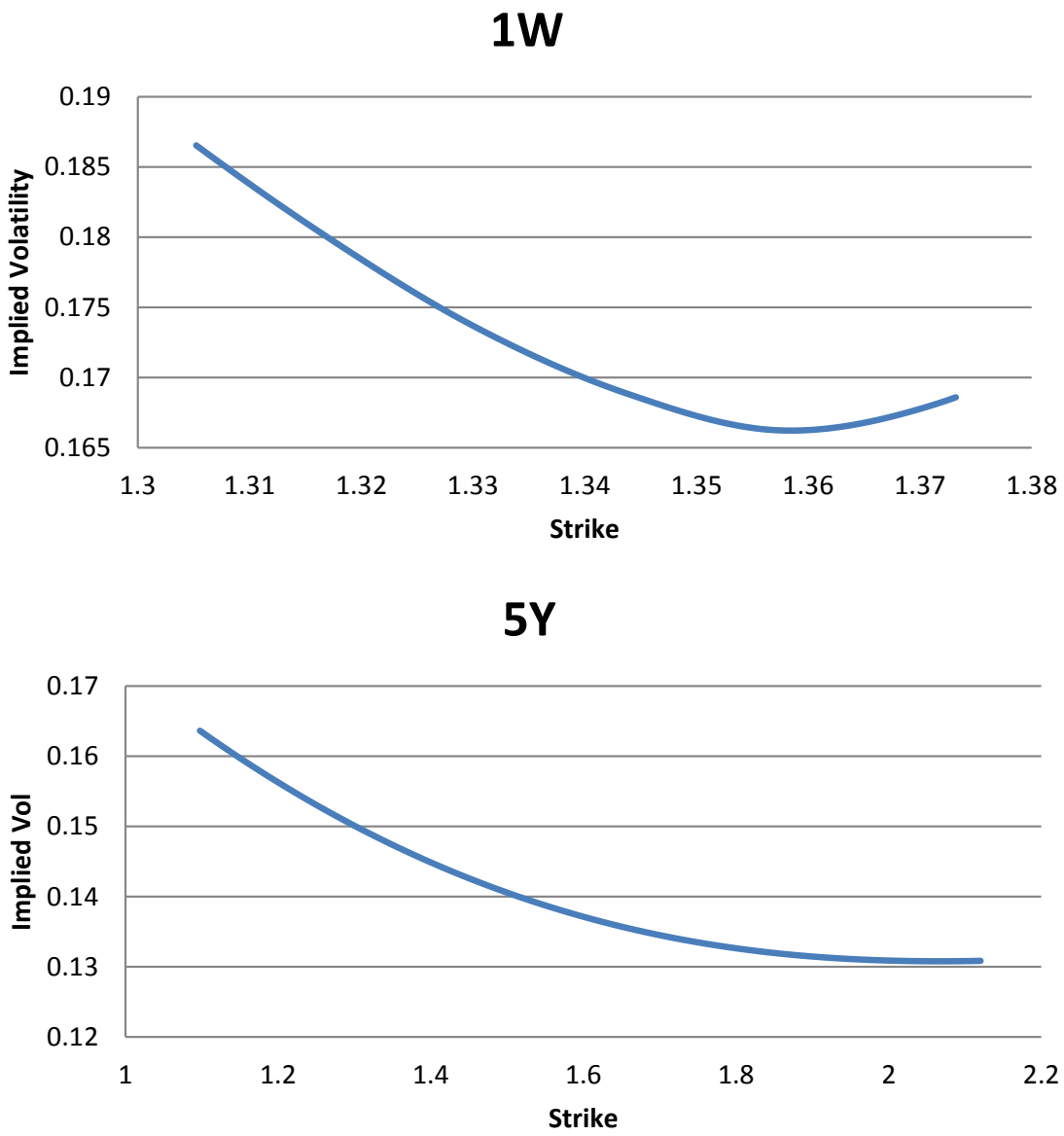


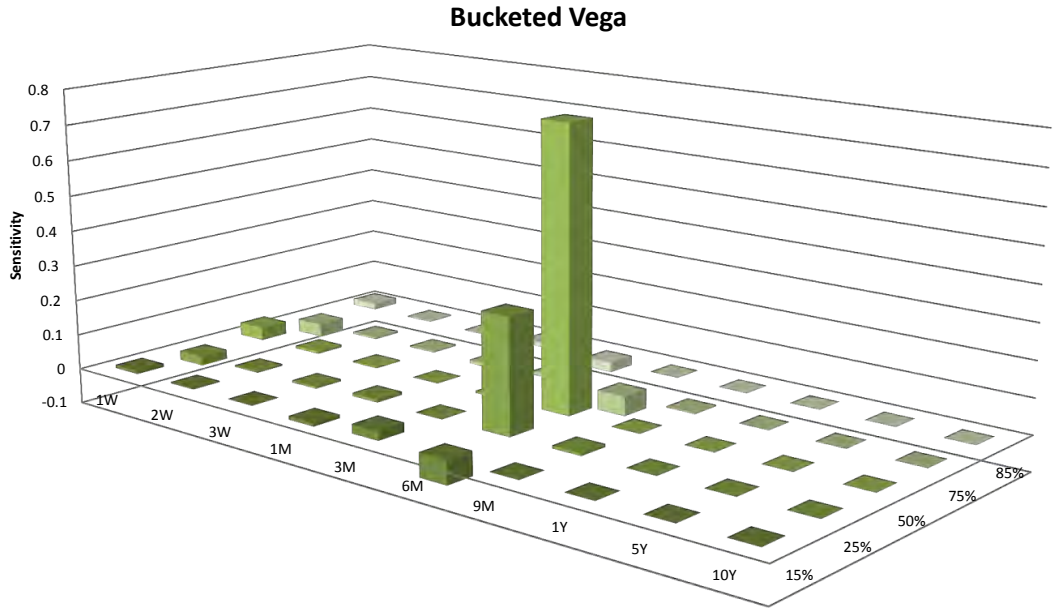
Figure 5: The smile at 1 week and 5 years.

5.2 Bucketed Vega

Another useful diagnostic is to measure the sensitivity of a representative point to changes in the input data. The point is an option with six month expiry and a strike of 1.4 (slightly ITM). Each of the 50 market data points has its implied volatility shifted by 1 basis point in turn, the implied volatility surface refitted, the local volatility surface recalculated, and finally a PDE solver run with the new surface. The sensitivity is defined as

$$s = \frac{\sigma_{\text{original}} - \sigma_{\text{bumped}}}{\epsilon}$$

and $\epsilon = 1$ basis point. Almost identical results are obtained by solving the backwards PDE (eqn. 6) as solving the forward PDE (eqn. 15). Figure 6 shows the result for the backwards PDE. There is a large amount of sensitivity to the 6M ATM (which corresponds to a strike of 1.44[13]) and the 6M 25% delta (with strike 1.32), with smaller amounts to the other 6M points, and very little sensitivity to the rest of the market data - which is desirable.



	1W	2W	3W	1M	3M	6M	9M	1Y	5Y	10Y
15%	9.71E-03	3.48E-03	1.11E-03	-1.02E-02	2.63E-02	-6.79E-02	-6.69E-04	3.35E-05	0.00E+00	0.00E+00
25%	-2.69E-02	-4.88E-03	-6.38E-03	9.17E-03	9.20E-04	3.20E-01	8.35E-03	-1.60E-04	0.00E+00	0.00E+00
50%	3.74E-02	7.42E-03	4.00E-03	2.76E-04	6.39E-03	7.85E-01	-2.73E-04	-2.45E-04	0.00E+00	0.00E+00
75%	-3.84E-02	-5.39E-03	-4.39E-03	-1.06E-02	1.08E-02	-5.67E-02	-3.54E-03	2.66E-05	0.00E+00	0.00E+00
85%	1.87E-02	2.87E-03	-1.75E-03	1.31E-02	-2.79E-02	-7.23E-04	2.44E-04	-4.23E-10	0.00E+00	0.00E+00

Figure 6: Bucketed vega - the sensitivity of the implied volatility of an option (expiry 6M, strike 1.4) to the implied volatility of market inputs.

[13]The ATM is taken to be the delta-neutral straddle (DNS), so the strike is given by $K = F\exp(\sigma_{ATM}^2 T/2)$.

5.3 Delta and Gamma

In this case, delta refers to the forward delta (or driftless delta), which is the sensitivity of the future value (FV) of an option to the relevant forward. This is known as pips forward delta in FX, and in a Black-Sholes-Merton world is given by

$$\Delta_{F;\text{Black}} = \omega N(\omega d_1) \tag{26}$$

For a stochastic volatility model where the implied volatility is a function of the forward (such as SABR), the full delta can be written as

$$\Delta_F = \Delta_{F;\text{Black}} + \nu_{F;\text{Black}} \frac{\partial \sigma}{\partial F} \tag{27}$$

where $\nu_{F;\text{Black}}$ is the Black forward vega (sensitivity of the forward value of an option to its implied volatility). For a model with some parameters, $\boldsymbol{\theta}$, that have been calibrated to the market, the term $\frac{\partial \sigma}{\partial F}$ is understood to be the sensitivity with those parameters fixed.[14]

For local volatility, it is assumed that the local volatility surface (parameterised by strike) is invariant to a change in the forward - this assumption will produce a different delta to that given by a stochastic volatility model even when they agree exactly on price.

Solving equation 6 for a particular strike and expiry, T, produces a set of forward values at different initial levels of the relevant forward, $F(0,T)$. The delta can then be read off by taking the difference quotient of these values[15]. To produce delta as a function of strike, the PDE is solved 100 times at different strikes, with the expiry fixed at six months.

Extracting deltas from the solution to equation 15 is more challenging since the spatial variable is now moneyness. Delta can be written as:

$$\Delta_F = \frac{\partial [F\hat{C}]}{\partial F} = \hat{C} + \frac{\partial \hat{C}}{\partial F} = \hat{C} - x\frac{\partial \hat{C}}{\partial x} + \text{surface delta} \tag{28}$$

The term *surface delta* comes from the fact that, if the local volatility parameterised by strike is invariant to the forward, the surface parameterised by moneyness cannot be. Writing the local volatility explicitly as a function of the forward, $\hat{\sigma}(T,x;F_T)$, the delta is written as:

$$\begin{aligned}\Delta_F =& \hat{C}(T,x;\hat{\sigma}(T,x;F_T)) - x\frac{\partial \hat{C}(T,x;\hat{\sigma}(T,x;F_T))}{\partial x} \\ &+ \lim_{\epsilon\to 0} \frac{\hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T)) - \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T))}{2\epsilon F_T}\end{aligned} \tag{29}$$

The PDE is solved three times; once with an unmodified surface, then once each with a surface made by fractionally shifting the forward up and down by an amount $\epsilon = 0.05$. The first term is obtained from the unmodified solution, the second by taking difference quotients, and the third by taking the difference between the two modified solutions.

Aside: If the third term is missed out, then the implicit assumption is that the local volatility surface parameterised by moneyness is invariant to the forward curve. This will give the same

[14] The $\frac{\partial \sigma}{\partial F}$ term is known as the backbone, and is sometimes defined extraneously to produce the smile dynamics 'expected' by the trader.

[15] To get the delta at the actual forward, if it does no lie on the grid, linear interpolation is used.

delta as a stochastic volatility model whose implied volatility is a function of moneyness only - this included SABR when $\beta = 1$ and the Heston model.

Figure 7 shows the smile produced from solving the forward and backwards PDE (recall that in the case of the backwards PDE, it is solved 100 times with different strikes). The agreement is good, only visibly deviating for very large strikes. Figure 8 shows the corresponding delta along with the Black delta. The agreement is good between the solutions of the two PDEs.

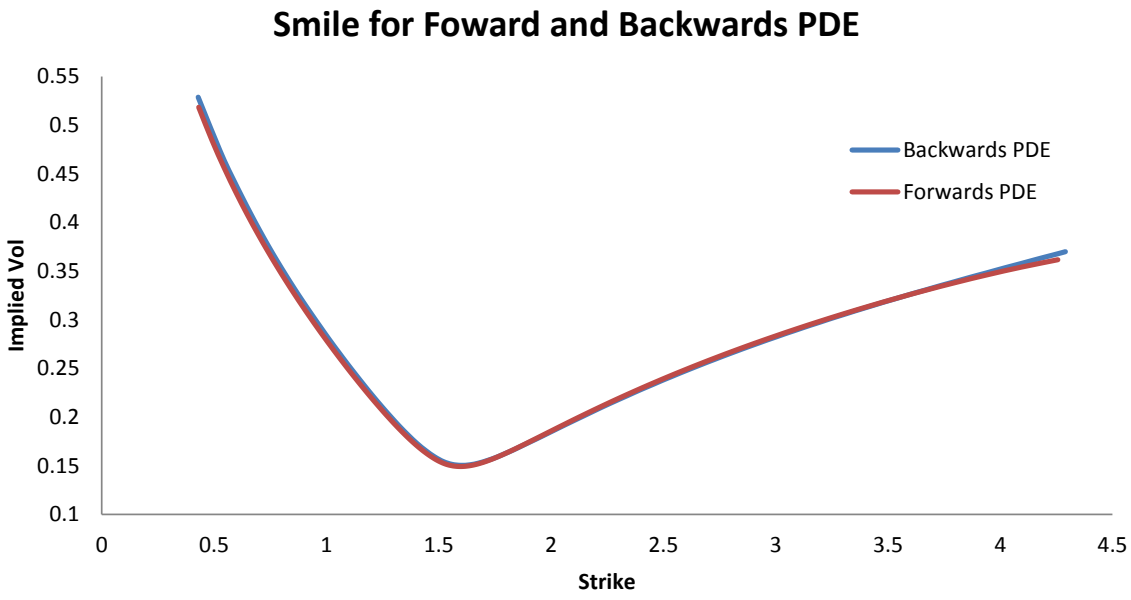


Figure 7: The volatility smile at six months from solving the forward and backwards PDE.

A delta below the Black delta indicates (from equation 27) that $\frac{\partial \sigma}{\partial F} < 0$. This is the case for strikes below ATM, and the opposite is true above ATM. From the shape of the smile, this suggests that the smile will move to the left as the forward increases. This is confirmed in figure 9 where smiles have been produced as above, but with the forward increased by 10%. This smile dynamic for local volatility is the opposite of what is seen in the market, as noted by Hagan (2002).

5.3.1 Gamma

As with delta, the true gamma can be written as

$$\Gamma_F = \Gamma_{F;\text{Black}} + 2\text{Vanna}_{F;\text{Black}} \frac{\partial \sigma}{\partial F} + \nu_{F;\text{Black}} \frac{\partial^2 \sigma}{\partial F^2} \tag{30}$$

There are two correction terms to the Black forward gamma. The local volatility gamma is easily found from the second order difference quotient of the solution to the backwards PDE, and repeated 100 times at different strikes. Finding gamma from the forward PDE is more

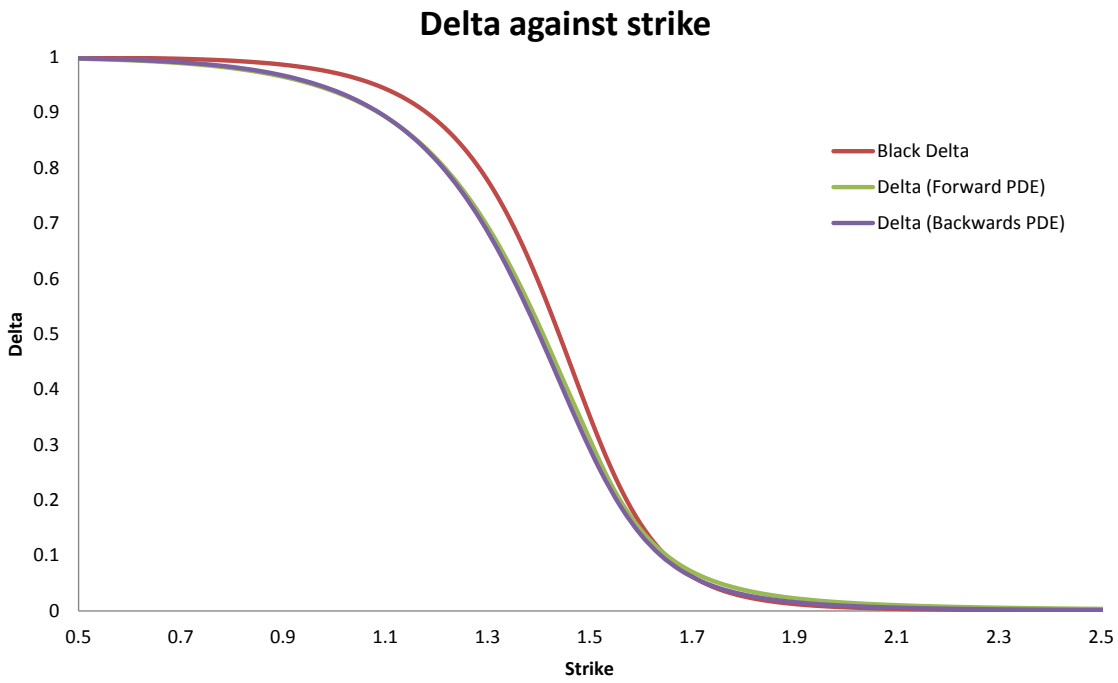


Figure 8: The Black delta and the local volatility delta

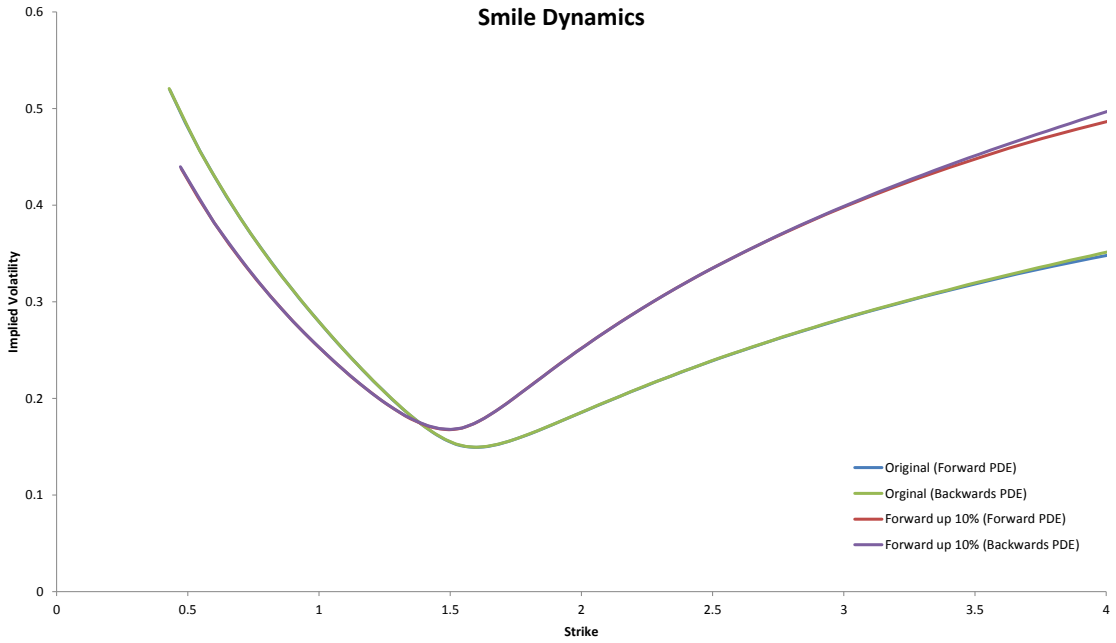


Figure 9: The change in the smile when the forward is increased by 10%.

involved. Using the same notation as for delta, it is written:

$$\begin{aligned}\Gamma_F = & \frac{x^2}{F_T}\frac{\partial^2 \hat{C}(T,x;\hat{\sigma}(T,x;F_T))}{\partial x^2} \\ & + \frac{\hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T)) - \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T))}{2\epsilon F_T} \\ & - x\frac{\frac{\partial \hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T))}{\partial x} - \frac{\partial \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T))}{\partial x}}{\epsilon F_T} \\ & + \frac{\hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T)) + \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T)) - 2\hat{C}(T,x;\hat{\sigma}(T,x;F_T))}{\epsilon^2 F_T}\end{aligned} \tag{31}$$

The first term is the gamma if the local volatility, parameterised by moneyness, were invariant to the forward curve. The second term is the surface delta we saw above. The third term could be called (with a slight abuse of terminology) the surface vanna - it is the change in the moneyness delta due to a change in the surface. Finally the last term could obviously be called the surface gamma. Figure 10 shows the gamma from the two PDE solutions along with the Black gamma. Again, the agreement is good between the solutions of the two PDEs, although it is not clear what is causing the shoulder to appear in the forward PDE gamma.[16]

[16]It was thought that this was due to well-known numerical problems with Crank-Nicolson, but running a fully implicit PDE solver does not cure it.

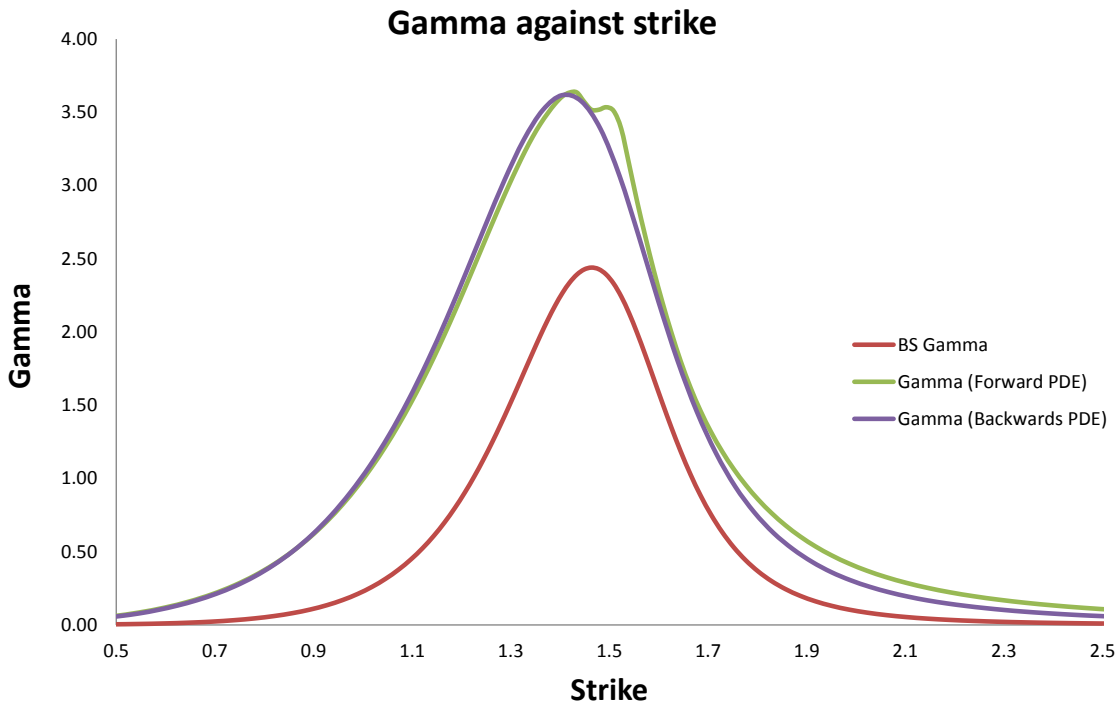


Figure 10: The Black gamma and the local volatility gamma

5.4 Volatility Greeks

Vega in a BSM world is well defined, as volatility is a parameter of the pricing formula. For a local volatility model, it could mean the sensitivity of price to any deformation of the local volatility surface. Standard deformations could be a parallel shift (e.g. each point is increased by 1 basis point of volatility) or a fractional shift (e.g. each point is increased by 0.01% of its value). Figure 11 show the results for a parallel shift; the surface is moved up and down by 1bp, and the vega calculated as the difference between the prices divided by 2 bps. As usual, we are taking an expiry of six months.

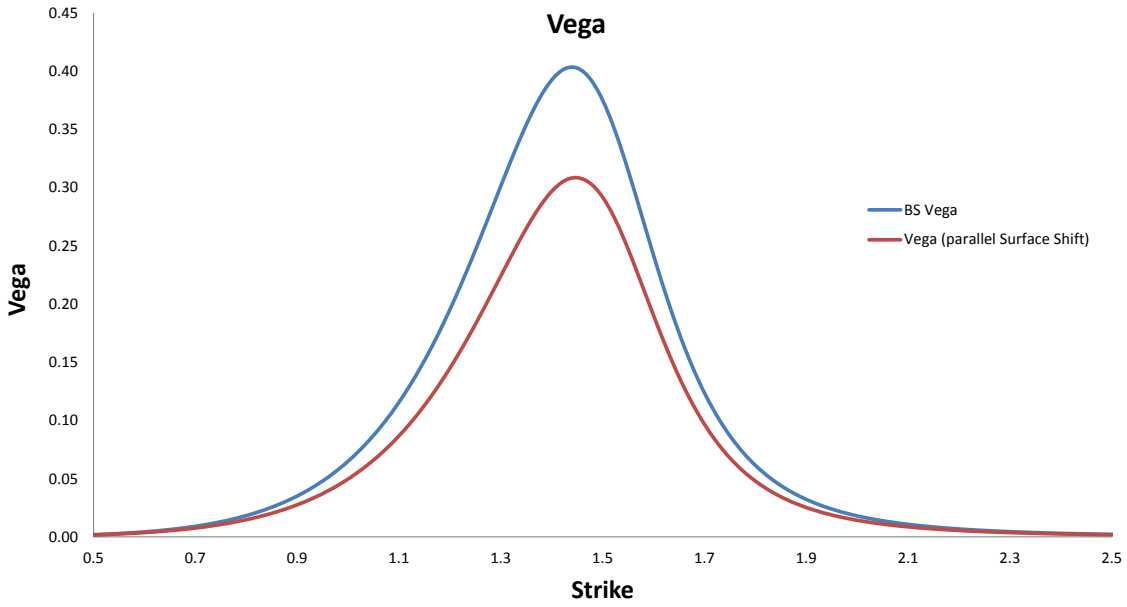


Figure 11: The Black vega and the local volatility vega computed from parallel shifts of the surface, for six month expiries

5.4.1 Vanna and Vomma

Vanna is the sensitivity of delta to the implied volatility (i.e. $\frac{\partial^2 C}{\partial \sigma \partial F}$) and vomma is the second derivative of price with respect to implied volatility (i.e. $\frac{\partial^2 C}{\partial \sigma^2}$). Again, both have very clear definitions in a BSM world. As with vega, we calculate the local volatility vomma as the second order difference quotient from parallel shifts in the surface. Vanna is more complex (mainly because computing delta from a moneyness parameterised forward PDE is complex), and can be

expressed as

$$\begin{aligned}\text{Vanna}_F = &\frac{\hat{C}(T,x;\hat{\sigma}(T,x;F_T)+\eta) - \hat{C}(T,x;\hat{\sigma}(T,x;F_T)-\eta)}{2\eta} \\ &- x\frac{\frac{\partial \hat{C}(T,x;\hat{\sigma}(T,x;F_T)+\eta)}{\partial x} - \frac{\partial \hat{C}(T,x;\hat{\sigma}(T,x;F_T)+\eta)}{\partial x}}{2\eta} \\ &+ \frac{1}{4\epsilon\eta}\Big(\hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T)+\eta) + \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T)-\eta) \\ &-\hat{C}(T,x;\hat{\sigma}(T,x;(1+\epsilon)F_T)-\eta) + \hat{C}(T,x;\hat{\sigma}(T,x;(1-\epsilon)F_T)+\eta)\Big)\end{aligned} \tag{32}$$

The first term is simply the vega of the fractional price, $\hat{C}$, the second is the vanna of the fractional price, and the third is the cross second order derivative to changes to the volatility surface due to deformation from change in the forward, and parallel shifts. Figures 12 and 13 show the Black and local volatility vanna and vomma.

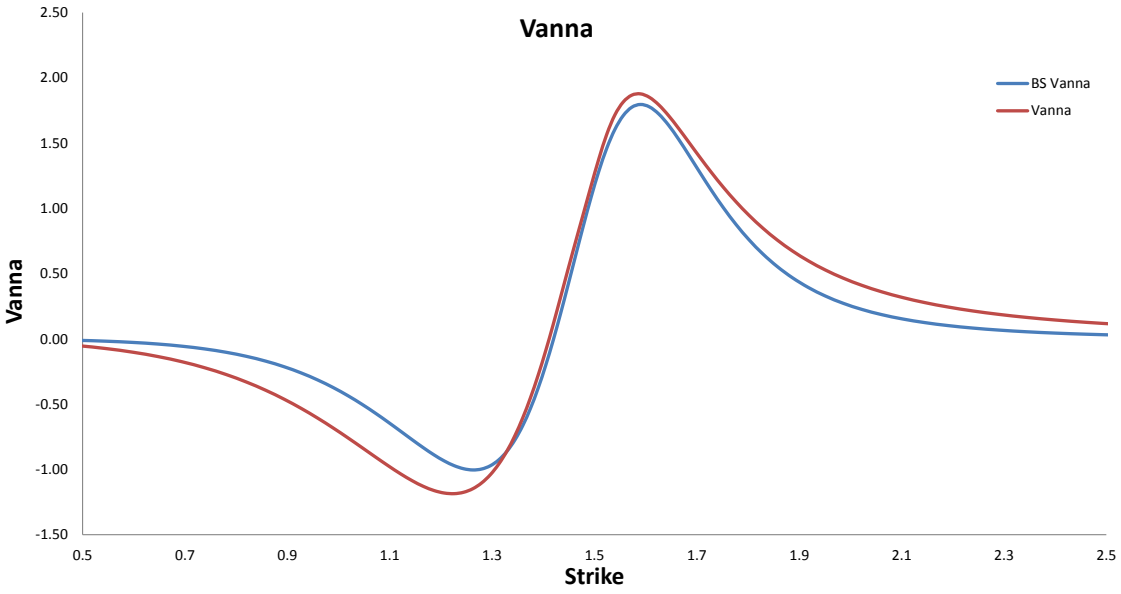


Figure 12: The Black vanna and the local volatility vanna computed from parallel shifts of the surface, for six month expiries

6 Conclusion

We have shown that our implementation of local volatility (which includes fitting a smooth implied volatility surface to market data, deriving a local volatility surface from this, and numerically solving one of two PDEs) reprices the market almost exactly. Furthermore, it produces

Figure 13: The Black vomma and the local volatility vomma computed from parallel shifts of the surface, for six month expiries

local behaviour of bucketed vega (i.e. the price of a representative option only depends on the values of nearby market inputs) and smooth greeks.

References

[Cla11] Iain J. Clark. *Foreign exchange option pricing*. Wiley Finance, 2011. 5

[Duf04] Daniel J. Duffy. A critique of the crank nicolson scheme strengths and weaknesses for financial instrument pricing. *Wilmott magazine*, 2004. 9

[Duf06] Daniel J. Duffy. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. Wiley, 2006. 9

[Dup94] Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994. Reprinted in Derivative Pricing:The Classical Collection, Risk Books (2004). 2

[Gyo86] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields*, 71:501–516, 1986. 1

[Whi13] Richard White. Numerical solutions to pdes with financial applications. Technical report, OpenGamma, 2013. developers.opengamma.com/quantitative-research/Numerical-Solutions-to-PDEs-with-Financial-Applications-OpenGamma.pdf. 9

OpenGamma Quantitative Research

1. Marc Henrard. Adjoint Algorithmic Differentiation: Calibration and implicit function theorem. November 2011.
2. Richard White. Local Volatility. January 2012.
3. Marc Henrard. My future is not convex. May 2012.
4. Richard White. Equity Variance Swap with Dividends. May 2012.
5. Marc Henrard. Deliverable Interest Rate Swap Futures: Pricing in Gaussian HJM Model. September 2012.
6. Marc Henrard. Multi-Curves: Variations on a Theme. October 2012.
7. Richard White. Option pricing with Fourier Methods. April 2012.
8. Richard White. Equity Variance Swap Greeks. August 2012.
9. Richard White. Mixed Log-Normal Volatility Model. August 2012.
10. Richard White. Numerical Solutions to PDEs with Financial Applications. February 2013.

About OpenGamma

OpenGamma helps financial services firms unify their calculation of analytics across the traditional trading and risk management boundaries.

The company's flagship product, the OpenGamma Platform, is a transparent system for front-office and risk calculations for financial services firms. It combines data management, a declarative calculation engine, and analytics in one comprehensive solution. OpenGamma also develops a modern, independently-written quantitative finance library that can be used either as part of the Platform, or separately in its own right.

Released under the open source Apache License 2.0, the OpenGamma Platform covers a range of asset classes and provides a comprehensive set of analytic measures and numerical techniques.

Find out more about OpenGamma ____ **Download the OpenGamma Platform**

Europe
OpenGamma
185 Park Street
London SE1 9BL
United Kingdom

North America
OpenGamma
230 Park Avenue South
New York, NY 10003
United States of America