5

Implied Volatility Formulas and Calibration

By *calibration* we mean adjusting the model parameters so that the model option prices reproduce in the best possible way the market prices for a given set of observed derivatives. Our approach consists of using, instead of the model prices, their approximations derived in the previous chapter. We then calibrate the group market parameters produced by our first-order approximation instead of the parameters of a fully specified stochastic volatility model. This *reduction of parameters* is a main feature of our method. Its full strength resides in the fact that these parameters are exactly those needed to price other derivatives to the same level of approximation, or to approximate hedging strategies, as will be shown in the following chapters.

To summarize the task ahead, we recall that our starting point is the following class of stochastic volatility models for the underlying under the real-world measure:

$$\left.\begin{aligned} dX_t &= \mu X_t dt + f(Y_t, Z_t) X_t \, dW_t^{(0)}, \\ dY_t &= \tfrac{1}{\varepsilon}\alpha(Y_t) dt + \tfrac{1}{\sqrt{\varepsilon}}\beta(Y_t)\, dW_t^{(1)}, \\ dZ_t &= \delta c(Z_t) dt + \sqrt{\delta}\; g(Z_t)\, dW_t^{(2)}, \end{aligned}\right\} \tag{5.1}$$

as introduced in Section 3.6. For the derivative pricing problem associated with the underlying X, we consider these models under a risk-neutral measure $\mathbb{P}^\star$:

$$\left.\begin{aligned} dX_t &= r X_t dt + f(Y_t, Z_t) X_t \, dW_t^{(0)\star}, \\ dY_t &= \Big(\tfrac{1}{\varepsilon}\alpha(Y_t) - \tfrac{1}{\sqrt{\varepsilon}}\beta(Y_t)\Lambda_1(Y_t, Z_t)\Big) dt + \tfrac{1}{\sqrt{\varepsilon}}\beta(Y_t)\, dW_t^{(1)\star}, \\ dZ_t &= \Big(\delta\; c(Z_t) - \sqrt{\delta}\; g(Z_t)\Lambda_2(Y_t, Z_t)\Big) dt + \sqrt{\delta}\; g(Z_t)\, dW_t^{(2)\star}, \end{aligned}\right\} \tag{5.2}$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

where Λ_1 and Λ_2 are market prices of volatility risk. We seek the price $P^{\varepsilon,\delta}$ of a European derivative with payoff $h(X_T)$ on expiration date T, which can be expressed as the risk-neutral expected discounted payoff:

$$P^{\varepsilon,\delta}(t,X_t,Y_t,Z_t) = \mathbb{E}^\star\left\{e^{-r(T-t)}h(X_T) \mid X_t,Y_t,Z_t\right\}, \tag{5.3}$$

where the function $P^{\varepsilon,\delta}(t,x,y,z)$ solves the partial differential equation (4.5) involving as coefficients the functions appearing in the model (5.2). Notice from (5.3) that pricing requires the current values of the volatility factors Y and Z, which are not directly observed, in addition to calibration of the model (5.2). Calibration to historical data and options data involves choosing and fitting the functions $f,\alpha,\beta,c,g,\Lambda_1,\Lambda_2$. This is computationally intensive and requires good initial guesses in order that the fitted parameters are realistic representations of volatility behavior. The perturbation theory developed in the previous chapter vastly simplifies this calibration problem. As stated in Theorem 4.11, for nice payoff functions h,

$$P^\star = P^\star_{BS} + (T-t)\left(V_0^\delta \frac{\partial P^\star_{BS}}{\partial \sigma} + V_1^\delta D_1\left(\frac{\partial P^\star_{BS}}{\partial \sigma}\right) + V_3^\varepsilon D_1 D_2 P^\star_{BS}\right) \tag{5.4}$$

is a price approximation with accuracy $P^{\varepsilon,\delta} = P^\star + O(\varepsilon+\delta)$. Observe that only the *group market parameters*

$$(\sigma^\star, V_0^\delta, V_1^\delta, V_3^\varepsilon) \tag{5.5}$$

are needed to compute the approximate price $P^\star$ (here, for simplicity, we omit the dependence on z). In this chapter, we show how to calibrate these parameters from the term structure of implied volatilities obtained from European call options.

In later chapters, we show that these same parameters are exactly those needed to price more complicated instruments such as American options or path-dependent options to the same level of accuracy. Thus, in the regime where ε and δ are small, corresponding to stochastic volatility models with fast and slow volatility factors, we can handle a wide class of pricing problems which would be very complicated without the perturbation theory.

5.1 Approximate Call Prices and Implied Volatilities

We consider a European call option with strike K and maturity T. The corresponding payoff function is $h(x) = (x-K)^+$ and the Black–Scholes price (1.3.4) at volatility $\sigma^\star$ is

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

$$P^\star_{BS} = xN(d^\star_1) - Ke^{-r\tau}N(d^\star_2), \tag{5.6}$$

where we use the notation $\tau = T - t$, and

$$d^\star_{1,2} = \frac{\log(x/K) + \left(r \pm \frac{1}{2}\sigma^{\star 2}\right)\tau}{\sigma^\star\sqrt{\tau}}. \tag{5.7}$$

We will now use formula (5.4) in the case where $P^\star_{BS}$ is given by (5.6). Notice that in this case the payoff function h is not smooth. However, it is continuous and piecewise smooth, and we refer to the remark at the end of Section 4.5 regarding the accuracy of approximation in this case.

It is convenient to rewrite (5.4) as

$$P^\star = P^\star_{BS} + \left\{ \tau V_0^\delta + \tau V_1^\delta D_1 + \frac{V_3^\varepsilon}{\sigma^\star} D_1 \right\} \frac{\partial P^\star_{BS}}{\partial \sigma}, \tag{5.8}$$

where we have used the relation

$$\frac{\partial P^\star_{BS}}{\partial \sigma} = \tau\sigma^\star D_2 P^\star_{BS}$$

between the *Vega* and the *Gamma*, and where $P^\star_{BS}$ and $\partial P^\star_{BS}/\partial\sigma$ are evaluated at $\sigma = \sigma^\star$. Recall also from Chapter 1 that in the case of a call, *Vega* is explicitly given by

$$\frac{\partial C_{BS}}{\partial \sigma} = \frac{x\sqrt{\tau}\, e^{-d_1^2/2}}{\sqrt{2\pi}}. \tag{5.9}$$

As explained in Section 2.1, we convert the price $P^\star$ to an implied volatility I defined by

$$C_{BS}(I) = P^\star, \tag{5.10}$$

where the left-hand side is the Black–Scholes call price at the volatility level I, and the right-hand side is the call price (5.3) that follows from the fully specified stochastic volatility model. It follows from (5.9) that we can invert for this (unique) implied volatility.

The next step consists of expanding the difference $I - \sigma^\star$ between the implied volatility I and the volatility used to compute $P^\star_{BS}$, in powers of $\sqrt{\varepsilon}$ and $\sqrt{\delta}$:

$$I - \sigma^\star = \sqrt{\varepsilon}\, I_{1,0} + \sqrt{\delta}\, I_{0,1} + \cdots, \tag{5.11}$$

where the argument below shows that indeed there is no order one term on the right-hand side. We then expand both sides of equation (5.10) by using (5.8) and (5.11):

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

$$C_{BS}(\sigma^\star) + (\sqrt{\varepsilon} I_{1,0} + \sqrt{\delta} I_{0,1}) \frac{\partial C_{BS}(\sigma^\star)}{\partial \sigma} + \cdots =$$
$$P^\star_{BS} + \left\{ \tau V_0^\delta + \tau V_1^\delta D_1 + \frac{V_3^\varepsilon}{\sigma^\star} D_1 \right\} \frac{\partial P^\star_{BS}}{\partial \sigma} + \cdots .$$

From its definition, we know that $P^\star_{BS} = C_{BS}(\sigma^\star)$, and we can now match separately the $\sqrt{\varepsilon}$ terms and the $\sqrt{\delta}$ terms:

$$\sqrt{\varepsilon} I_{1,0} \frac{\partial C_{BS}(\sigma^\star)}{\partial \sigma} = \frac{V_3^\varepsilon}{\sigma^\star} D_1 \frac{\partial P^\star_{BS}}{\partial \sigma}, \tag{5.12}$$
$$\sqrt{\delta} I_{0,1} \frac{\partial C_{BS}(\sigma^\star)}{\partial \sigma} = \left\{ \tau V_0^\delta + \tau V_1^\delta D_1 \right\} \frac{\partial P^\star_{BS}}{\partial \sigma}, \tag{5.13}$$

where we recall that V_3^ε is proportional to $\sqrt{\varepsilon}$, and V_0^δ and V_1^δ are proportional to $\sqrt{\delta}$.

By differentiating the expression for *Vega* in (5.9) with respect to x, we deduce

$$D_1 \frac{\partial C_{BS}}{\partial \sigma} = \left(1 - \frac{d_1}{\sigma\sqrt{\tau}} \right) \frac{\partial C_{BS}}{\partial \sigma}. \tag{5.14}$$

From the relations (5.12), (5.13), and (5.14) evaluated at $\sigma = \sigma^\star$, we find

$$\begin{aligned} \sqrt{\varepsilon} I_{1,0} &= \frac{V_3^\varepsilon}{\sigma^\star} \left(1 - \frac{d_1^\star}{\sigma^\star\sqrt{\tau}} \right) \\ &= \frac{V_3^\varepsilon}{2\sigma^\star} \left(1 - \frac{2r}{\sigma^{\star 2}} \right) + \left(\frac{V_3^\varepsilon}{\sigma^{\star 3}} \right) \frac{\log(K/x)}{\tau}, \end{aligned} \tag{5.15}$$
$$\begin{aligned} \sqrt{\delta} I_{0,1} &= \tau V_0^\delta + \tau V_1^\delta \left(1 - \frac{d_1^\star}{\sigma^\star\sqrt{\tau}} \right) \\ &= \tau \left\{ V_0^\delta + \frac{V_1^\delta}{2} \left(1 - \frac{2r}{\sigma^{\star 2}} \right) \right\} + \left(\frac{V_1^\delta}{\sigma^{\star 2}} \right) \log(K/x). \end{aligned} \tag{5.16}$$

In terms of the reduced variable *log-moneyness to maturity ratio* defined by

$$\text{LMMR} = \frac{\log(K/x)}{T-t}, \tag{5.17}$$

the first-order approximation for the implied volatility,

$$I \approx \sigma^\star + \sqrt{\varepsilon} I_{1,0} + \sqrt{\delta} I_{0,1},$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

takes the simple form

$$I \approx b^\star + \tau b^\delta + \left(a^\varepsilon + \tau a^\delta\right) \mathrm{LMMR}, \tag{5.18}$$

where the parameters $(b^\star, b^\delta, a^\varepsilon, a^\delta)$ are defined in terms of $(\sigma^\star, V_0^\delta, V_1^\delta, V_3^\varepsilon)$ by

$$b^\star = \sigma^\star + \frac{V_3^\varepsilon}{2\sigma^\star}\left(1 - \frac{2r}{\sigma^{\star 2}}\right), a^\varepsilon = \frac{V_3^\varepsilon}{\sigma^{\star 3}}, \tag{5.19}$$

$$b^\delta = V_0^\delta + \frac{V_1^\delta}{2}\left(1 - \frac{2r}{\sigma^{\star 2}}\right), a^\delta = \frac{V_1^\delta}{\sigma^{\star 2}}. \tag{5.20}$$

In Figure 5.1, we show an example of an implied volatility surface predicted by this model. Notice the negative slope in the strike variable which becomes progressively more pronounced for short maturities. For fixed maturity the implied volatility is actually affine in the log-moneyness, $\log(K/x)$, for the approximation (5.18).

Since V_3^ε is of order $\sqrt{\varepsilon}$ whereas V_0^δ and V_1^δ are both of order $\sqrt{\delta}$, the relations (5.19) and (5.20) show that the intercept $b^\star$ is of order one, the slope coefficients a^ε and a^δ are of order $\sqrt{\varepsilon}$ and $\sqrt{\delta}$, respectively, and b^δ is of order $\sqrt{\delta}$.

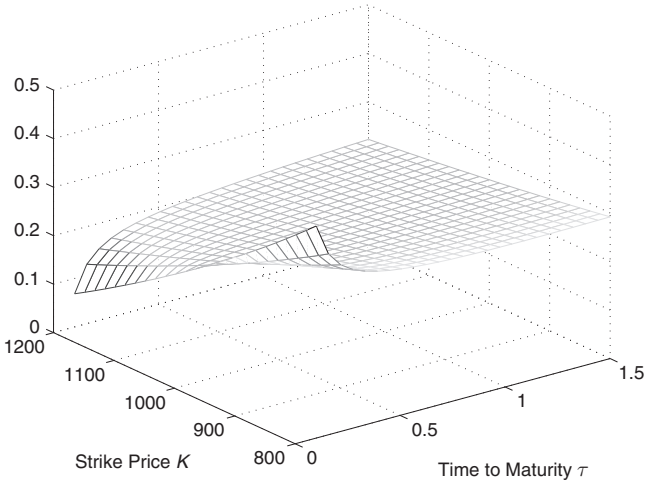

Figure 5.1 The implied volatility surface approximation described in (5.18) for $(a^\varepsilon, a^\delta, b^\delta, b^*) = (-0.0791, -0.1183, 0.0141, 0.2328)$. For a fixed time to maturity the surface is affine in $\log(K/x)$.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

In practice, formula (5.18) will be fitted to the observed term structure of implied volatility, as we discuss in detail in the following sections. From the fitted parameters $(a^\varepsilon, b^\star, a^\delta, b^\delta)$, we are able to calibrate the market group parameters $(\sigma^\star, V_0^\delta, V_1^\delta, V_3^\varepsilon)$ by inverting the relations above, retaining only the terms up to order $\sqrt{\varepsilon}$ and $\sqrt{\delta}$. From the first equation in (5.19), we deduce the following quadratic equation for $\sigma^\star$:

$$\frac{a^\varepsilon \sigma^{\star 2}}{2} + \sigma^\star - (b^\star + r a^\varepsilon) = 0.$$

Note that we have $\sigma^\star = b^\star + \mathcal{O}(\sqrt{\varepsilon})$ and $a^\varepsilon = \mathcal{O}(\sqrt{\varepsilon})$, therefore, to the same order of approximation:

$$\frac{a^\varepsilon b^{\star 2}}{2} + \sigma^\star - (b^\star + r a^\varepsilon) = 0.$$

We then find

$$\sigma^\star = b^\star + a^\varepsilon \left(r - \frac{b^{\star 2}}{2} \right).$$

From (5.20) it follows that the parameter V_3^ε can be written

$$V_3^\varepsilon = a^\varepsilon \sigma^{\star 3} = a^\varepsilon b^{\star 3} + \cdots,$$

where again only terms up to order $\sqrt{\varepsilon}$ are shown. The derivation of expressions for V_0^δ and V_1^δ from (5.20) is similar.

To the accuracy of our first-order approximation, the *calibration formulas* are:

$$\sigma^\star = b^\star + a^\varepsilon \left(r - \frac{b^{\star 2}}{2} \right), V_3^\varepsilon = a^\varepsilon b^{\star 3}, \tag{5.21}$$

$$V_0^\delta = b^\delta + a^\delta \left(r - \frac{b^{\star 2}}{2} \right), V_1^\delta = a^\delta b^{\star 2}, \tag{5.22}$$

and hence all the parameters that are needed for pricing can be obtained from the fitting of the term structure of implied volatility.

Note that by the result of Proposition 4.5 at the end of Section 4.2.2, if f is nondecreasing in y then the slope a^ε has the sign of ρ_1 as expected. Similarly, using the remark at the end of Section 4.2.4, one sees that in general a^δ has the sign of ρ_2.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

5.1.1 Relation to Historical Data

We showed above how to get the volatility level $\sigma^\star$ from the implied volatility term structure. This volatility level was introduced in (4.62) by

$$\sigma^\star = \sqrt{\bar{\sigma}^2(z) + 2V_2^\varepsilon(z)}, \tag{5.23}$$

where $\bar{\sigma}$ is obtained by averaging the two-factor volatility with respect to the invariant distribution of the fast factor:

$$\bar{\sigma}^2 = \bar{\sigma}^2(z) = \left\langle f^2(\cdot, z) \right\rangle.$$

Note that this parameter is therefore evaluated at the current level z of the slow volatility factor. The volatility $\bar{\sigma}$ can be obtained by constructing an estimator that computes averages of squared returns data over a time period that is short compared to the slow factor and long compared to the fast factor. The main advantage of constructing the price approximation relative to the volatility level $\sigma^\star$ is that we avoid the complications associated with constructing such an estimator and can rely on the term structure only. From (5.19) and (5.23), we find that $b^\star$ can be expressed by

$$b^\star = \bar{\sigma} + \frac{V_2^\varepsilon}{\bar{\sigma}} + \frac{V_3^\varepsilon}{2\bar{\sigma}}\left(1 - \frac{2r}{\bar{\sigma}^2}\right) + \cdots,$$

where only terms up to order $\sqrt{\varepsilon}$ are shown. It follows that $b^\star$ is a correction to $\bar{\sigma}$; the corrections due to V_2^ε and V_3^ε come from the market price of risk and from the leverage effect associated with the correlation ρ_1, respectively.

5.2 Calibration Procedure

In this section, we discuss the practical details of calibration of parameters to market implied volatility data. On a given day t, we have a finite set of implied volatilities $\{I(T_i, K_{ij})\}$ which are parameterized by expiration dates T_i and strike prices K_{ij}. Notice that options of different maturities may, and typically do, have different strikes. Clearly there is no unique way to fit the formula (5.18) to the usually large dataset to recover estimates of the four parameters $(a^\varepsilon, b^\star, a^\delta, b^\delta)$. However, the data is sparse in the T direction as compared with the K direction and our experience is that it is best to first fit the skew maturity by maturity and then across the term structure, as we now explain.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

For fixed T_i, we linearly regress the implied volatilities for strikes $(K_{ij})_j$ on the corresponding LMMRs

$$(\text{LMMR})_{ij} = \frac{\log(K_{ij}/x)}{\tau_i},$$

where x is the present value of the underlying and $\tau_i = T_i - t$, the time to maturity for this subset of options. We use the least-squares criterion to obtain the estimates $\hat{a}_i$ and $\hat{b}_i$ which solve

$$\min_{a_i,b_i} \sum_j \left(I(T_i,K_{ij}) - (a_i(\text{LMMR})_{ij} + b_i)\right)^2. \tag{5.24}$$

Then we regress these estimates on affine functions of τ_i to obtain estimates of the intercept $\hat{a}^\varepsilon$ and slope $\hat{a}^\delta$ which solve

$$\min_{a_0,a_1} \sum_i \left\{\hat{a}_i - (a_0 + a_1\tau_i)\right\}^2, \tag{5.25}$$

and the intercept $\hat{b}^\star$ and slope $\hat{b}^\delta$ which solve

$$\min_{b_0,b_1} \sum_i \left\{\hat{b}_i - (b_0 + b_1\tau_i)\right\}^2.$$

This two-step fitting procedure gives us estimates of the four group market parameters we need. Some key issues that we comment on in the context of S&P 500 data below are goodness-of-fit and stability over time of the estimated parameters.

5.3 Illustration with S&P 500 Data

We illustrate the fitting procedure using S&P 500 options data obtained from the Wharton Research Data Service (WRDS) database. Our dataset contains implied volatilities from closing S&P 500 European option prices. However, as discussed in Figlewski (2010, section 4), the data needs some treatment to remove inconsistencies and unreliable entries. These include very low priced out-of-the-money contracts where bid–ask spreads are more than 100% of the midpoint price; quotes existing for deep out-of-the-money or in-the-money options which have not been trading recently; and a pronounced "jump" in the implied volatility curve at the point where implied volatilities switch from being computed from put option prices to call option prices. We roughly follow the data-cleaning procedure described in Figlewski (2010), with some modifications, and describe its impact on the test case of April 19, 2005 data.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

5.3.1 Data Cleaning

From the WRDS database, we extract daily closing bid and ask quotes for European options on the S&P 500 index, along with each option's strike and expiration date and the WRDS-computed implied volatility. On the illustration date April 19, 2005, there are originally 566 option entries.

- We first filter out datapoints which have
  - bid quotes less than \$0.50 (this removes 70 entries on the illustration date);
  - no implied volatility value (this removes a further 90);
  - incorrect expiration dates (an occasional problem, but none on the test date).
- In the WRDS data there is a jump between the implied volatility curves coming from calls or puts with the same maturity. To smooth this out, following the general procedure described in Figlewski (2010), we blend the implied volatility values for options with strikes between certain cutoffs L and H, with $L < H$.
  - *Choosing cutoffs*: For a given maturity, let $\mathscr{K}$ be the set of strikes for which both call and put implied volatilities are available, and let x denote the day's closing level of the S&P 500 index. We set $L = \max(0.85x, \min(\mathscr{K}))$ and $H = \min(1.15x, \max(\mathscr{K}))$. Essentially we set the cutoffs to $+/-15\%$ of at-the-money, and then modify them if necessary since we can only blend when we have a put–call pair. Occasionally there is still a handful of strikes between L and H for which only one type of option is traded, in which case these unpaired values are discarded since they cannot be blended.
  - We discard *puts* with strikes above H and *calls* with strikes below L, that is those that are too far *in*-the-money to be sufficiently liquid. On the test date, this removes 63 datapoints.
  - *Blending*: For each maturity and strike $K \in (L,H)$, let $I_p(K)$ and $I_c(K)$ denote the put and call implied volatilities in the database where they both exist. With $w = w(K) = (H-K)/(H-L)$, we set the implied volatility value $I(K)$ for that strike to be

$$I(K) = wI_p(K) + (1-w)I_c(K).$$

    On the test date, there are 260 options in the blending region, one put and one call for each strike. Blending halves this number and therefore reduces the data by a further 130 points. So finally, there are 213 datapoints left on April 19, 2005, which is 37.6% of the original

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

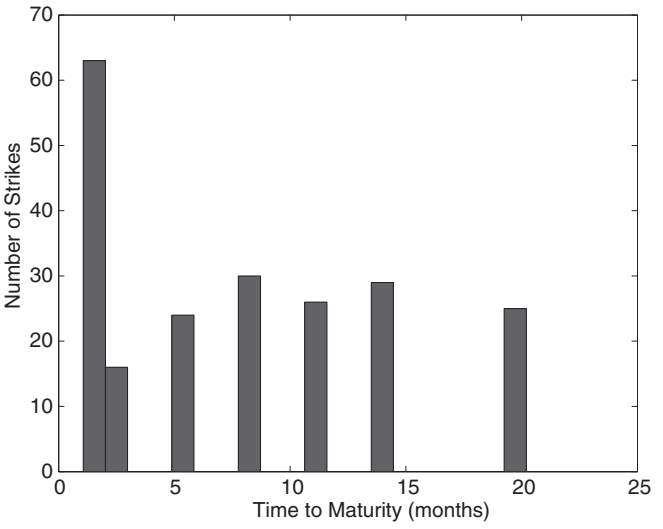

Figure 5.2 Number of strikes for each maturity: S&P 500 options on April 19, 2005.

data size for that day. Figure 5.2 shows the number of strikes for each maturity.

– In summary: for strikes $\leq L$, we use only puts, for strikes $\geq H$, we use only calls, and for strikes between L and H, we blend the two.

- For the month of April 2005, there are 12,173 original datapoints and we remove about 14% that have low bid prices, 19% where no implied volatility is given, and 10% that are too far in-the-money. The final number of datapoints after blending is 4339, which is 35.6% of the original.

We next proceed to examine the fits of the asymptotic formulas on the data.

5.3.2 *Fast Volatility Factor Only*

We first look at the performance using the fast scale theory only, ignoring the slow scale. Figure 5.3 shows the fit using only the fast-factor approximation

$$I \approx b^\star + a^\varepsilon(\text{LMMR}),$$

which corresponds to assuming $\delta = 0$. Each strand in Figure 5.3 comes from options of different maturities (with the shortest maturities on the leftmost strand, and the maturity increasing going clockwise). Clearly the

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

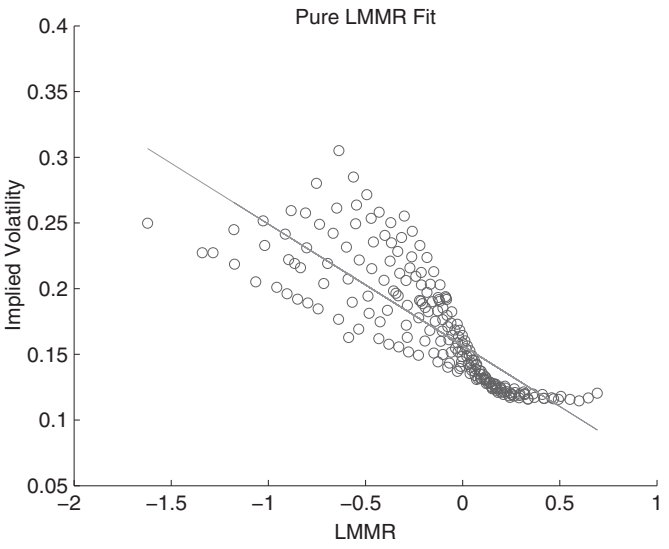

Figure 5.3 S&P 500 implied volatilities as a function of LMMR on April 19, 2005. The circles are from S&P 500 data, and the line $b^\star + a^\varepsilon(LMMR)$ shows the result using the estimated parameters from only an LMMR (fast factor) fit.

single-factor theory struggles to capture the range of maturities and so runs through the middle.

5.3.3 Slow Volatility Factor Only

In Figure 5.4, we show the result of the calibration using only the slow-factor approximation, which corresponds to assuming $\varepsilon = 0$:

$$I \approx \bar{\sigma} + b^\delta \tau + a^\delta(\mathrm{LM}),$$

where we denote $\mathrm{LM} = \tau\mathrm{LMMR}$, the log-moneyness. We plot the maturity adjusted implied volatility defined by $I - b^\delta \tau$. The fit as a function of the regressor LM is shown in Figure 5.4, with the maturities increase going counterclockwise from the top-leftmost strand. Again, the single-factor theory struggles to capture the range of maturities.

5.3.4 Fast and Slow Volatility Factors

Finally, in fitting the two-factor volatility approximation (5.18), we follow the two-step procedure of fitting the skew to obtain $\hat{a}_i$ and $\hat{b}_i$ for different maturities τ_i. These are then fitted to an affine functions of τ to give estimates of a^ε, $b^\star$, a^δ, and b^δ. A plot of this second term-structure fit is

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

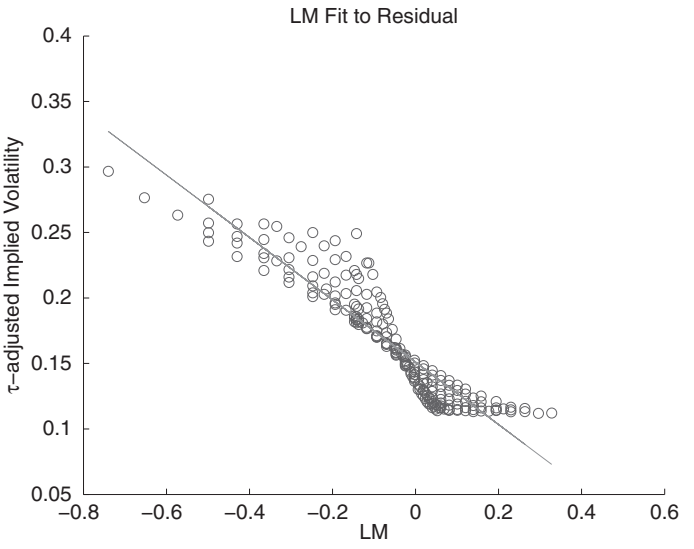

Figure 5.4 τ-adjusted implied volatility $I - b^{\delta}\tau$ as a function of LM. The circles are from S&P 500 data on April 19, 2005, and the line $\sigma^{\star} + a^{\delta}(LM)$ shows the fit using the estimated parameters from only a slow factor fit.

shown in Figure 5.5. Observe that in the regime where our approximation is valid the parameters $a^{\varepsilon}, a^{\delta}$, and b^{δ} are expected to be small, while $b^{\star}$ is the leading-order magnitude of volatility: this is also what we see in Figure 5.5.

In Figure 5.6, we show the calibrated multiscale approximation (5.18) to all the data on April 19, 2005. We see the ability to capture the range of maturities is much improved. The largest misfitting is at the level of the shortest maturities (the leftmost strand). One way to handle this discrepancy uses a periodic modulation of the fast time scale with period corresponding to the monthly expiration cycles of traded options, and we discuss this in Section 5.4. The rightmost *turn of the skew* also comes from the shortest maturity and can be captured using the second-order asymptotic theory derived in Section 9.3 and implemented in this chapter in Section 5.5.2.

The picture remains largely similar in the aftermath of the financial crisis. Figure 5.7 shows the analogs of Figures 5.5 and 5.6 using data from September 19, 2009.

5.3.5 Parameter Stability

A crucial feature of a model is *time stability* of its fitted parameters, since this time stability means that the model captures well the aspects of the dynamics of the underlying, which is essential to ensure consistency when

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

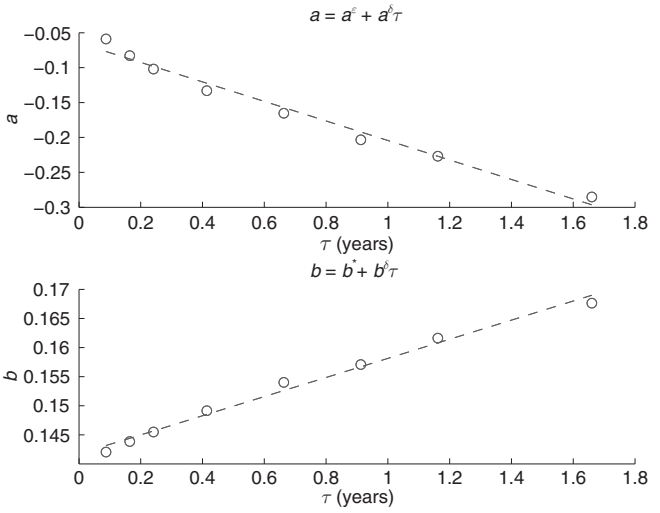

Figure 5.5 Term-structure fits on April 19, 2005. The circles in the top plot are the slope coefficients $\hat{a}_i$ of LMMR fitted in the first step of the regression. The solid line is the straight line $(a^{\varepsilon} + a^{\delta}\tau)$ fitted in the second step of the regression. The bottom plot shows the corresponding picture for the skew intercepts $\hat{b}_i$ fitted to the straight line $b^{\star} + b^{\delta}\tau$. The estimates are $(a^{\varepsilon}, a^{\delta}, b^{\delta}, b^{*}) = (-0.0646, -0.1397, 0.0164, 0.1417)$.

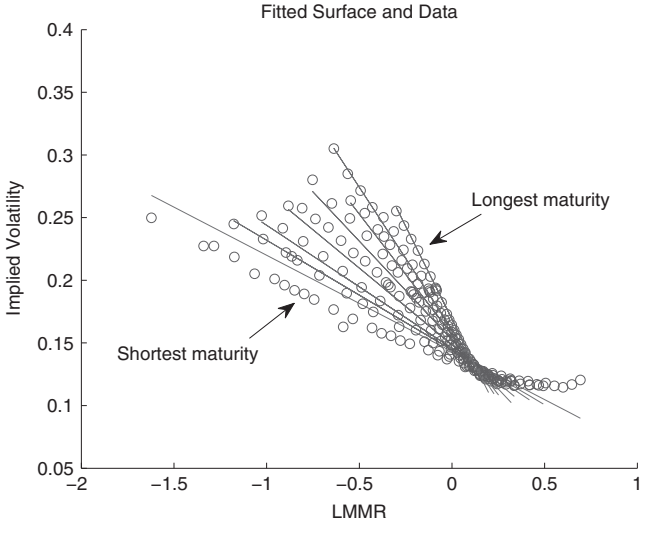

Figure 5.6 Data and calibrated fit on April 19, 2005. The circles are from S&P 500 data, and the lines are the formula (5.18) using $(a^{\varepsilon}, b^{\star}, a^{\delta}, b^{\delta})$ estimated from the full fast and slow factor fit. The average *relative* fitting error is 3.75%.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

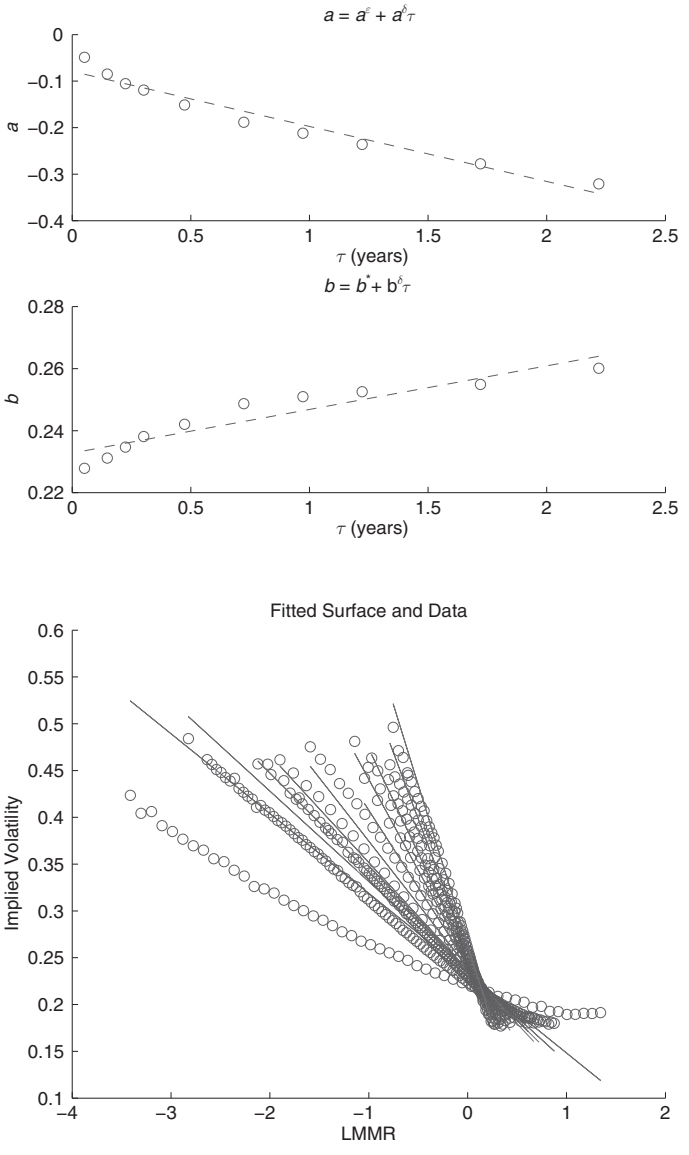

Figure 5.7 Term-structure fits (top) and calibrated fit (bottom) on September 28, 2009. The estimates are $(a^{\varepsilon}, a^{\delta}, b^{\delta}, b^{*}) = (-0.0791, -0.1183, 0.0141, 0.2328)$ and the average *relative* fitting error is 4.65%.

pricing and hedging path-dependent options, for example. Practically, this means that we are looking at the time evolution (with respect to t) of the estimated parameters $(a^{\varepsilon}, b^{\star}, a^{\delta}, b^{\delta})$. This is shown in Figure 5.8 over the period January 2000 through October 2009, using S&P 500 options data, but excluding the shortest maturity. We show also the corresponding

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

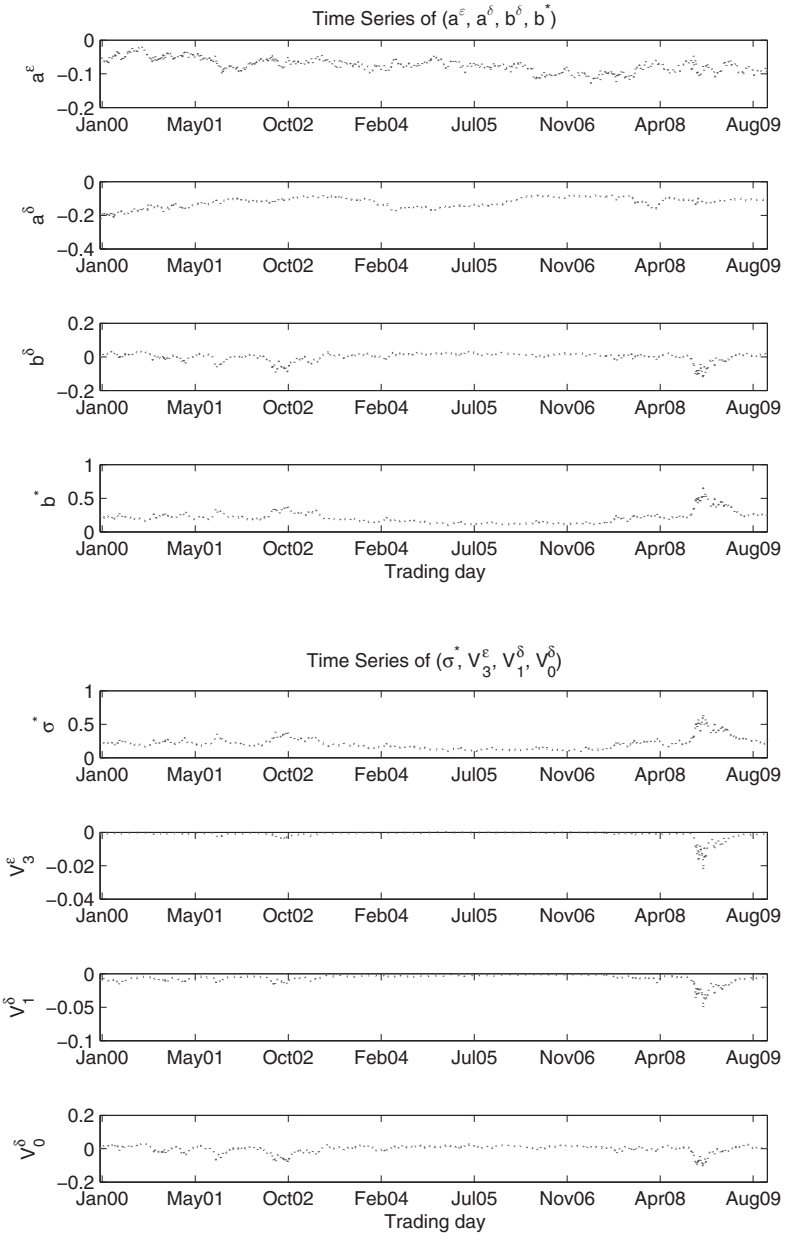

Figure 5.8 Stability of fitted parameters. The top picture shows daily calibrated values of $(a^\varepsilon, b^\star, a^\delta, b^\delta)$, with the corresponding $(\sigma^\star, V_0^\delta, V_1^\delta, V_3^\varepsilon)$ in the bottom picture.

plots for the market group parameters $(\sigma^\star, V_0^\delta, V_1^\delta, V_3^\varepsilon)$ computed from the calibration formulas (5.21)–(5.22).

In these, one sees clearly the financial crisis that erupted in September 2008: in particular, its impact is felt more in $(b^\star, b^\delta)$ than in $(a^\varepsilon, a^\delta)$,

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

indicating a more dramatic jump in implied volatility levels rather than skew slope.

The mean values of the group parameters over this period are:

$$\sigma^\star = 0.2054, \qquad V_0^\delta = 0.0008, \qquad V_1^\delta = -0.0059, \qquad V_3^\varepsilon = -0.0010.$$

The first gives an indication of the "average" S&P 500 volatility level of 20%. The other parameters, the V's, are small as the asymptotic theory demands.

5.4 Maturity Cycles

In the previous calibration we did not use the shortest maturity options, which have the steepest skew. When this maturity is included, an interesting periodicity reveals itself in the fitted slopes of the skew. To see this clearly, we consider for the moment just a fast-factor fit to the first three maturities. We keep the shortest maturity option until three days to expiration, in general the data comprises one-month, two-month, and three-month options, which we fit to the formula

$$I \approx b^\star + a^\varepsilon \mathrm{LMMR}$$

to obtain daily slope estimates a^ε. Figure 5.9 shows these through the year 2007. The ticks on the graph correspond to the expiration dates of S&P 500 options, which is the third Friday of each month. There is clearly a periodic phenomenon coinciding with the expiration dates. These dates (which are known in advance) therefore seem to have some bearing on the evolution of the implied volatility surface. This pattern contains information about

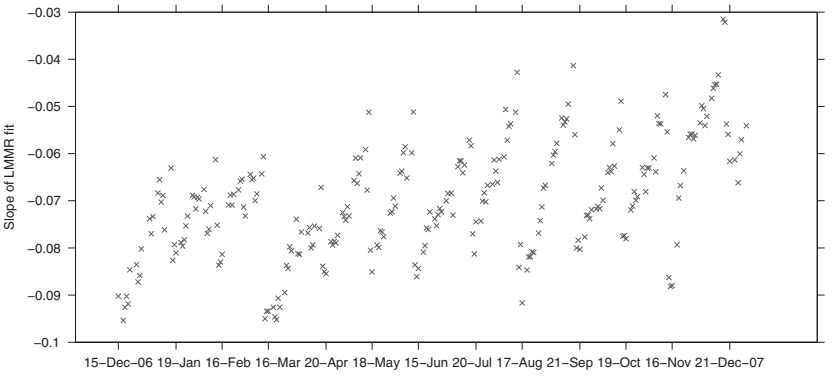

Figure 5.9 Fitted skew slopes a^ε through 2007 using only the first three maturities. The ticks on the graph correspond to the expiration dates of S&P 500 options.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

the systematic behavior of market option prices which is not easy to see directly from implied volatility data, but is revealed clearly once the regression with respect to LMMR has been carried out to filter out stochastic volatility effects.

The "jumps" in the fitted skew slope, when the cycle returns to the beginning (it is not continuous), coincide with the closest-to-maturity options disappearing from the data. The fact that some options are expiring has a mild "feedback effect" on the entire options market (with the common underlying), probably through the volatility of the underlying itself. When those options disappear, the shortest length of options in the data jumps from near zero to about 30 days. As a consequence of there not being a very steep skew visible on that day, the estimated skew slope becomes much less negative as the next expiration date is suddenly much further away.

Rather than proposing a mechanism that might explain this, we model the empirical observation directly through a periodic variation in the *local speed* of the fast process driving volatility: it is still fast, but beomes faster periodically leading up to the expiration dates. In the following sections we present the theoretical aspects of this extension and in Section 5.4.5 we illustrate that this greatly improves the stability of the parameter fitting. The "calendar effect" we observe in S&P 500 options is important primarily for options expiring on the next few maturity dates. This cycle effect may not be present in different markets with different expiration date structures, e.g. OTC markets where there is an almost continuous range of maturity dates.

5.4.1 Perturbation with Time-Varying Coefficients

We outline the extension of the asymptotic theory of Chapter 4 to a specific instance of time-dependent parameters in the fast volatility process. Here we choose to introduce variation in the time scale coefficient ε. For our data fitting below, we will make this time scale dependent on the time until the next maturity date, and we will find that the periodic pattern seen in Figure 5.9 is consistent with a local mean reversion speed that increases as we approach each expiration date, and then drops back down until the next one is neared.

Our model under the risk-neutral measure becomes

$$dX_t = rX_t\,dt + f(Y_t, Z_t)X_t\,dW_t^{(0)\star}, \tag{5.26}$$

$$dY_t = \left(\frac{1}{\varepsilon\nu(t)}\alpha(Y_t) - \frac{1}{\sqrt{\varepsilon\nu(t)}}\beta(Y_t)\Lambda_1(Y_t)\right)dt + \frac{1}{\sqrt{\varepsilon\nu(t)}}\beta(Y_t)\,dW_t^{(1)\star},$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

$$dZ_t = \Big(\delta c(Z_t) - \sqrt{\delta}\; g(Z_t)\Lambda_2(Y_t, Z_t)\Big)\, dt + \sqrt{\delta}\; g(Z_t)\, dW_t^{(2)\star}, \tag{5.27}$$

where we have introduced the time-dependent scaling factor $\nu(t)$ into the fast factor Y. Typically this factor will become small as t approaches a maturity date. As usual, the standard Brownian motions $\left(W_t^{(0)\star}, W_t^{(1)\star}\right)$ are correlated according to the following cross-variation:

$$d\langle W^{(0)\star}, W^{(1)\star}\rangle_t = \rho_1 dt, \tag{5.28}$$

where $|\rho_1| < 1$. The volatility of the underlying asset X is $f(Y_t, Z_t)$, where f is a positive function satisfying the conditions in Theorem 4.10. Other aspects of the model are described in more detail in Section 4.1. The special case with $\nu(t) = 1$ corresponds to the time-independent theory discussed in Chapter 4.

As before, we seek an approximation for option prices in the regime where ε and δ are small. The price of a European derivative is given by

$$P^\varepsilon(t, X_t, Y_t, Z_t) = \mathbb{E}^\star\left\{e^{-r(T-t)}h(X_T) \mid X_t, Y_t, Z_t\right\}, \tag{5.29}$$

and is also characterized as the solution of the parabolic partial differential equation

$$\begin{aligned} \mathscr{L}^{\varepsilon,\delta}(t)P^\varepsilon &= 0, \\ P^\varepsilon(T, x, y) &= h(x), \end{aligned} \tag{5.30}$$

where the operator $\mathscr{L}^{\varepsilon,\delta}(t)$ can be expressed as a sum of components:

$$\mathscr{L}^\varepsilon(t) = \frac{1}{\varepsilon\nu(t)}\mathscr{L}_0 + \frac{1}{\sqrt{\varepsilon\nu(t)}}\mathscr{L}_1 + \mathscr{L}_2 + \sqrt{\delta}\mathscr{M}_1 + \delta\mathscr{M}_2 + \sqrt{\frac{\delta}{\varepsilon\nu(t)}}\mathscr{M}_3,$$

with the operators $\mathscr{L}_i$ being defined in (4.8), (4.9), and (4.10) and the $\mathscr{M}_i$ in (4.11), (4.12), and (4.13). Thus, we find that the problem described by (5.30) corresponds exactly to the one described by (4.5) with the replacements

$$\mathscr{L}_0 \longmapsto \frac{1}{\nu(t)}\mathscr{L}_0, \quad \mathscr{L}_1 \longmapsto \frac{1}{\sqrt{\nu(t)}}\mathscr{L}_1, \quad \mathscr{M}_3 \longmapsto \frac{1}{\sqrt{\nu(t)}}\mathscr{M}_3.$$

The argument presented in Section 4.2 can now be repeated with some slight modifications.

We first expand the price as in Section 4.2.1:

$$P^\varepsilon = P_0 + \varepsilon^{1/2}P_{1,0} + \delta^{1/2}P_{0,1} + \cdots.$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

With our usual definition of the constant effective volatility

$$\bar{\sigma}(z) = \sqrt{\langle f^2(\cdot,z)\rangle},$$

one can follow the lines of the derivation of the first-order approximation in Chapter 4, and obtain that the leading-order term P_0 is the Black–Scholes price P_{BS} which solves

$$\begin{aligned}\mathscr{L}_{BS}(\bar{\sigma}(z))P_{BS} &= 0,\\ P_{BS}(T,x) &= h(x).\end{aligned}$$

Next, one finds that the first correction $P^{\varepsilon}_{1,0} = \sqrt{\varepsilon}P_{1,0}$ is modified because the operator $\mathscr{A}^{\varepsilon}$ in (4.32) is replaced by

$$\begin{aligned}&\sqrt{\varepsilon}\left\langle \frac{1}{\sqrt{\nu(t)}}\mathscr{L}_1\left(\frac{1}{\nu(t)}\mathscr{L}_0\right)^{-1}(\mathscr{L}_2 - \langle\mathscr{L}_2\rangle)\right\rangle\\ &\quad = \sqrt{\varepsilon\nu(t)}\left\langle \mathscr{L}_1\mathscr{L}_0^{-1}(\mathscr{L}_2 - \langle\mathscr{L}_2\rangle)\right\rangle \;=\; \sqrt{\nu(t)}\mathscr{A}^{\varepsilon}.\end{aligned}$$

It follows that $P^{\varepsilon}_{1,0}$ solves the inhomogeneous problem

$$\begin{aligned}\mathscr{L}_{BS}(\bar{\sigma})P^{\varepsilon}_{1,0} &= -\sqrt{\nu(t)}\left(V^{\varepsilon}_2 D_2 P_{BS} + V^{\varepsilon}_3 D_1 D_2 P_{BS}\right),\\ P^{\varepsilon}_{1,0}(T,x) &= 0.\end{aligned}$$

Using the argument in Section 4.2.3, one can easily check that the solution is given explicitly in terms of P_{BS} by

$$P^{\varepsilon}_{1,0} = \left(\int_t^T \sqrt{\nu(s)}\,ds\right)\left(V^{\varepsilon}_2 D_2 P_{BS} + V^{\varepsilon}_3 D_1 D_2 P_{BS}\right).$$

We define now the time-averaged quantity

$$\overline{\nu^{1/2}_{t,T}} = \frac{1}{T-t}\int_t^T \sqrt{\nu(s)}\,ds.$$

It is straightforward to see that the slow scale correction is unaffected by the time-varying coefficient in the fast factor, and therefore the first-order approximation in (4.58) becomes

$$\begin{aligned}\tilde{P}^{\varepsilon,\delta} = P_{BS} + (T-t)\Bigg[&V^{\delta}_0(z)\frac{\partial}{\partial\sigma} + V^{\delta}_1(z)D_1\left(\frac{\partial}{\partial\sigma}\right) + \overline{\nu^{1/2}_{t,T}}\left(V^{\varepsilon}_2(z)D_2\right.\\ &\left.+V^{\varepsilon}_3(z)D_1D_2\right)\Bigg]P_{BS}.\end{aligned} \tag{5.31}$$

Clearly, the effect of the time-dependent parameter at this level is the replacement

$$V^{\varepsilon}_2(z) \mapsto \overline{\nu^{1/2}_{t,T}}\,V^{\varepsilon}_2(z), \qquad V^{\varepsilon}_3(z) \mapsto \overline{\nu^{1/2}_{t,T}}\,V^{\varepsilon}_3(z).$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

Following the parameter-reduction argument given in Section 4.3, one can deduce that the first-order price approximation takes the form

$$\tilde{P}^\star := P_{BS}(\sigma^\star_{t,T}) + (T-t)\left[V_0^\delta(z)\frac{\partial}{\partial\sigma} + V_1^\delta(z)D_1\left(\frac{\partial}{\partial\sigma}\right) + \overline{\nu_{t,T}^{1/2}}\,V_3^\varepsilon(z)D_1D_2\right]P_{BS}(\sigma^\star_{t,T}), \tag{5.32}$$

where the (time-dependent) corrected volatility $\sigma^\star_{t,T}$ is given by

$$\sigma^\star_{t,T} = \sqrt{\bar{\sigma}^2 + 2\,\overline{\nu_{t,T}^{1/2}}\,V_2^\varepsilon}. \tag{5.33}$$

This generalization involves the time-dependent Black–Scholes operator with deterministic time-varying square volatility $\bar{\sigma}^2 + 2\sqrt{\nu(t)}V_2^\varepsilon$, and the fact that the corresponding Black–Scholes price is obtained as the Black–Scholes price with time-averaged square volatility as explained in Section 2.2.1. The accuracy results given in Section 4.5 can be generalized to this time-dependent case and the price correction due to the slow factor is not affected by the time variation in the parameters to the accuracy of our approximation.

5.4.2 Calendar Time

We will choose $\nu(t)$ to be a power of the time to next maturity date. Let T_n denote the maturity dates (for example, the third Fridays of the month in the case of S&P 500 options). We define $n(t)$, the time to the next maturity, by

$$n(t) = \inf\{n : T_n - t \geq \Delta t\}, \tag{5.34}$$

with $\Delta t > 0$ a small cutoff which will prevent blowup of ν^{-1} at maturity dates. Then our maturity cycles model is

$$\nu(t) = c_p\left(T_{n(t)} - t\right)^p, \qquad p > 0, \tag{5.35}$$

where p is a parameter and c_p is a normalization constant to be chosen. We suppose that the option expiration dates are approximately ΔT apart (for example, one month). Then c_p is chosen so that

$$\frac{\int_0^{\Delta T}(\nu(s))^{1/2}\,ds}{\Delta T} = 1. \tag{5.36}$$

This gives

$$c_p^{1/2} = \frac{1}{(\Delta T)^{p/2}}\left(1 + \frac{p}{2}\right).$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

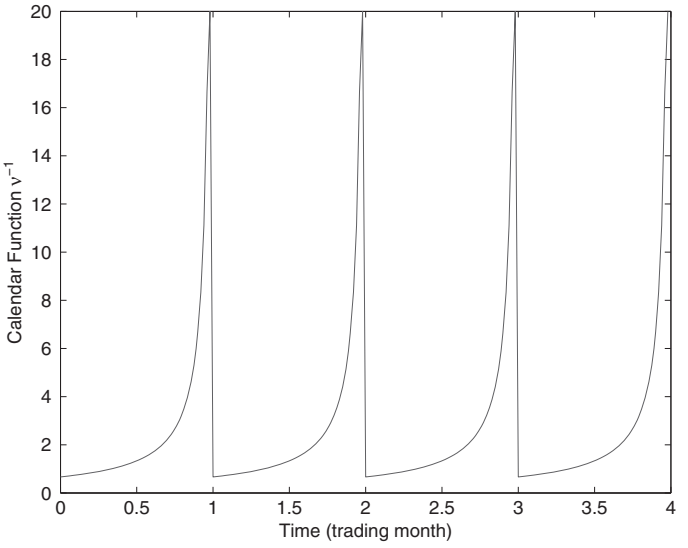

Figure 5.10 The calendar function $1/\nu(t)$ defined in (5.35), with $p = 1$, as a function of trading month.

The choice $p = 0$ corresponds to a constant rate of volatility mean-reversion, as we had in Chapter 4. The data presented in Section 5.4.5 is consistent with the choice $p = 1$. In Figure 5.10 we plot the "calendar function" ν^{-1}; in this case, it is small at the beginning of the trading month and becomes very large close to the expiration dates. A very large ν^{-1} corresponds to a very rapid mean-reversion rate of the Y process.

For large times to maturity the time-dependent case corresponds approximately to the constant parameter case since the monthly variation in the parameter averages out. To show this, assume that the period between the expiration dates ΔT is constant and that the time to maturity is decomposed as $T - t = m_0 \Delta T + \eta$ with m_0 an integer and $0 \leq \eta < \Delta T$. Also assume for simplicity the cutoff $\Delta t = 0$ in (5.34). We then find, in view of (5.35) and (5.36),

$$\begin{aligned}
\overline{\nu_{t,T}^{1/2}} &= \frac{1}{T-t}\int_t^T (\nu(s))^{1/2}\,ds = \frac{1}{T-t}\int_0^{m_0\Delta T+\eta} (\nu(T-s))^{1/2}\,ds \\
&= \frac{\sqrt{c_p}}{m_0\Delta T+\eta}\left[\int_t^{t+\eta}(t+\eta-s)^{p/2}\,ds + m_0\int_0^{\Delta T}(\Delta T-s)^{p/2}\,ds\right] \\
&= \frac{1}{(\Delta T)^{p/2}}\,\frac{\left(\eta^{1+\frac{p}{2}} + m_0(\Delta T)^{1+\frac{p}{2}}\right)}{m_0\Delta T+\eta}.
\end{aligned}$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

Therefore, when $T-t$ is within one ΔT of maturity and so $m_0 = 0$, we have

$$\overline{\nu_{t,T}^{1/2}} = \left(\frac{T-t}{\Delta T}\right)^{p/2},$$

and for longer maturities when $m_0 \Delta T >> \eta$,

$$\overline{\nu_{t,T}^{1/2}} \approx 1,$$

so we return to the time-independent theory.

5.4.3 Calibration with Time-Dependent Parameters

In this section we generalize the analysis presented in Section 5.1 to the case when the price approximation is given by (5.32). Thus, we consider a European call option with strike K and maturity T so that the payoff function is $h(x) = (x-K)^+$. The call price at the *constant* volatility level σ is denoted $C_{BS}(\sigma)$ and given by the Black–Scholes formula in (5.6) for the case $\sigma = \sigma^\star$. The implied volatility I associated with our stochastic volatility model solves

$$C_{BS}(I) = P^\varepsilon,$$

the volatility level implied by the Black–Scholes formula for the price given by the fully specified stochastic volatility model.

Consider first the case of *only* the fast volatility factor in the model. In the case of a call, our price approximation takes the form $P^\varepsilon \approx P_{BS}(\sigma_{t,T}^\star) + P_{1,0}^\varepsilon(\sigma_{t,T}^\star)$. We next identify the approximation for the implied volatility skew that follows from this first-order price approximation. As before, we expand as

$$I = \sigma_{t,T}^\star + \sqrt{\varepsilon} I_{1,0} + \cdots,$$

then the expression (5.15) generalizes to

$$\sqrt{\varepsilon} I_{1,0} = \left(\frac{V_3^\varepsilon}{2\sigma_{t,T}^\star}\left(1 - \frac{2r}{(\sigma_{t,T}^\star)^2}\right) + \left(\frac{V_3^\varepsilon}{(\sigma_{t,T}^\star)^3}\right)\frac{\log(K/x)}{\tau}\right)\overline{\nu_{t,T}^{1/2}},$$

and (5.18) to

$$I \approx b_{t,T}^\star + a_{t,T}^\star \,\mathrm{LMMR}\,\overline{\nu_{t,T}^{1/2}},$$

where

$$a_{t,T}^\varepsilon = \frac{V_3^\varepsilon}{(\sigma_{t,T}^\star)^3},$$

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

$$b^{\star}_{t,T} = \sigma^{\star}_{t,T} + \frac{V_3^{\varepsilon}}{2\sigma^{\star}_{t,T}}\left(1 - \frac{2r}{(\sigma^{\star}_{t,T})^2}\right)\overline{\nu^{1/2}_{t,T}}.$$

Recall that V_2^{ε} and V_3^{ε} are of order $\sqrt{\varepsilon}$. Moreover, that

$$\sigma^{\star} = \sqrt{\bar{\sigma}^2 + 2V_2^{\varepsilon}},$$
$$\sigma^{\star}_{t,T} = \sqrt{\bar{\sigma}^2 + 2\,\overline{\nu^{1/2}_{t,T}}\,V_2^{\varepsilon}},$$

so that we can write

$$a^{\varepsilon}_{t,T} = \frac{V_3^{\varepsilon}}{(\sigma^{\star})^3} + \mathscr{O}(\varepsilon) = a^{\varepsilon} + \mathscr{O}(\varepsilon), \tag{5.37}$$

$$b^{\star}_{t,T} = \bar{\sigma} + \frac{V_2^{\varepsilon}}{\bar{\sigma}}\overline{\nu^{1/2}_{t,T}} + \frac{V_3^{\varepsilon}}{2\sigma^{\star}}\left(1 - \frac{2r}{(\sigma^{\star})^2}\right)\overline{\nu^{1/2}_{t,T}} + \mathscr{O}(\varepsilon) \tag{5.38}$$

$$= \bar{\sigma} + (b^{\star} - \bar{\sigma})\overline{\nu^{1/2}_{t,T}} + \mathscr{O}(\varepsilon), \tag{5.39}$$

with a^{ε} and $b^{\star}$ defined in (5.19).

Recall that in the case where the slow factor is also present, the price approximation can be written in the form (5.32). Since the price correction that is due to the slow factor is not affected by the time dependence in the parameters, we find that the calibration formulas (5.21) and (5.22) for the group market parameters associated with the slow factor, V_0^{δ} and V_1^{δ}, remain unchanged. Now, incorporating back the slow factor leads to the following first-order approximation for the implied volatility:

$$\begin{aligned} I &\approx \bar{\sigma} + (b^{\star} - \bar{\sigma})\overline{\nu^{1/2}_{t,T}} + a^{\varepsilon}\,\mathrm{LMMR}\,\overline{\nu^{1/2}_{t,T}} + b^{\delta}\tau + a^{\delta}\mathrm{LM} \\ &= \bar{\sigma} + \Delta b\,\overline{\nu^{1/2}_{t,T}} + a^{\varepsilon}\,\mathrm{LMMR}_{\nu} + b^{\delta}\tau + a^{\delta}\mathrm{LM}, \end{aligned} \tag{5.40}$$

with

$$\mathrm{LMMR}_{\nu} = \mathrm{LMMR}\,\overline{\nu^{1/2}_{t,T}},$$
$$\Delta b = b^{\star} - \bar{\sigma}.$$

Notice that by observing the implied volatilities for various maturities and values for the moneyness we can estimate the parameters a^{ε}, $b^{\star}$, and $\bar{\sigma}$ (as well as b^{δ} and a^{δ}). To the accuracy of our first-order approximation we then obtain *calibration formulas* for V_2^{ε} and V_3^{ε} from (5.37), (5.38), and (5.39) by

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

$$\frac{V_2^\varepsilon}{\bar\sigma} = \Delta b + a^\varepsilon\left(r - \frac{b^{\star 2}}{2}\right), \quad V_3^\varepsilon = a^\varepsilon (b^\star)^3.$$

We assume that $\nu(t)$ is a known function, therefore these parameters are sufficient to obtain a first-order approximation for $\sigma^\star_{t,T}$. Thus, from the calibration of the implied volatility term structure we can identify the parameters that are needed to compute the price approximation in (5.32).

In the next sections we continue our example using S&P 500 data to illustrate some practical aspects of the calibration. As mentioned earlier, these data are consistent with the model (5.35) for $p = 1$. We therefore examine the form for the price and the implied volatility in this special case. In our discussion below we assume that the intervals between the maturity dates are constant and equal to ΔT. We have

$$\overline{\nu_{t,T}^{1/2}} = \begin{cases} \sqrt{\frac{T-t}{\Delta T}} & \text{for} \quad T-t < \Delta T, \\ 1 & \text{for} \quad T-t \gg \Delta T, \end{cases}$$

from which it follows that

$$\mathrm{LMMR}_\nu = \begin{cases} \frac{\log(K/x)}{\sqrt{(T-t)\Delta T}} & \text{for} \quad T-t < \Delta T, \\ \frac{\log(K/x)}{T-t} & \text{for} \quad T-t \gg \Delta T. \end{cases}$$

In fact, a common observation is that for fixed $K \neq x$ the observed term structure (variation with time to maturity $T-t$) behaves like $(T-t)^{-1/2}$ rather than $(T-t)^{-1}$. The time-dependent theory presented above provides a bridge to this observation and below we examine this transition in more detail by returning to the S&P 500 data, first looking at some practical aspects of the data fitting.

5.4.4 Practical Fitting with Time-Dependent Parameters

We will now fit the implied volatility surface to (5.40) to obtain estimates of $(a^\varepsilon, b^\star, \bar\sigma, b^\delta, a^\delta)$. As usual, on a given day t, we have a finite set of implied volatility $\{I(T_i, K_{ij})\}$ data which are parameterized by expiration dates T_i and strike prices K_{ij}. Again there are far more observation points in the K dimension than the T dimension, and we find it convenient to implement a two-step regression procedure where we first carry out a linear regression with respect to moneyness for fixed t and T.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

First, for each maturity date T_i, we linearly regress the implied volatilities for the various strikes $(K_{ij})_j$ on the variable

$$(\mathrm{LM})_{ij} = \log(K_{ij}/x),$$

with $X_t = x$ being the present value for the underlying. We use the least-squares criterion to obtain the coefficients $\hat{a}_i$ and $\hat{b}_i$ which solve

$$\min_{a_i,b_i} \sum_j \left(I(T_i,K_{ij}) - (a_i(\mathrm{LM})_{ij} + b_i)\right)^2.$$

This step gives estimates

$$\hat{a}_i \approx a^{\varepsilon} \frac{\overline{\nu_{t,T_i}^{1/2}}}{\tau_i} + a^{\delta}, \qquad \hat{b}_i \approx b^{\star}_{t,T_i} \;=\; \bar{\sigma} + \Delta b\, \overline{\nu_{t,T_i}^{1/2}} + b^{\delta}\tau_i,$$

for each time to maturity $\tau_i = T_i - t$.

Next, we obtain estimates $\widehat{\Delta b}$, $\widehat{\bar{\sigma}}$ and $\widehat{b^{\delta}}$ by regressing on the factors $\overline{\nu_{t,T_i}^{1/2}}$ and τ_i as

$$\min_{\sigma,\Delta b,b^{\delta}} \sum_i \left\{\hat{b}_i - (\sigma + \Delta b\, \overline{\nu_{t,T_i}^{1/2}} + b^{\delta}\tau_i)\right\}^2. \tag{5.41}$$

Finally, to get estimates $\widehat{a^{\delta}}$ and $\widehat{a^{\varepsilon}}$, we regress the estimated skew slopes $\hat{a}_i$ on $\overline{\nu_{t,T_i}^{1/2}}/\tau_i$:

$$\min_{a^{\delta},a^{\varepsilon}} \left\{\hat{a}_i - (a^{\varepsilon} \frac{\overline{\nu_{t,T_i}^{1/2}}}{\tau_i} + a^{\delta})\right\}^2. \tag{5.42}$$

This then gives the group market parameters we need. Notice that we obtain estimates for each day t and can examine the stability properties by plotting these as a function of time and the goodness-of-fit properties by plotting the fit with respect to the calibrated model for a given day.

5.4.5 Maturity Cycles for S&P 500 Data

We illustrate the performance of the fitting, with the maturity cycles, on one day's data in Figures 5.11 and 5.12.

We argue that the above data analysis illustrates how just as implied volatility is a convenient transformation for capturing the deviations of options prices from the Black–Scholes theory, the LMMR representation is useful for picking out secondary phenomena relative to the basic skew captured by a stochastic volatility theory. We next illustrate this further by looking at aspects of the skew that cannot be captured by the first-order theory described above, but that can be handled by higher-order terms in the asymptotic price approximation.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

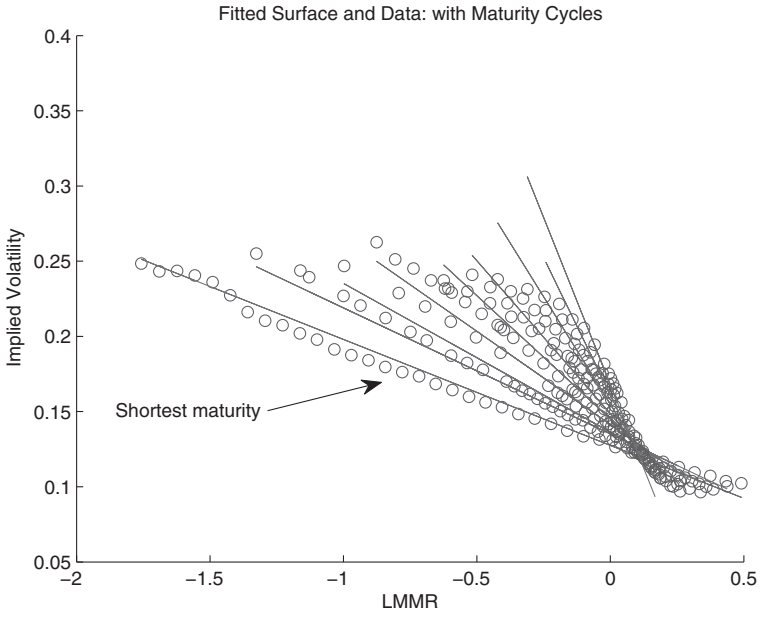

Figure 5.11 Fit of maturity cycles approximation formula (5.40) on S&P 500 options data from July 26, 2006. Notice the better fit to the shortest maturity.

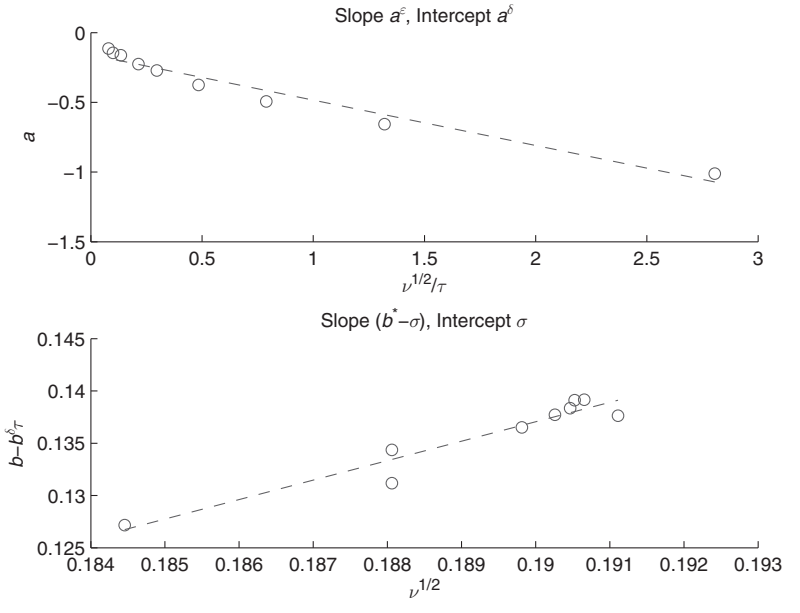

Figure 5.12 Term-structure fit of maturity cycles approximation formula (5.40) on S&P 500 options data from July 26, 2006. The top plot shows the fit of $\hat{a}$ against $\frac{\overline{v_{t,T}^{1/2}}}{\tau}$ in the formula (5.42). The bottom plot illustrates the fit (5.41) of $\hat{b}$ by plotting $\hat{b} - b^{\delta}\tau$ against $\overline{v_{t,T}^{1/2}}$.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

5.5 Higher-Order Corrections

We have discussed fitting of the implied volatility surface for a wide range of maturities. Fitting of the surface for very long or short-dated options led us to generalize the model by introducing, respectively, time-varying coefficients and a slow volatility factor. In this section we examine the implied volatility surface for extreme strikes giving large values for the moneyness. The surface often turns or shows "wings" for extreme strikes, see for instance Figure 5.13. The first-order perturbation theory gives skews that are affine in log-moneyness and does not capture these wings well. A natural extension is now to introduce higher-order terms in the expansion. The next set of terms turns out to give models that allow for skews that are quartic (fourth-order) polynomials in log-moneyness. Recall that in our perturbation approach we first define a class of stochastic volatility models containing fast and slow volatility factors. We then expand the corresponding pricing equation with respect to the two small parameters ε and δ, with

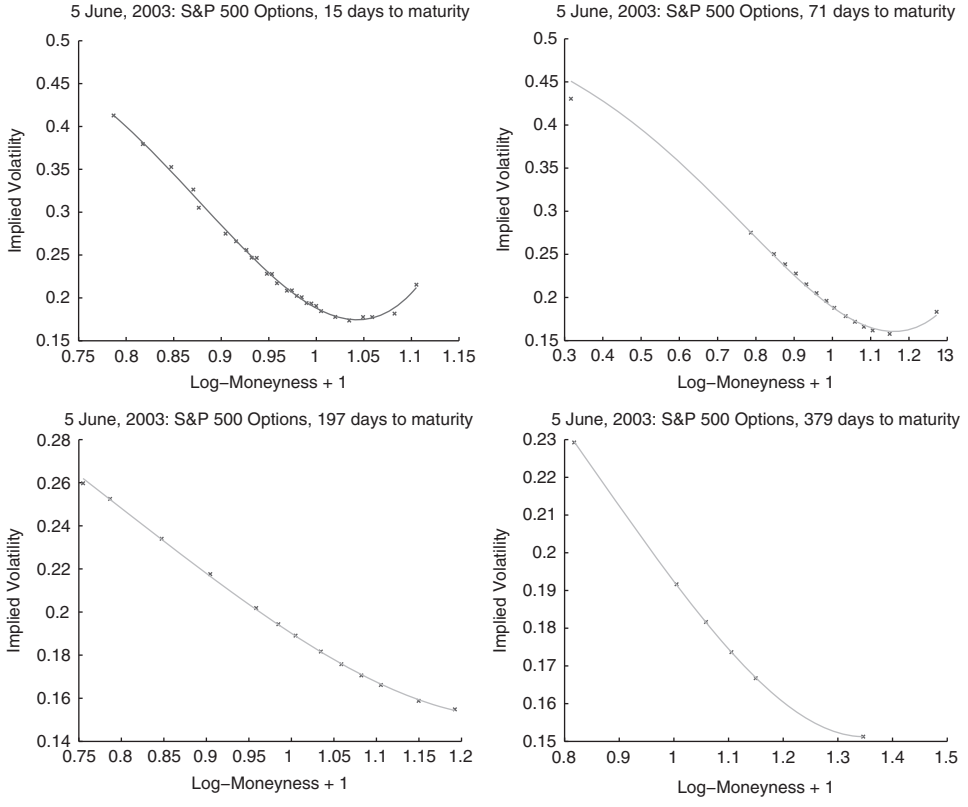

Figure 5.13 S&P 500 implied volatility data on June 5, 2003 and quartic fits to the asymptotic theory for four maturities.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

ε being the time scale of the fast factor and δ being the reciprocal of the time scale of the slow factor. In the notation introduced in Chapter 4, we write the price as

$$P^{\varepsilon,\delta} = P_0 + \sqrt{\varepsilon}P_{1,0} + \sqrt{\delta}P_{0,1} + \varepsilon P_{2,0} + \sqrt{\varepsilon\delta}P_{1,1} + \delta P_{0,2} + \cdots,$$

with $P_0 + \sqrt{\varepsilon}P_{1,0} + \sqrt{\delta}P_{0,1}$ corresponding to the first-order theory. In the second-order theory we include the next three terms as well. Indeed, this increases the number of parameters or degrees of freedom in the model for the implied volatility. The first-order theory involves the four group market parameters $(a^{\varepsilon}, b^{\star}, a^{\delta}, b^{\delta})$, whereas the second-order theory gives eleven parameters. By including the second-order terms we thus improve the quality of the fit to the skew and the accuracy of the pricing formulas. Here, we describe the approximation for the implied volatility that follows from including the second-order terms and how the extended model captures the observed wings in the skew. The theoretical derivation is presented in Sections 9.3 and 9.4. Notice also that in this section we consider the case with time-independent parameters.

The derivation of the second-order approximation to the implied volatility presented in Section 9.3.4 shows that outside of a small terminal layer (very close to expiration), the formula (5.18) for the implied volatilities now generalizes to

$$I \approx \sum_{j=0}^{4} a_j(\tau)\,(\mathrm{LMMR})^j + \frac{1}{\tau}\Phi_t, \tag{5.43}$$

where as before $\tau = T - t$, $\mathrm{LMMR} = \log(K/x)/(T-t)$, and Φ_t varies with the fast volatility factor Y_t . The LMMR coefficients are all third-order polynomials in the time to maturity and have the form:

$$a_1(\tau) = \sum_{k=0}^{3} a_{1,k}\tau^k, \quad a_2(\tau) = \sum_{k=0}^{3} a_{2,k}\tau^k,$$
$$a_3(\tau) = \sum_{k=2}^{3} a_{3,k}\tau^k, \quad a_4(\tau) = \sum_{k=2}^{3} a_{4,k}\tau^k.$$

The coefficients $a_{j,k}$ are now the relevant parameters and we discuss next practical aspects of fitting these to the observed implied volatility surface.

5.5.1 Second-Order Fitting

We describe how to calibrate the model (5.43) to the observed implied volatility records. Again, we employ a two-stage fitting procedure that

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

recognizes the thinness of data in the maturity dimension, relative to the many available strikes. The procedure now also makes use of the decomposition of the implied volatility into first and second-order parts. Since the second-order terms are correction terms that are associated with additional degrees of freedom, we first fit the data to the leading or first-order model exactly as in Section 5.2, and then we fit the residual to the additional degrees of freedom provided by the second-order theory. In other words, by viewing the wings as small corrections to the linear skew, we avoid the "tail wagging the dog" phenomenon.

The first stage in the fitting procedure is thus as described in Section 5.2. That is, for each day t we carry out the least squares regression in (5.24) to obtain $\hat{a}_i^I$ and $\hat{b}_i^I$ by

$$\min_{a_i^I, b_i^I} \sum_j \left(I(T_i, K_{ij}) - (a_i^I (\mathrm{LMMR})_{ij} + b_i^I) \right)^2,$$

for a given maturity τ_i. Note that the coefficient $\hat{b}_i^I$ contains a small, rapidly fluctuating component due to the term Φ_t. Here, however, our primary focus is the behavior of the surface in the moneyness dimension rather than the maturity dimension, as was our focus above.

Then we identify second-order effects in the implied volatility by computing the residuals

$$I^{II}(T_i, K_{ij}) = \frac{I(T_i, K_{ij}) - (a_i^I (\mathrm{LMMR})_{ij} + b_i^I)}{((\mathrm{LMMR})_{ij} + 1)^2}.$$

Observe that we introduced a shift in the denominator to avoid divide by zero issues. The residual is then fitted to a quadratic polynomial in LMMR as above,

$$\min_{c_i^{II}, a_i^{II}, b_i^{II}} \sum_j \left(I^{II}(T_i, K_{ij}) - (c_i^{II} (\mathrm{LMMR})_{ij}^2 + a_i^{II} (\mathrm{LMMR})_{ij} + b_i^{II}) \right)^2.$$

The final stage of the fitting procedure consists of fitting these coefficients that are parameterized by the time to maturity τ_i to a cubic polynomial in τ, again using the least squares criterion. The coefficients of the polynomial $a_1(\tau)$ solve

$$\min_{a_{1,k}} \sum_i \left(\hat{a}_i^I - \sum_{k=0}^{3} a_{1,k} \tau_i^k \right)^2,$$

and similarly for the other coefficients.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

5.5.2 Capturing the S&P 500 Wings

We continue our example with the S&P 500 data and now carry out the second-order fitting procedure described in the previous section.

First, we show some typical quartic fits of S&P 500 implied volatilities for a few maturities in Figure 5.13. These plots show that the quartic produced by second-order approximation becomes important in capturing the turn of the skew, in particular for short maturities.

Figure 5.14 shows the fits of the observed $a_j(\tau_i)$'s as introduced in (5.43) to their calibrated term-structure formulas for S&P 500 data on June 5, 2003.

The final step is to recover the parameters needed for pricing from the estimates of $\{a_{j,k}\}$, the analog of the group market parameters (5.5) in the first-order theory which are obtained via the calibration formulas (5.21) and (5.22). In the second-order theory these relations are no longer linear, and a nonlinear inversion algorithm is required based on formulas such as (9.53) for the fast factor presented in Chapter 9. This aspect has to be treated case by case in order to take advantage of the particular features of the market under study. For instance, in FX markets, the correlation between the

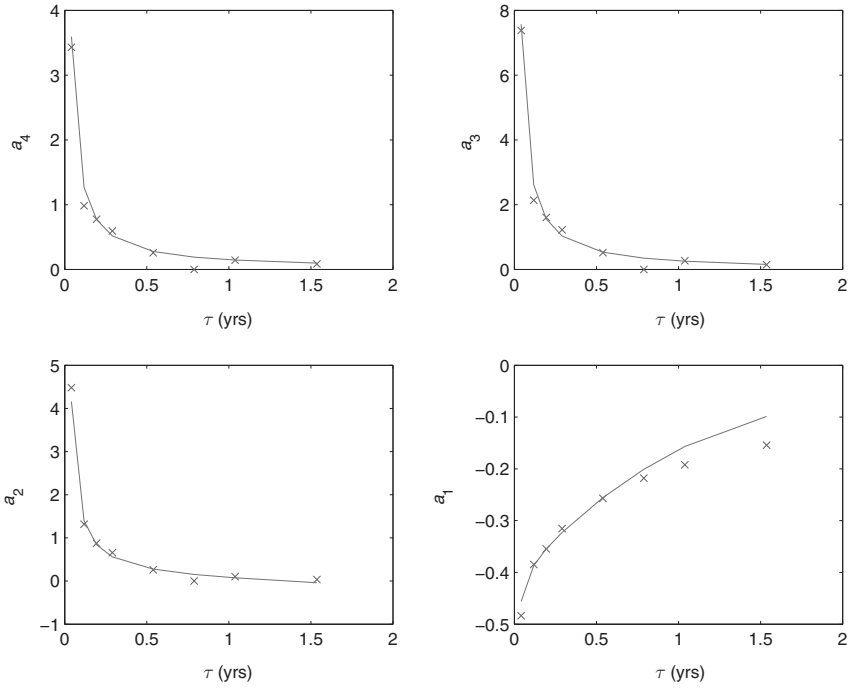

Figure 5.14 S&P 500 term-structure fit using second-order approximation. Data from June 5, 2003.

https://doi.org/10.1017/CBO9781139020534.006 Published online by Cambridge University Press

underlying and its volatility tends to be zero, which reduces the complexity of the implementation of the second-order theory.

Notes

The first-order calibration with fast and slow scales appeared in Fouque *et al.* (2003a), the time-dependent parameter derivation and calibration appeared in Fouque *et al.* (2004a), and the second-order calibration appeared in Fouque *et al.* (2004b). We thank Dan Lacker for research assistance with the data calibration in this chapter.

Phenomena with market dynamics that change as the time to the next "witching hour" gets small are well known. In fact, there are also special end-of-the-year effects, but we do not deal with those here.