

Chapter 9

Linking static and dynamic properties of stochastic volatility models

A stochastic volatility model can be assessed by studying its dynamic properties: the volatilities and correlations of volatilities, be they spot-starting or forward-starting and their joint dynamics with the spot. One can also choose to focus on the (static) smile it produces and examine the strike and maturity dependence of the implied volatilities it generates.

While static and dynamic features of a model are both determined by the joint dynamics of spot and forward variances in the model, how strong is the relationship between both? Which elements of this connection are specific to stochastic volatility models? Would it be possible to parametrize differently a given model so that the vanilla smile is (approximately) unchanged while the dynamics is different?

We establish a link between the rate at which the ATMF skew decays with maturity and the SSR, already introduced in the context of the local volatility model, whose definition we extend to the case of stochastic volatility models. Our analysis will allow us to split stochastic volatility models into two classes.¹

9.1 The ATMF skew

The ATMF skew vanishes for vanishing volatility of volatility. At order one it is given by (8.52). Setting $\varepsilon = 1$ and using the definition of $C^{x\xi}$ in (8.13) yields:

$$\mathcal{S}_T = \frac{1}{2\sqrt{T}} \frac{1}{\left(\int_0^T \xi_0^\tau d\tau\right)^{3/2}} \int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du \quad (9.1)$$

which can be rewritten as in (8.24), page 315:

$$\mathcal{S}_T = \frac{1}{2\hat{\sigma}_T^3 T} \int_0^T \frac{T-\tau}{T} \frac{\langle d\ln S_\tau \, d\hat{\sigma}_T^2(\tau) \rangle_0}{d\tau} d\tau \quad (9.2)$$

¹These results were first published in [11].

9.2 The Skew Stickiness Ratio (SSR)

The dynamics of a given model is reflected in the covariance of the spot and forward variances. Conditional on a move of the spot, different models generate different scenarios for implied volatilities. This can be quantified by focusing on the ATMF volatility and computing the regression coefficient of $\delta\hat{\sigma}_{F_T T}$ with respect to $\delta \ln S$.

It seems reasonable and natural to normalize this regression coefficient by the ATMF skew. We thus introduce the Skew Stickiness Ratio (SSR) \mathcal{R}_T , defined by:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T} \frac{E[d \ln S \, d\hat{\sigma}_{F_T(S)T}]}{E[(d \ln S)^2]} \quad (9.3)$$

where the notation $\hat{\sigma}_{F_T(S)T}$ emphasizes that the strike whose implied volatility we consider is not fixed. Unless necessary we will use the lighter notation $\hat{\sigma}_{F_T T}$.

\mathcal{R}_T is dimensionless – its value is known for some classes of models:

- In jump-diffusion or Lévy models with independent stationary increments for $\ln S$ – such as the model used in Section 5.3.2 – implied volatilities are a function of moneyness $\frac{K}{S}$ only: as S moves $\hat{\sigma}_{F_T T}$ is unchanged: $\mathcal{R}_T = 0, \forall T$.
- We have already encountered the SSR in the context of the local volatility model. Because implied volatilities are *functions* of (t, S) , formula (9.3) simplifies to:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T} \frac{d\hat{\sigma}_{F_T(S)T}}{d \ln S} = \frac{1}{\mathcal{S}_T} \left(\frac{d\hat{\sigma}_{KT}}{d \ln K} \Big|_{F_T} + \frac{d\hat{\sigma}_{KT}}{d \ln S} \Big|_{K=F_T} \right)$$

which agrees with the definition used for the SSR in Section 2.5.2. We know from Sections 2.5.3.1 and 2.4.6 that for time-independent local volatility functions for all maturities, or for general local volatility functions in the limit $T \rightarrow 0$, $\mathcal{R}_T = 2$. For equity index smiles, \mathcal{R}_T starts from 2 for short maturities and typically reaches values above 2 for long maturities – see for example Figure 2.4, page 59, and Figure 9.9, page 380.

We now compute \mathcal{R}_T in a general stochastic volatility model at lowest order in volatility of volatility. \mathcal{S}_T vanishes for vanishing volatility of volatility and starts with a term of order one. To get \mathcal{R}_T at lowest order we thus need to compute the covariance $E[dS \, d\hat{\sigma}_{F_T T}]$ at order one. From (8.21a), the difference between ATMF and VS volatilities is of order one. For the purpose of computing $E[dS \, d\hat{\sigma}_{F_T T}]$ at order one, we can thus conveniently replace the ATMF volatility with the VS volatility:

$$E[dS \, d\hat{\sigma}_{F_T T}] \simeq E[dS \, d\hat{\sigma}_T]$$

From the definition of the VS volatility $\widehat{\sigma}_T^2(t) = \frac{1}{T-t} \int_t^T \xi_t^u du$ we have:

$$\begin{aligned} E[d \ln S_t d\widehat{\sigma}_T(t)] &= \frac{1}{2(T-t)\widehat{\sigma}_T(t)} \int_t^T E[d \ln S_t d\xi_t^u] du \\ &= \frac{1}{2(T-t)\widehat{\sigma}_T(t)} \int_t^T \mu(t, u, \xi_0) du dt \end{aligned}$$

Setting $t = 0$ and dividing by $E[(d \ln S)^2]$ which is equal to $\xi_0^0 dt$, and by S_T , we get the following expression for \mathcal{R}_T at lowest non-trivial order in volatility of volatility:

$$\mathcal{R}_T = \frac{\int_0^T \xi_0^\tau d\tau}{T\xi_0^0} \frac{T \int_0^T \mu(0, u, \xi_0) du}{\int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du} \quad (9.4)$$

9.3 Short-maturity limit of the ATMF skew and the SSR

Let us take the limit $T \rightarrow 0$ in expression (9.1). We get:

$$\mathcal{S}_0 = \lim_{T \rightarrow 0} \frac{1}{2\sqrt{T}} \frac{1}{\left(\int_0^T \xi_0^\tau d\tau\right)^{3/2}} \int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du = \frac{\mu(0, 0, \xi_0)}{4(\xi_0^0)^{3/2}}$$

This recovers (8.36): at order one in volatility of volatility the short ATMF skew is a direct measure of the instantaneous spot/volatility covariance. Turning now to the SSR:

$$\begin{aligned} \mathcal{R}_0 &= \lim_{T \rightarrow 0} \frac{\int_0^T \xi_0^\tau d\tau}{T\xi_0^0} \frac{T \int_0^T \mu(0, u, \xi_0) du}{\int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du} = \lim_{T \rightarrow 0} \frac{T \int_0^T du}{\int_0^T d\tau \int_\tau^T du} \\ &= 2 \end{aligned}$$

Thus, in stochastic volatility models, the short limit of the SSR is 2, as in the local volatility model. This “2” is the 2 in the denominator of (8.36).

9.4 Model-independent range of the SSR

Let us assume that the term structure of VS volatilities is flat – $\xi_0^\tau \equiv \xi_0$ – and that the model at hand is time-homogeneous, that is the spot/variance covariance function $\mu(\tau, u, \xi)$ only depends on $u - \tau$. We now use the economical notation $\mu(u - \tau)$:

$$\mu(\tau, u, \xi_0) = \lim_{d\tau \rightarrow 0} \frac{1}{d\tau} E[d \ln S_\tau d\xi_\tau^u] \equiv \mu(u - \tau)$$

One of the integrals in (9.1) can be done analytically and we get the following simple expressions for \mathcal{S}_T and \mathcal{R}_T :

$$\mathcal{S}_T = \frac{1}{2\xi_0^{3/2}T^2} \int_0^T (T-t)\mu(t)dt \quad (9.5)$$

$$\mathcal{R}_T = \frac{\int_0^T \mu(t)dt}{\int_0^T (1-\frac{t}{T})\mu(t)dt} \quad (9.6)$$

These expressions for \mathcal{S}_T and \mathcal{R}_T only involve the spot/variance covariance function μ – can we derive general properties without assuming a specific form for $\mu(t)$?

Let us make the natural assumption that $\mu(t)$ decays monotonically towards zero as $t \rightarrow \infty$. Define $g(\tau)$ as:

$$g(\tau) = \int_0^\tau \mu(t)dt \quad (9.7)$$

\mathcal{S}_T and \mathcal{R}_T can be rewritten as:

$$\mathcal{S}_T = \frac{1}{2\xi_0^{3/2}T^2} \int_0^T g(\tau)d\tau \quad \mathcal{R}_T = \frac{g(T)}{\frac{1}{T} \int_0^T g(\tau)d\tau} \quad (9.8)$$

$g(\tau) = 0$ for $\tau = 0$ and is either increasing concave if $\mu(t) \geq 0$, or decreasing convex if $\mu(t) \leq 0$.

- \mathcal{R}_T is the ratio of $g(T)$ to its average value over $[0, T]$, thus $\mathcal{R}_T \geq 1$.
- $g(\tau)$ is either positive and concave, or negative and convex. Thus, $\frac{g(\tau)}{g(T)} \geq \frac{\tau}{T}$. This yields: $\mathcal{R}_T = \frac{1}{\frac{1}{T} \int_0^T \frac{g(\tau)}{g(T)} d\tau} \leq \frac{1}{\frac{1}{T} \int_0^T \frac{\tau}{T} d\tau} = 2$.

Thus, for a time-homogeneous model such that its spot/variance covariance function decays monotonically, and for a flat term structure of VS volatilities, we get, at order one in volatility of volatility, the following model-independent range for \mathcal{R}_T :

$$\mathcal{R}_T \in [1, 2] \quad (9.9)$$

In diffusive stochastic volatility models – with the assumptions we have made – the SSR cannot go below 1. To get lower values for \mathcal{R}_T one presumably needs to incorporate a jump or Lévy component in the process for $\ln S_t$.

(9.9) strictly holds for a flat term structure of VS volatilities. Glancing again at the definition of the SSR in (9.3) we can see that the numerator is proportional to the

instantaneous volatility of S_t while the denominator is proportional to its square, thus the SSR is inversely proportional to the short VS volatility.

It can be made artificially small or large by shifting the short end of the variance curve – one should bear this in mind when assessing the SSR of a given market smile.

An example of the impact of the term structure of VS volatilities is discussed on page 366.

For $T \rightarrow 0$, assuming that $\mu(t)$ is smooth as $t \rightarrow 0$, $g(\tau) = \tau\mu(0)$: expression (9.8) for \mathcal{R}_T again yields:

$$\mathcal{R}_0 = 2$$

9.5 Scaling of ATMF skew and SSR – a classification of models

To investigate further the connection between \mathcal{S}_T and \mathcal{R}_T we need a characterization of the rate of decay of $\mu(t)$. Let us assume that for $t \rightarrow \infty$ $\mu(t)$ decays with a characteristic exponent γ :

$$\mu(t) \propto \frac{1}{t^\gamma} \quad (9.10)$$

Consider $g(\tau) = \int_0^\tau \mu(t)dt$. For large τ , it is equal to $C + \alpha\tau^{1-\gamma}$, where C, α are constants. If $\gamma > 1$ it tends towards C while for $\gamma < 1$ it is equivalent to $\alpha\tau^{1-\gamma}$.

Now turn to $\int_0^T g(\tau)d\tau$: for large T it is equal to $B + CT + \frac{\alpha}{2-\gamma}T^{2-\gamma}$. For $T \rightarrow \infty$ this quantity is equivalent to CT if $\gamma > 1$ while it is equivalent to $\frac{\alpha}{2-\gamma}T^{2-\gamma}$ if $\gamma < 1$. As a result we get, using formulas (9.8), two types of behavior for \mathcal{S}_T and \mathcal{R}_T , which leads to a division of stochastic volatility models into two classes.

For long maturities:

- (Type I) If $\gamma > 1$:

$$\mathcal{S}_T \propto \frac{1}{T} \text{ and } \lim_{T \rightarrow \infty} \mathcal{R}_T = 1 \quad (9.11)$$

- (Type II) If $\gamma < 1$:

$$\mathcal{S}_T \propto \frac{1}{T^\gamma} \text{ and } \lim_{T \rightarrow \infty} \mathcal{R}_T = 2 - \gamma \quad (9.12)$$

We leave it to the reader to check that exponential decay of $\mu(t)$ produces Type I behavior.

Both Type I and Type II scalings are compactly summarized in the following relationship. For $T \rightarrow \infty$:

$$\mathcal{S}_T \propto \frac{1}{T^{2-\mathcal{R}_\infty}} \quad (9.13)$$

9.6 Type I models – the Heston model

The fact that for a fast-decaying spot/variance covariance function S_T decays like $\frac{1}{T}$ is not unexpected. Indeed in this case, in the limit of long time intervals, increments of $\ln S_t$ become independent and identically distributed. Cumulants of $\ln S_t$ then scale linearly with T , thus the skewness s_T of $\ln S_T$ scales like $\frac{1}{\sqrt{T}}$. Consequently, at order one in volatility of volatility, or equivalently at order one in s_T , the skew/skewness relationship (8.23) implies that the ATMF skew scales like $\frac{1}{T}^2$.²

The Heston model provides an example of Type I behavior. $\mu(t)$ is exponentially decaying – see expressions (8.48) – hence we expect Type I scaling (9.35). Let us verify this by using expressions (6.19) for $\widehat{\sigma}_{F_T T}$ and S_T at order one in volatility of volatility for long maturities.

The ATMF skew is given by (6.19a):

$$S_T = \frac{\rho\sigma}{2kT} \frac{1}{\sqrt{V^0}} \quad (9.14)$$

S_T indeed has Type I scaling. Consider now the SSR. $\widehat{\sigma}_{F_T T}$ in (6.19a) is a function of V – its covariance with $d \ln S$ is thus given by:

$$\begin{aligned} \frac{E[d\widehat{\sigma}_{F_T T} d \ln S_t]}{E[(d \ln S_t)^2]} &= \frac{\partial \widehat{\sigma}_{F_T T}}{\partial V} \frac{E[dV d \ln S]}{V dt} \\ &= \frac{1}{2kT\sqrt{V^0}} \frac{E[dV d \ln S]}{V dt} \\ &= \frac{1}{2kT\sqrt{V^0}} \rho\sigma \end{aligned} \quad (9.15)$$

where in $\frac{\partial \widehat{\sigma}_{F_T T}}{\partial V}$ we have kept the contribution at zeroth order in volatility of volatility to get the covariance of $d\widehat{\sigma}_{F_T T}$ and $d \ln S_t$ at order one. Dividing (9.14) by (9.15) yields $\mathcal{R}_T = 1$: we have confirmed by hand that the Heston model is indeed of Type I – this was already observed in [8].

²This $\frac{1}{T}$ scaling of the ATMF skew is also shared by jump-diffusion and Lévy models, at order one in the skewness of $\ln S_T$, as in these models, increments of $\ln S_t$ are indeed independent and identically distributed. The skewness of $\ln S_T$ exactly scales like $\frac{1}{\sqrt{T}}$, and at order one in this skewness, the skew scales like $\frac{1}{T}$.

However, while $\lim_{T \rightarrow \infty} \mathcal{R}_T = 1$ in Type I models, the behavior of \mathcal{R}_T in jump/Lévy models is different. Because implied volatilities are a function of K/S only, $\mathcal{R}_T = 0$, $\forall T$.

9.7 Type II models

In Type I models, the long-maturity scaling of the ATMF skew and the limit of the SSR are fixed and do not depend on γ . Depending on the size of γ , the long-maturity regime sets in for shorter or longer maturities, but the limiting behavior of the ATMF skew and of the SSR is universal and bears no trace of the precise underlying dynamics.

Type II models are richer as both the scaling of \mathcal{S}_T and the limit of \mathcal{R}_T are non-trivial and reflect the characteristic exponent of the decay of the spot/variance covariance function $\mu(t)$.

Moreover, the ATMF skew of market smiles typically decays algebraically, with an exponent around $-\frac{1}{2}$. This calls for a Type II model; hence the following natural questions:

- Is it practically possible to build a Type II model?
- Is the *dynamics* of market smiles consistent with Type II?

Type II scaling in the two-factor model

Consider an N -factor model of the type studied in Section 7.3:

$$\begin{cases} dS_t = (r - q)S_t dt + \sqrt{\xi_t^S} S_t dW_t^S \\ d\xi_t^T = \omega \xi_t^T \sum_i w_i e^{-k_i(T-t)} dW_t^i \end{cases}$$

Assuming a flat term structure of VS volatilities the model is time-homogeneous and the spot/variance covariance function reads:

$$\mu(\tau) = \omega \xi_0^{\frac{3}{2}} \sum_i w_i \rho_{iS} e^{-k_i \tau}$$

where ρ_{iS} is the correlation between W^S and W^i . Inserting this expression into formulas (9.5) and (9.6) yields the following expressions for the ATMF skew and the SSR:

$$\mathcal{S}_T = \frac{\omega}{2} \sum_i w_i \rho_{iS} \frac{k_i T - (1 - e^{-k_i T})}{(k_i T)^2} \quad (9.16a)$$

$$\mathcal{R}_T = \frac{\sum_i w_i \rho_{iS} \frac{1 - e^{-k_i T}}{k_i T}}{\sum_i w_i \rho_{iS} \frac{k_i T - (1 - e^{-k_i T})}{(k_i T)^2}} \quad (9.16b)$$

$\mu(\tau)$ is a linear combination of exponentials. As $\tau \rightarrow \infty$, $\mu(\tau) \propto e^{-\min_i k_i T}$. Thus for $T \rightarrow \infty$ this model is of type I: $\mathcal{S}_T \propto \frac{1}{T}$ and $\lim_{T \rightarrow \infty} \mathcal{R}_T = 1$; this can be

checked explicitly on expressions (9.16). The fact that $\mu(\tau)$ is a linear combination of exponentials is dictated by the property that this form allows for a Markov representation of the model with a number of state variables – besides S_t – equal to the number of exponentials – see Section 7.1.

We could try and use a model whose dynamics for forward variances reads:

$$d\xi_t^T = \omega \xi_t^T \frac{1}{(T-t+\theta)^\gamma} dW \quad (9.17)$$

where θ is a (small) offset parameter. In this model $\mu(\tau)$ has the desired power law scaling for large τ : $\mu(\tau) \propto \frac{1}{\tau^\gamma}$. Unfortunately we lose the Markov representation: it is no longer possible to express the set of forward variances ξ^T as a function of a finite number of state variables. Such a model is not usable in practice.

Luckily, by suitably choosing model parameters it is possible to get Type II scaling in an N -factor model – in fact a two-factor model – over a range of maturities that is sufficient for practical purposes, even though Type I scaling eventually kicks in for (very) long maturities.

Let us use the two-factor model of Section 8.7 with the parameters listed in Table 8.2, page 329. Remember that (a) $\nu, \theta, k_1, k_2, \rho$ have been selected so that the volatilities of VS volatilities decay approximately as a power law with exponent 0.4, (b) ρ_{SX^1} and ρ_{SX^2} have been chosen so that the ATMF skew decays approximately as a power law with exponent $\frac{1}{2}$. This is illustrated in Figure 8.5.

At order one in volatility of volatility the ATMF skew is given by expression (8.54), page 329. The SSR is calculated using expression (9.4) and we have:

$$\mathcal{S}_T = \frac{\nu \alpha_\theta}{\hat{\sigma}_T^3 T^2} \int_0^T dt \sqrt{\xi_0^t} \int_t^T \xi_0^u \left[(1-\theta) \rho_{SX^1} e^{-k_1(u-t)} + \theta \rho_{SX^2} e^{-k_2(u-t)} \right] du \quad (9.18)$$

$$\mathcal{R}_T = \frac{1}{\sqrt{\xi_0^0}} \frac{\hat{\sigma}_T^2 T \int_0^T \xi_0^t \left[(1-\theta) \rho_{SX^1} e^{-k_1 t} + \theta \rho_{SX^2} e^{-k_2 t} \right] dt}{\int_0^T dt \sqrt{\xi_0^t} \left(\int_t^T \xi_0^u \left[(1-\theta) \rho_{SX^1} e^{-k_1(u-t)} + \theta \rho_{SX^2} e^{-k_2(u-t)} \right] du \right)} \quad (9.19)$$

For a flat term structure of VS volatilities:

$$\mathcal{S}_T = \nu \alpha_\theta \left[(1-\theta) \rho_{SX^1} \frac{k_1 T - (1 - e^{-k_1 T})}{(k_1 T)^2} + \theta \rho_{SX^2} \frac{k_2 T - (1 - e^{-k_2 T})}{(k_2 T)^2} \right] \quad (9.20)$$

$$\mathcal{R}_T = \frac{(1-\theta) \rho_{SX^1} \frac{1-e^{-k_1 T}}{k_1 T} + \theta \rho_{SX^2} \frac{1-e^{-k_2 T}}{k_2 T}}{(1-\theta) \rho_{SX^1} \frac{k_1 T - (1 - e^{-k_1 T})}{(k_1 T)^2} + \theta \rho_{SX^2} \frac{k_2 T - (1 - e^{-k_2 T})}{(k_2 T)^2}} \quad (9.21)$$

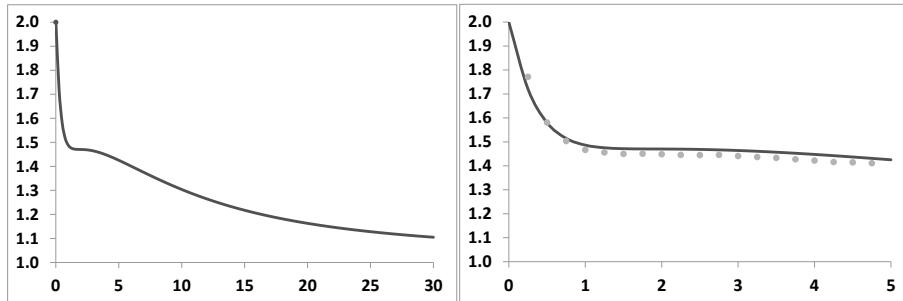


Figure 9.1: The SSR of the two-factor model, computed with formula (9.20) and parameters in Table 8.2, page 329. Left: maturities up to 30 years – right: zoom on maturities less than 5 years. The dots are the result of a Monte Carlo evaluation of the SSR – see Section 9.8 below.

\mathcal{R}_T , computed using (9.21), appears in Figure 9.1, as a function of T .

Consider first the left-hand graph. The short-maturity limit of the SSR is 2, as shown in Section 9.3. As $T \rightarrow \infty$, the SSR tends to 1 – as it should. Notice however the shoulder for a value of the SSR around 1.5, which appears in more detail in the right-hand graph. The SSR for maturities from 1 to 5 years is stable around 1.5. For this range of maturities, the model obeys the Type II scaling rules in (9.36): the value of the SSR is 2 minus the characteristic exponent of the decay of the ATMF skew – in our case $\frac{1}{2}$.

Note in the right-hand graph how the approximate expression (9.16b), obtained at order one in volatility of volatility, agrees well with the actual value of the SSR.

Another example is shown in Figure 9.2: here we have chosen ρ_{SX^1} , ρ_{SX^2} so that the ATMF skew decays approximately with a characteristic exponent equal to 0.3 for maturities up to 5 years (left-hand graph). As the right-hand side graph shows, we get an SSR around 1.7, i.e. $2 - 0.3$ over this range.³

Formula (9.16b) for the SSR has been derived for a flat term structure of VS volatilities. When this is not the case, the prefactor $\xi^u \sqrt{\xi^t}$ in expression (8.50), page 327, for $\mu(t, u, \xi)$ in the two-factor model is not constant and, at order one in volatility of volatility, both the numerator and denominator of \mathcal{R}_T have to be calculated by numerical integration.

The SSR thus depends on the term structure of VS volatilities. For example, for a VS term structure that increases (decreases) from 20% to 25% for $T = 1$ year, the SSR for this maturity is 1.46 (1.50). Let us take $T = 5$ years; for a VS term structure that increases (decreases) from 20% to 30%, the SSR is 1.43 (1.49). These values

³With parameters $\nu, \theta, k_1, k_2, \rho$ already set so as to generate the desired term structure of volatilities of VS volatilities, ρ_{SX^1}, ρ_{SX^2} are the only handles left to control the term structure of (a) the ATMF skew, (b) the SSR. Introducing more factors allows for more flexibility, but in the author's experience two factors are sufficient for capturing typical market ATMF skews and SSR.

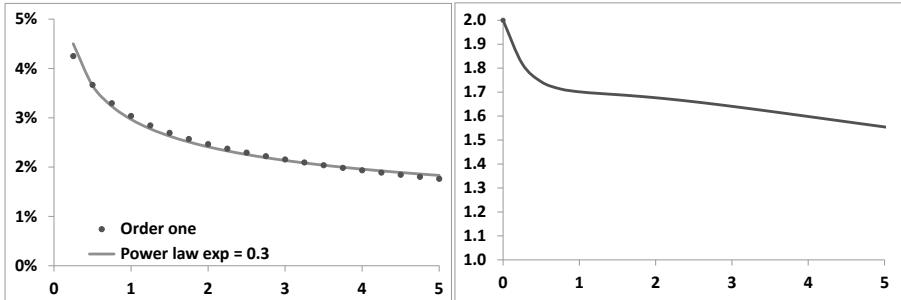


Figure 9.2: Left: The ATM skew measured as the difference of implied volatilities for strikes $0.95F_T$ and $1.05F_T$ at order one in volatility of volatility (dots, formula (9.16a)) and a power law benchmark with exponent 0.3 (line). Right: the SSR at order one for maturities up to 5 years computed with expression (9.16b). The parameters are those of Table 8.1, with $\rho_{SX^1} = -35\%$, $\rho_{SX^2} = -83\%$.

are obtained in a Monte Carlo simulation of the two-factor model – see Section 9.8 below – with the same parameters as in Figure 9.1. For all practical purposes, the accuracy of formula (9.16b) for \mathcal{R}_T is thus adequate as long as VS term structures are not too steep.

The definition of the SSR in (9.3) involves in the denominator the instantaneous variance of S at $t = 0$. In case short VS volatilities are, say, substantially larger than longer-dated ones, we get a smaller value for the SSR than that given by expression (9.16b). For the purpose of evaluating the SSR at $t = 0$, approximation (9.16b) is thus incorrect.

However, it is the average level of spot/volatility covariance generated by the model up to the maturity of our exotic option, rather than its particular value at $t = 0$, that is practically relevant. It is then advisable to replace the short-dated variance in the denominator of (9.3) with the square of the VS volatility for maturity T . Approximation (9.16b) is then almost exact.

In conclusion, while the two-factor model is strictly speaking a Type I model, by suitably choosing parameters we obtain Type II scaling for a range of maturities that is practically relevant. For these maturities we are able to generate a decay of the ATM skew with an exponent less than 1 and the relationship between \mathcal{S}_T and \mathcal{R}_T in (9.13) is approximately obeyed.

Type II scaling in reality

We are able to obtain Type II scaling in a two-factor model, but are actual market smiles consistent with Type II behavior? This is characterized in (9.36) by two features: one static, one dynamic:

- The ATM skew decays with a non-trivial exponent: $\mathcal{S}_T \propto \frac{1}{T^\gamma}$ with $\gamma < 1$

- The SSR is different than 1 and is related to γ : for long maturities $\mathcal{R}_T \rightarrow 2 - \gamma$

The decay of the ATM skew of market smiles is consistent with Type II, but what about the SSR? From the expression of the SSR in (9.3) we define the realized SSR as:

$$\mathcal{R}_T^r = \frac{\sum_i \ln \frac{S_{i+1}}{S_i} (\hat{\sigma}_{T,i+1} - \hat{\sigma}_{T,i})}{\sum_i S_{T,i} (\ln \frac{S_{i+1}}{S_i})^2} \quad (9.22)$$

where $\hat{\sigma}_{T,i}$ (resp. $S_{T,i}$) is the ATM implied volatility (resp. ATMF skew) at time i for residual maturity T .

\mathcal{R}_T^r for $T = 2$ years is shown in Figure 9.3 for the Euro Stoxx 50 and S&P 500 indexes. We have used for simplicity ATM rather than ATMF volatilities.

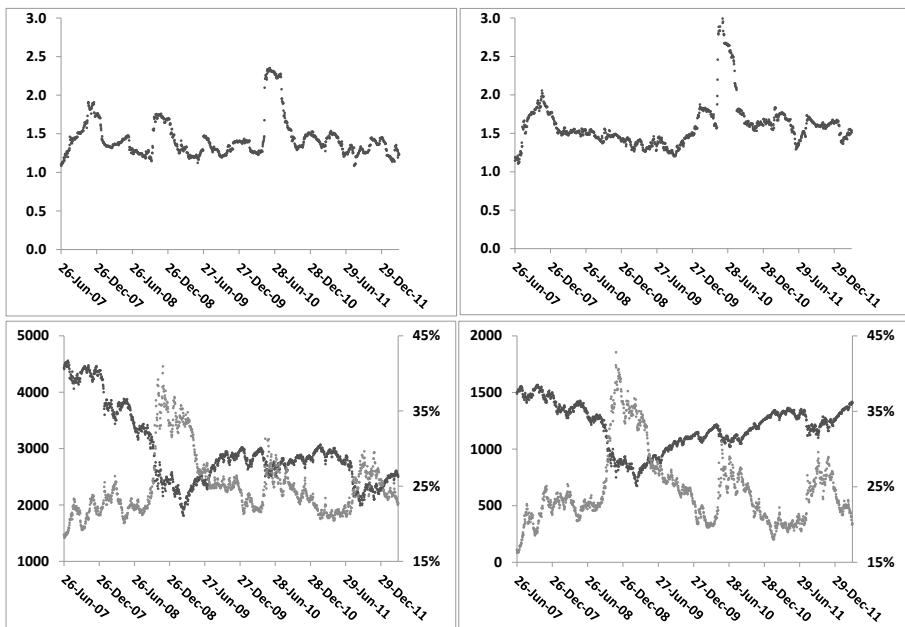


Figure 9.3: Top: 3-months realized SSR (\mathcal{R}_T^r) for the Euro Stoxx 50 (left) and S&P 500 (right) indexes, for $T = 2$ years. Bottom: the underlying (dark dots, left axis) and 2-year ATM implied volatility (lighter dots, right axis).

For both indexes, the realized SSR hovers around 1.5, with occasional spikes generated by simultaneous and opposite moves of S and $\hat{\sigma}_{F_T(S)T}$. The magnitude of the spikes and of the fluctuations around 1.5 is not relevant; it would be larger (resp. smaller) had we estimated \mathcal{R}_T^r on a shorter (resp. longer) window than 3 months – the duration of the spikes is indeed about 3 months. Notice also how similar ATM implied volatilities are for both indexes. Realized values for the SSR for longer-dated maturities – say $T = 5$ years – are comparable to the 2-year case.

In conclusion, the dynamics of equity market smiles is consistent with Type II behavior: $\mathcal{S}_T \propto \frac{1}{T^\gamma}$ with $\gamma < 1$ and for large T , $\gamma + \mathcal{R}_T$ is approximately equal to 2.

9.8 Numerical evaluation of the SSR

At order one in volatility of volatility, for the sake of computing the SSR the ATMF volatility can be substituted with the VS volatility and this leads to the analytic expression (9.16a) for a flat term structure of VS volatilities.

Generally, however, the SSR can be easily computed numerically. In the two-factor model, the ATMF volatility is a function of X^1, X^2 . At $t = 0$:

$$\widehat{\sigma}_{F_T T} \equiv \widehat{\sigma}_{F_T T}(X_0^1, X_0^2)$$

Expanding at first order in dX^1, dX^2 :

$$d\widehat{\sigma}_{F_T T} = \frac{d\widehat{\sigma}_{F_T T}}{dX^1} dX^1 + \frac{d\widehat{\sigma}_{F_T T}}{dX^2} dX^2$$

From the definition of the SSR in (9.3) and using that $E[(d \ln S_t)^2] = \xi_0^0 dt$, $E[d \ln S_t dX_t^1] = \rho_{SX^1} \sqrt{\xi_0^0} dt$ and likewise for X^2 , we get:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T} \frac{E[d \ln S_t d\widehat{\sigma}_{F_T T}]}{E[(d \ln S_t)^2]} = \frac{1}{\mathcal{S}_T} \frac{1}{\sqrt{\xi_0^0}} \left(\frac{d\widehat{\sigma}_{F_T T}}{dX^1} \rho_{SX^1} + \frac{d\widehat{\sigma}_{F_T T}}{dX^2} \rho_{SX^2} \right)$$

Thus \mathcal{R}_T can be simply evaluated numerically by computing $\widehat{\sigma}_{F_T T}$ with two different initial values for (X, Y) :

$$\mathcal{R}_T \simeq \frac{1}{\mathcal{S}_T} \frac{1}{\sqrt{\xi_0^0}} \frac{\widehat{\sigma}_{F_T T}(X_0^1 + \varepsilon \rho_{SX^1}, X_0^2 + \varepsilon \rho_{SX^2}) - \widehat{\sigma}_{F_T T}(X_0^1, X_0^2)}{\varepsilon}$$

with ε a small offset. Typically we take $X_0^2 = X_0^1 = 0$.

In stochastic volatility models defined by the dynamics of the instantaneous variance V_t , the SSR is simply computed by shifting V as volatilities for a fixed moneyness are a function of V . For example, in the Heston model, using the notations of Chapter 6, \mathcal{R}_T is simply given by:

$$\mathcal{R}_T \simeq \frac{1}{\mathcal{S}_T} \frac{\widehat{\sigma}_{F_T T}(V + \varepsilon \rho \sigma) - \widehat{\sigma}_{F_T T}(V)}{\varepsilon}$$

9.9 The SSR for short maturities

In both Type I and Type II models the short-maturity limit of the SSR is 2.

In a stochastic volatility model, in the limit $T \rightarrow 0$, the VS and ATM volatilities are identical. We will thus use the notation $\hat{\sigma}_0$ for both.

From the expression of the SSR in (9.3), in the special case of vanishing maturities, we can derive two definitions of the *realized* SSR according to whether we choose:

- the realized value of the instantaneous variance $(\ln \frac{S_{i+1}}{S_i})^2$
- or the implied value of the realized variance $\hat{\sigma}_{T,i}^2 \Delta t$

in the denominator of (9.3).

For reasons that will be clearer shortly, we choose the second convention. For small T we define $\mathcal{R}_T^{r,\text{short}}$ as:

$$\mathcal{R}_T^{r,\text{short}} = \frac{\sum_i \ln \frac{S_{i+1}}{S_i} (\hat{\sigma}_{T,i+1} - \hat{\sigma}_{T,i})}{\Delta t \sum_i S_{T,i} \hat{\sigma}_{T,i}^2} \quad (9.23)$$

where Δt is the duration of one (trading) day.

\mathcal{R}_T^r for $T = 1$ month is shown in Figure 9.4 for the Euro Stoxx 50 and S&P 500 indexes.

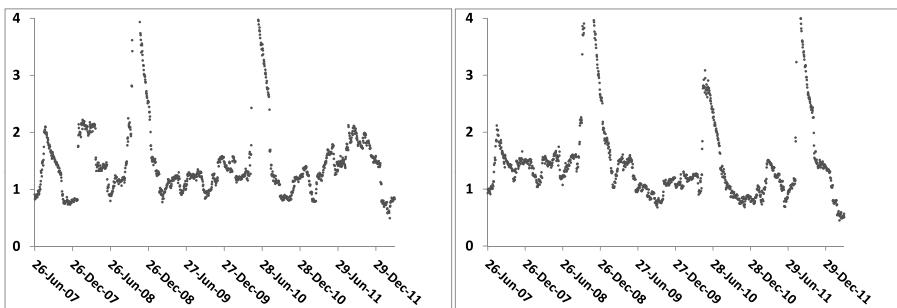


Figure 9.4: 3-months realized SSR (\mathcal{R}_T^r) for the Euro Stoxx 50 (left) and S&P 500 (right) indexes, for $T = 1$ month.

The average value of \mathcal{R}_T^r is substantially lower than the model-independent “implied” value of 2. Does this point to a discrepancy between the dynamics of stochastic models and the dynamics of market smiles?

Consider the example of realized and implied volatilities – typically, for equity indexes, VS implied volatilities are systematically higher on average than realized volatilities. This in itself does not signal a modeling inconsistency. What is important is that the difference between both volatilities can be materialized as the P&L of an option position – here a VS.

The natural question is thus: is it possible to materialize the difference $2 - \mathcal{R}_T^r$ as the P&L of an option position?

9.10 Arbitraging the realized short SSR

The SSR expresses the implied spot/volatility covariance in units of the ATM skew. Arbitraging $\mathcal{R}_T^{r,\text{short}}$ entails being able to materialize this covariance as a P&L. In what follows we set $T = 1$ month and assume zero interest rate and repo for simplicity.

To single out the realized spot/volatility covariance as a carry P&L we risk-manage our option position using the lognormal model of Section 8.5.1. In the limit $T \rightarrow 0$ the variance curve collapses to a single volatility $\hat{\sigma}_0$.

Since any stochastic volatility model calibrated to the market smile yields the same implied value for the instantaneous covariance of S and $\hat{\sigma}_0$ we could use any dynamics for $\hat{\sigma}_0$. However, in the lognormal model the short ATM skew is independent on $\hat{\sigma}_0$, a property that is also approximately shared by market smiles.

Choosing the lognormal model for risk-managing our option position then reduces chances that the mark-to-market P&L generated by remarking the ATM skew to market is large.

9.10.1 Risk-managing with the lognormal model

Consider the limit $T \rightarrow 0$ in the pricing equation (8.1): P becomes a function of $t, S, \hat{\sigma}_0$: $P(t, S, \xi) \equiv P(t, S, \hat{\sigma}_0)$. In the lognormal model for the instantaneous volatility the pricing equation is given by:

$$\frac{dP}{dt} + \frac{\hat{\sigma}_0^2}{2} S^2 \frac{d^2 P}{dS^2} + \frac{\nu^2}{2} \hat{\sigma}_0^2 \frac{d^2 P}{d\hat{\sigma}_0^2} + \rho\nu\hat{\sigma}_0^2 S \frac{d^2 P}{dS d\hat{\sigma}_0} = 0 \quad (9.24)$$

where ν is the (lognormal) volatility of $\hat{\sigma}_0$ and ρ the correlation of $\hat{\sigma}_0$ with S .

We have derived in Section 8.5.1 the smile at order two in volatility of volatility, in the limit $T \rightarrow 0$. At this order, expressions (8.39a) and (8.39b) for the ATM skew and curvature yield the following expression for the price of a vanilla option of strike K , maturity T :

$$P = P_{BS}(t, S, \hat{\sigma}(x), K, T) \quad (9.25a)$$

$$\hat{\sigma}(x) = \hat{\sigma}_0 + \mathcal{S}x + \frac{\mathcal{C}}{2}x^2 \quad (9.25b)$$

where P_{BS} is the Black-Scholes formula, $x = \ln(\frac{K}{S})$ and \mathcal{S}, \mathcal{C} are given by:

$$\mathcal{S} = \frac{\rho\nu}{2} \quad \mathcal{C} = \frac{2 - 3\rho^2}{6\hat{\sigma}_0}\nu^2$$

For a market smile of the form (9.25b) ρ, ν are given by:

$$\rho = \frac{2\mathcal{S}}{\sqrt{3\hat{\sigma}_0\mathcal{C} + 6\mathcal{S}^2}} \quad \nu = \sqrt{3\hat{\sigma}_0\mathcal{C} + 6\mathcal{S}^2} \quad (9.26)$$

In the limit $T \rightarrow 0$, \mathcal{S} and \mathcal{C} , derived at order two in volatility of volatility, do not depend on T and the ATM implied volatility is identical to the VS volatility. This identity is exact; indeed consider a diffusive model whose instantaneous volatility at $t = 0$ is σ . Using expression (8.64) with $\widehat{\sigma} = 0$, the price of a European option whose payoff is $f(S)$ is given, at order one in T by: $P = f(S_0) + \frac{\sigma^2 T}{2} S_0^2 \frac{d^2 f}{dS_0^2}$ where S_0 is the spot value at $t = 0$. Denote now by $\widehat{\sigma}$ the implied volatility of this option. At order one in T , $\widehat{\sigma}$ is such that $P = f(S_0) + \frac{\widehat{\sigma}^2 T}{2} S_0^2 \frac{d^2 f}{dS_0^2}$. This yields $\widehat{\sigma} = \sigma$: the implied volatilities of all European payoffs f such that $\frac{d^2 f}{dS_0^2} \neq 0$ are identical, and equal to the instantaneous volatility.⁴

Consider now a long position in a delta-hedged European option, risk-managed with the lognormal model. We will calibrate \mathcal{S} , \mathcal{C} – or equivalently ρ , ν – to the market smile near the money on a daily basis. Denote by $\Pi(t, S, \widehat{\sigma}_0, \mathcal{S}, \mathcal{C})$ the value of our European option position. We do not consider for now the mark-to-market P&L generated by a change of \mathcal{S} , \mathcal{C} – or equivalently of ρ , ν – and instead focus on the carry P&L. This P&L during δt , at order one in δt and two in δS and $\delta \widehat{\sigma}_0$, is given by:

$$P\&L = \frac{d\Pi}{dt} \delta t + \frac{d\Pi}{d\widehat{\sigma}_0} \delta \widehat{\sigma}_0 + \frac{1}{2} \frac{d^2 \Pi}{dS^2} \delta S^2 + \frac{1}{2} \frac{d^2 \Pi}{d\widehat{\sigma}_0^2} (\delta \widehat{\sigma}_0)^2 + \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \delta S \delta \widehat{\sigma}_0 \quad (9.27)$$

While the delta hedge removes the term in δS , there remains a contribution in $\delta \widehat{\sigma}_0$ as we are not vega-hedged.

Using (9.24), at order one in δt and two in δS and $\widehat{\sigma}_0$ our P&L is given by:

$$P\&L = \frac{d\Pi}{d\widehat{\sigma}_0} \delta \widehat{\sigma}_0 \quad (9.28a)$$

$$+ \frac{1}{2} S^2 \frac{d^2 \Pi}{dS^2} \left(\left(\frac{\delta S}{S} \right)^2 - \widehat{\sigma}_0^2 \delta t \right) + \frac{1}{2} \widehat{\sigma}_0^2 \frac{d^2 \Pi}{d\widehat{\sigma}_0^2} \left(\left(\frac{\delta \widehat{\sigma}_0}{\widehat{\sigma}_0} \right)^2 - \nu^2 \delta t \right) \quad (9.28b)$$

$$+ S \widehat{\sigma}_0 \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \left(\frac{\delta S}{S} \frac{\delta \widehat{\sigma}_0}{\widehat{\sigma}_0} - \rho \nu \widehat{\sigma}_0 \delta t \right) \quad (9.28c)$$

From (9.25b) the ATMF skew \mathcal{S} is $\frac{\rho\nu}{2}$ – this is equivalent to the statement $\mathcal{R}_{T=0} = 2$. The last piece (9.28c) can thus be rewritten as:

$$S \widehat{\sigma}_0 \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \left(\frac{\delta S}{S} \frac{\delta \widehat{\sigma}_0}{\widehat{\sigma}_0} - \rho \nu \widehat{\sigma}_0 \delta t \right) = S \widehat{\sigma}_0 \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \left(\frac{\delta S}{S} \frac{\delta \widehat{\sigma}_0}{\widehat{\sigma}_0} - 2S \widehat{\sigma}_0 \delta t \right) \\ = S \widehat{\sigma}_0^2 \mathcal{S} \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \left(\frac{\frac{\delta S}{S} \delta \widehat{\sigma}_0}{S \widehat{\sigma}_0^2 \delta t} - 2 \right) \delta t \quad (9.29a)$$

$$= S \widehat{\sigma}_0^2 \mathcal{S} \frac{d^2 \Pi}{dS d\widehat{\sigma}_0} \left(\mathcal{R}_{T=0}^{r,\text{short}} - 2 \right) \delta t \quad (9.29b)$$

⁴This is not the case for vanilla options with strikes different than S_0 as for these options $\frac{d^2 f}{dS_0^2} = 0$ – the contribution of order one in T vanishes.

From (9.29a) it is apparent that, in defining the realized SSR, the short implied – rather than the realized – variance should be used in the denominator, in the definition of $\mathcal{R}_T^{r,\text{short}}$, hence expression (9.23) for $\mathcal{R}_T^{r,\text{short}}$.

9.10.2 The realized skew

The P&L above can be equivalently expressed as a difference between the implied skew \mathcal{S} and the instantaneous *realized* skew \mathcal{S}^r which we define as:

$$\mathcal{S}^r = \frac{1}{2\hat{\sigma}_0\delta t} \frac{\delta S}{S} \frac{\delta\hat{\sigma}_0}{\hat{\sigma}_0} = \left(\frac{\mathcal{R}_{T=0}^{r,\text{short}}}{2} \right) \mathcal{S} \quad (9.30)$$

Expression (9.30) shows that the relative mismatch of realized to implied skew is equal to that of realized to implied SSR. The cross gamma/theta P&L 9.29b reads:

$$2S\hat{\sigma}_0^2 \frac{d^2\Pi}{dSd\hat{\sigma}_0} (\mathcal{S}^r - \mathcal{S}) \delta t \quad (9.31)$$

Note that the P&L in (9.28) is not the total P&L incurred on our option position. Additional P&L is generated by daily recalibration of \mathcal{S} and \mathcal{C} . Only if this P&L is small and if the contributions in (9.28a) and (9.28b) are vanishing or negligible are we able to isolate the P&L of interest (9.29b) or equivalently (9.31).

9.10.3 Splitting the theta into three pieces

Expression (9.25a) for P is correct at order two in volatility of volatility: P does not exactly solve (9.24). Can we still use (9.25a) for P&L accounting? (9.24) expresses that the theta $\frac{dP}{dt}$ can be broken up in three pieces which match each of the second-order gamma contributions in (9.28).

In our model theta is equal to the Black-Scholes theta since $\hat{\sigma}(x)$ in (9.25b) does not depend on T . In the Black-Scholes model all of the theta is ascribed to the spot gamma – the second piece in the right-hand side of (9.28) – with a break-even level, the implied volatility, which is strike-dependent. In contrast, in the stochastic volatility model we are using, this theta is distributed across three gammas with break-even levels that are *not* strike-dependent. Checking whether (9.24) holds amounts to checking how well the following equality holds

$$\frac{\hat{\sigma}(x)^2}{2} S^2 \frac{d^2P_{BS}}{dS^2} = \frac{\hat{\sigma}_0^2}{2} S^2 \frac{d^2P}{dS^2} + \frac{\nu^2}{2} \hat{\sigma}_0^2 \frac{d^2P}{d\hat{\sigma}_0^2} + \rho\nu\hat{\sigma}_0^2 S \frac{d^2P}{dSd\hat{\sigma}_0} \quad (9.32)$$

where the three contributions in the right-hand side are called, respectively: spot theta, vol theta, cross theta.

Let us take the typical example of a one-month maturity smile with $\hat{\sigma}_0 = 20\%$, $\mathcal{S} = -0.7$, $\mathcal{C} = 0.4$. With these parameters, the implied volatilities of the 90%, 100%, 110% strikes are, respectively, 27.6%, 20%, 13.5%. From (9.26) we have: $\rho = -78.5\%$, $\nu = 178.3\%$.

Figure 9.5 shows on the left the three thetas as a function of K , for $S = 100$. The sum of these three thetas, along with the Black-Scholes theta – the left-hand side of (9.32) where a different implied volatility is used for each strike – is shown on the right.

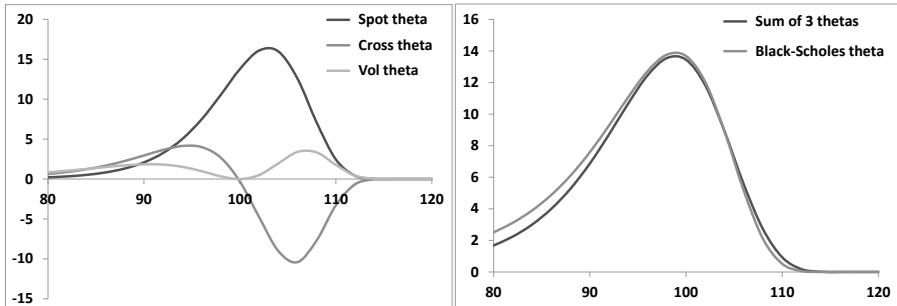


Figure 9.5: Left: the three terms in the right-hand side of (9.32) as a function of K , for $S = 100$. Right: the sum of the three thetas compared to the Black-Scholes theta $\frac{\hat{\sigma}(x)^2}{2} S^2 \frac{d^2 P_{BS}}{dS^2}$.

As we can see our approximation applied to the lognormal model does a decent job at splitting the Black-Scholes theta in three pieces, whose break-even levels are independent on the option's strike. The three thetas do not exactly add up to the Black-Scholes theta for strikes far out of the money, but the agreement is satisfactory for strikes between 95% and 105%, which is the range we use in our tests.

As is clear from Figure 9.5, the spot theta dominates by far – as our objective is to isolate the cross theta, we need to ensure that the spot theta of our position vanishes. The vol theta, on the other hand, is much smaller and almost cancels for a spread position.

9.10.4 Backtesting on the Euro Stoxx 50 index

We now use historical implied volatilities of the Euro Stoxx 50 index from April 2007 to March 2012 to backtest the following dynamical option trading strategy. We sell one-month options of strike 95% and buy the appropriate number of one-month options of strike 105% so that the spot gamma vanishes. We delta-hedge this position until the next (trading) day, when it is unwound and a new position is started. In order to approximately maintain a constant level of cross-gamma, we trade a constant notional – 100€ – of the 95%-strike option. For typical levels of 1-month market skew, for one 95% option sold, we need to buy about 0.5 options of strike 105%.

We use daily one-month implied volatilities for strikes 95%, 100%, 105% to determine the skew \mathcal{S} and curvature \mathcal{C} and back out ρ and ν using (9.26). These, together with the ATM volatility $\hat{\sigma}_0$, are fed to the lognormal model which we use

(a) to determine the ratio of the 105% options to 95% options, (b) to compute the delta.

Calibrated values of \mathcal{S} (multiplied by -10) and \mathcal{C} appear in the left-hand graph of Figure 9.6. \mathcal{C} is very noisy – indeed over the strike range [95%, 105%], the one-month smile is almost a straight line: in our tests we thus set the product \mathcal{C} equal to $\frac{2}{\sigma_0}$ – the precise value of \mathcal{C} used hardly affects our results since we only use 95% and 105% strikes.

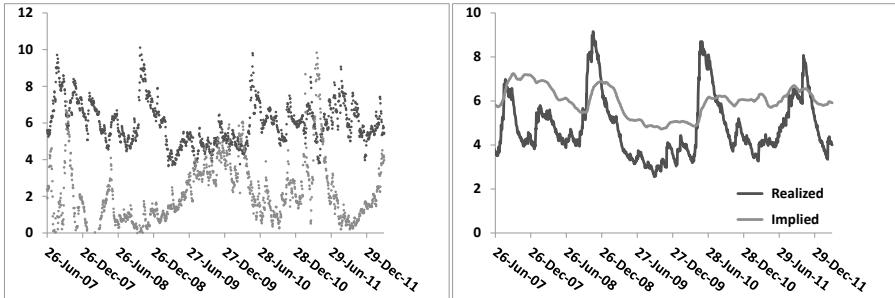


Figure 9.6: Left: daily values of skew \mathcal{S} (darker dots) and curvature \mathcal{C} (lighter dots) of one-month smiles of the Euro Stoxx 50 index. Right: 3-months exponentially weighted moving averages of the implied ATM skew \mathcal{S} and its realized counterpart $\mathcal{S}^{\text{real}}$ for a one-month maturity – see (9.30). Values of \mathcal{S} in both graphs, and of $\mathcal{S}^{\text{real}}$ have been multiplied by -10 to correspond approximately to the 95%/105% skew in volatility points.

The right-hand graph shows the implied and realized 1-month skew for the Euro Stoxx 50 index. In our historical sample, the average value of the *implied* 95%/105% skew is 6 points of volatility, while the average value of its *realized* counterpart – defined in (9.30) – is 4.8 points; this is 20% less than its implied counterpart.

Consequently, from (9.30), the average value of the realized SSR is about 20% lower than its implied value of 2; this agrees with the average value of realized SSR in Figure 9.4.

Our daily P&L between day i and day $i + 1$ is very plainly given by:

$$\begin{aligned} P\&L_{\text{Total}} &= [\Pi_i(t_{i+1}, S_{i+1}, \hat{\sigma}_{0i+1}, \mathcal{S}_{i+1}, \mathcal{C}_{i+1}) - \Pi_i(t_i, S_i, \hat{\sigma}_{0i}, \mathcal{S}_i, \mathcal{C}_i)] \\ &\quad - \frac{d\Pi_i}{dS_i}(S_{i+1} - S_i) \end{aligned}$$

where Π_i is the market value of the option portfolio purchased at time t_i . This P&L can be broken down in three contributions:

$$P\&L_{\text{Total}} = P\&L_{\text{Carry}}^{\text{vega-hedged}} + P\&L_{\text{Vega}} + P\&L_{\text{MtM}}^{\mathcal{S}, \mathcal{C}}$$

where:

$$\begin{aligned} P\&L_{\text{Carry}}^{\text{vega-hedged}} &= [\Pi_i(t_{i+1}, S_{i+1}, \hat{\sigma}_{0i+1}, \mathcal{S}_i, \mathcal{C}_i) - \Pi_i(t_i, S_i, \hat{\sigma}_{0i}, \mathcal{S}_i, \mathcal{C}_i)] \\ &\quad - \frac{d\Pi_i}{dS_i}(S_{i+1} - S_i) - \frac{d\Pi_i}{d\hat{\sigma}_{0i}}(\hat{\sigma}_{0i+1} - \hat{\sigma}_{0i}) \\ P\&L_{\text{Vega}} &= \frac{d\Pi_i}{d\hat{\sigma}_{0i}}(\hat{\sigma}_{0i+1} - \hat{\sigma}_{0i}) \\ P\&L_{\text{MtM}}^{S,C} &= \Pi_i(t_{i+1}, S_{i+1}, \hat{\sigma}_{0i+1}, \mathcal{S}_{i+1}, \mathcal{C}_{i+1}) - \Pi_i(t_{i+1}, S_{i+1}, \hat{\sigma}_{0i+1}, \mathcal{S}_i, \mathcal{C}_i) \end{aligned}$$

$P\&L_{\text{MtM}}^{S,C}$ is the mark-to-market P&L generated by recalibrating \mathcal{S} and \mathcal{C} to the market smile at time t_{i+1} .

At order one in $(t_{i+1} - t_i)$ and order two in $(S_{i+1} - S_i)$ and $(\hat{\sigma}_{0i+1} - \hat{\sigma}_{0i})$, $P\&L_{\text{Carry}}^{\text{Vega-hedged}}$ is the sum of the three gamma/theta contributions in (9.28b) and (9.28c).

Our option position has vanishing spot gamma by construction. Moreover, inspection of the left-hand graph of Figure 9.5 suggests that the volatility gamma/theta P&L is small: ideally P&L (9.28b) will be negligible so that $P\&L_{\text{Carry}}^{\text{Vega-hedged}}$ closely tracks the quantity of interest, that is the cross-gamma/theta P&L (9.28c) which we aim to single out. Our P&L is however polluted by the contributions of $P\&L_{\text{MtM}}^{S,C}$ and $P\&L_{\text{Vega}}$.

How well $(P\&L_{\text{Carry}}^{\text{vega-hedged}} + P\&L_{\text{MtM}}^{S,C})$, that is $(P\&L_{\text{Total}} - P\&L_{\text{Vega}})$ correlates to the cross-gamma/theta P&L is assessed in the left-hand graph of Figure 9.7.

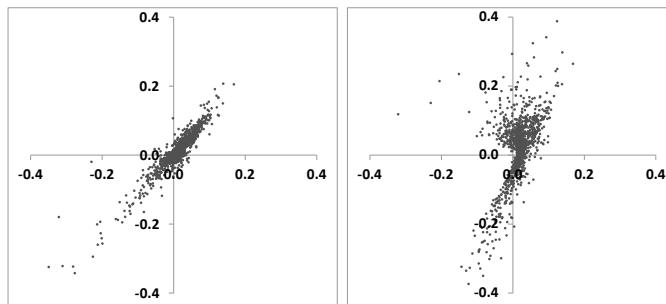


Figure 9.7: Scatter plots of daily values of $P\&L_{\text{Total}} - P\&L_{\text{Vega}}$ (left-hand graph) and $P\&L_{\text{Total}}$ (right-hand graph) as a function of P&L (9.28c).

$P\&L_{\text{MtM}}^{S,C}$ contributes in fact little noise; the reason we do not get a straight line is mostly due to higher-order terms in $P\&L_{\text{Carry}}^{\text{vega-hedged}}$. Still, the total daily P&L corrected for the vega contribution captures the daily cross gamma/theta P&L with acceptable accuracy.

Inclusion of the vega contribution – see right-hand graph – reduces the correlation of both P&Ls but, as we will now see, the cumulative vega P&L is small enough that our strategy's P&L is still mostly attributable to the cross gamma/theta P&L.

Figure 9.8 shows the cumulative P&Ls of our arbitrage strategy. That the bulk of the P&L is contributed by the cross gamma/theta P&L is manifest.

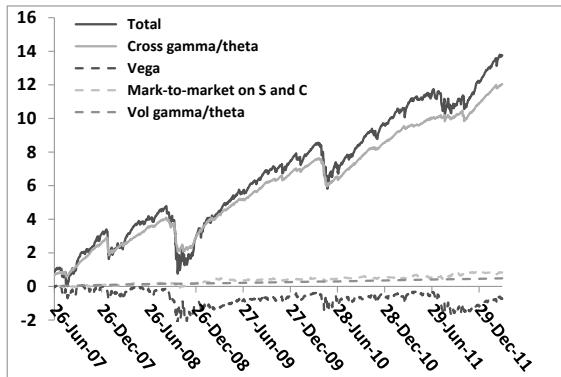


Figure 9.8: Cumulative P&Ls of the realized skew arbitrage strategy.

The volatility gamma/theta P&L, as well $P\&L_{\text{MtM}}^{S,C}$ and $P\&L_{\text{Vega}}$, are both small, even though the latter is very noisy. As a rule, sharp market moves generate simultaneous downward moves of the spot and upward moves of implied volatilities. Our gamma-neutral position has negative vega. On these days, our option position loses money both because of our positive cross-gamma and negative vega positions: this is confirmed by the simultaneous drops in the cumulative cross gamma/theta and vega P&Ls in Figure 9.8. These drops indeed coincide with peaks of the realized skew seen in Figure 9.6.

Remember that in our historical sample the *realized* ATM skew is on average 20% lower than the *implied* ATM skew – or equivalently that the realized SSR is 20% lower than its model-independent value of 2. Our position has vanishing spot gamma; moreover Figure 9.8 shows that the volatility gamma/theta P&L is small. Thus, almost all of the theta is generated by the cross theta. The cumulative theta P&L of our strategy is 45.7€. 20% of this, that is about 9€ is a number that is indeed consistent with the cumulative cross gamma/theta P&L.

9.10.5 The “fair” ATMF skew

In conclusion, just as the difference between realized and implied volatility can be arbitrated – that is materialized as the P&L of an option strategy – for short maturities, the difference between the realized value of the SSR and its model-independent value of 2 – or equivalently the difference between the realized ATM skew and the implied ATM skew – can be materialized, to a good approximation, as the P&L of a dynamical option strategy.⁵

⁵In practice, we would presumably buy a little more 105% strike options in order to reduce our vega position. This would generate a positive spot gamma position – which helps in case of large market

There is then no inconsistency in the fact that $\mathcal{R}^{\text{real}} < 2$. The short ATM skew quantifies the *implied* value of the instantaneous spot/volatility covariance, which may differ from its *realized* value. In our sample, it is on average 20% lower than its *implied* counterpart. Our trading strategy approximately materializes this difference as a P&L.

We have shown that this is equivalent to materializing as a P&L the difference between implied skew and realized skew:

$$\mathcal{S}^r = \frac{1}{2\hat{\sigma}_0\delta t} \left\langle \frac{\delta S}{S} \frac{\delta\hat{\sigma}_0}{\hat{\sigma}_0} \right\rangle$$

Note that the fair level of the ATM skew is not determined by the covariance of the spot and *realized* volatility; rather it is given by the covariance of spot and ATM *implied* volatility, a circumstance that may surprise at first.

Imagine there is no options' market – our only hedge instrument is the underlying itself – and we are asked to quote “fair” vanilla option prices, “fair” meaning that we do not make or lose money on average. We then need to model the process of the instantaneous historical realized variance – or equivalently the process of expected future historical variances, that is forward variances. We use pricing equation (7.4), which expresses that, on average no money is made or lost as we delta-hedge our option. The only difference now is that ξ^τ is no longer a market VS forward variance, but instead the expected future realized instantaneous variance.⁶

The results of Chapter 8 apply: at order one in volatility of volatility the “fair” ATM skew is given by the weighted integral of the *historical* covariance of spot and future (*expected*) *realized* variance.⁷

Now assume that there exists instead a market of ATM options – or variance swaps. Just as in the above backtest, we use these options to cancel the spot gamma, thus removing the sensitivity to *realized* volatility.

Since the vanilla hedge has to be dynamically readjusted, our position becomes sensitive to the joint dynamics of the spot and *implied* volatilities. Practically, this is

moves but costs some theta otherwise. Using the 100% strike along with the 95% and 105% strikes in order to cancel both gamma and vega does not work, as the resulting position is long the 95% and 105% strikes and short the 100% strike: this is mostly a volatility gamma/theta position. Also, unwinding and restarting a new position on a daily basis is not practical: in our backtest, factoring in a bid/offer spread of 0.2 points of volatility on each leg of our spread position wipes out the strategy's P&L.

⁶Interestingly the ξ^τ are still driftless. The ξ^τ in (7.4) are driftless because they can be delta-hedged by taking a position in market instruments (VSs) that require no financing. In the present context, the ξ^τ are driftless in the historical probability measure, just because they are expectations (of future realized variances).

⁷One can also avoid modeling variances altogether by using the hedged Monte Carlo technique of Bouchaud, Potters and Sestovic – see [17], or equivalently, a method proposed by Bruno Dupire. It consists in (a) using consecutive sequences of historical returns as Monte Carlo paths, (b) simulating the daily delta-hedging – at a given implied volatility – of a vanilla option, (c) finding the implied volatility such that the average of the final payoff minus the P&L from the delta hedge equals the Black-Scholes price. Underlying this technique is the unstated – and strong – assumption of stationarity: we are averaging over different volatility regimes and the conditionality on spot level or past return history is lost.

materialized, at order one in volatility of volatility, as the cross gamma of spot and *implied* volatility.

9.10.6 Relevance of model-independent properties

It is not clear that other model-independent properties can be established, without resorting to more or less reasonable additional assumptions. For example, Peter Carr and Roger Lee show in [27] that if implied volatilities of power payoffs⁸ are uncorrelated with the spot process, the density of the realized quadratic variation up to T can be extracted from the vanilla smile of maturity T , hence payoffs on realized variance can be replicated by dynamical trading in vanilla options.

Are these model-independent rules practically relevant? This is assessed by studying whether a violation of these rules can be arbitAGED, i.e. materialized as a P&L. It is not clear in particular that a violation of (9.13) could be practically arbitAGED.

This issue is connected to the general question of the practical relevance of calibration. Our study of the arbitrage of the short SSR is a sobering illustration of how difficult it can be to lock the value of a model parameter – in our case the covariance of spot and ATM implied volatility – by dynamically trading vanilla options.

Only when one is able to do so does it make sense to entrust a model with the task of backing the value of a dynamical parameter out of the vanilla smile.

9.11 Conclusion

Because the skew in stochastic volatility models is generated by the covariance of spot and forward variances, some features of the underlying spot/variance dynamics can be recovered from the resulting smile. Provided some reasonable assumptions hold – in particular time homogeneity – for a flat term structure of VS volatilities, at order one in the volatility of volatility, the Skew Stickiness ratio is bounded above and below:

$$\mathcal{R}_T \in [1, 2]$$

We also show that \mathcal{R}_T – a quantity that characterizes the model’s dynamics – is related to the scaling of the ATMF skew \mathcal{S}_T with maturity a static property of the smile. For short maturities, \mathcal{R}_T tends to the universal value of 2. For long maturities its scaling depends on the characteristic exponent γ of the underlying

⁸See Section 4.3 for the definition of power payoffs. In the derivation of [27] the assumption of no correlation between S and the instantaneous volatility is made. What matters however from a trading point of view is that S and implied volatilities of power payoffs be uncorrelated. Only then does the spot/implied volatility cross-gamma P&L vanish.

spot/variance covariance function. Stochastic volatility models fall into one of two classes depending on the value of γ .

- Type I models are such that $\gamma > 1$: for large T , $\mathcal{S}_T \propto \frac{1}{T}$ and $\mathcal{R}_T \rightarrow 1$
- Type II models are such that $\gamma < 1$: for large T , $\mathcal{S}_T \propto \frac{1}{T^\gamma}$ and $\mathcal{R}_T \rightarrow 2 - \gamma$

Thus, Type II models allow for a slower decay of the ATMF skew – which is consistent with market skews, for which $\gamma \simeq \frac{1}{2}$ – and a non-trivial value of the long-maturity limit of the SSR, which, again, is compatible with the observed dynamics of market smiles. Moreover, the characteristic exponent of the spot/variance covariance function can be backed out of the scaling of \mathcal{S}_T and the long-maturity limit of \mathcal{R}_T .

This connection between \mathcal{S}_T and \mathcal{R}_T is a distinguishing feature of time-homogeneous stochastic volatility models – it is summarized by the following formula: for long maturities,

$$\mathcal{S}_T \propto \frac{1}{T^{2-\mathcal{R}_\infty}}$$

In time-homogeneous Jump-Lévy models, by contrast, while for long maturities $\mathcal{S}_T \propto \frac{1}{T}$, \mathcal{R}_T vanishes for all T .

While putting together a genuine Type II model is difficult – because one needs to reconcile a non-trivial scaling of the spot/variance covariance function with the requirement of a low-dimensional Markov representation – Type II behavior can be achieved for a decent range of maturities in a model driven by simple Ornstein–Uhlenbeck processes.

We have provided evidence that suitable parametrization of a simple two-factor model generates Type II scaling for a range of maturities that is practically relevant – while also enforcing the desired scaling of volatilities of volatilities.

Having analyzed how dynamical properties of spot and time-homogeneous stochastic volatility models are related to the smile they produce, the following questions naturally arise:

- How do they compare with dynamical properties of a local volatility model calibrated to the same smile?
- How do they compare with the *realized* behavior of spot and implied volatilities?

9.11.1 SSR in local and stochastic volatility models – and in reality

The behavior of \mathcal{R}_T as a function of T is structurally different in stochastic volatility models than in the local volatility model. In either Type I or Type II stochastic volatility models, the SSR starts from 2, then *decreases* either towards 1 or towards a non-trivial value in $[1, 2]$. In the local volatility model, instead, the SSR

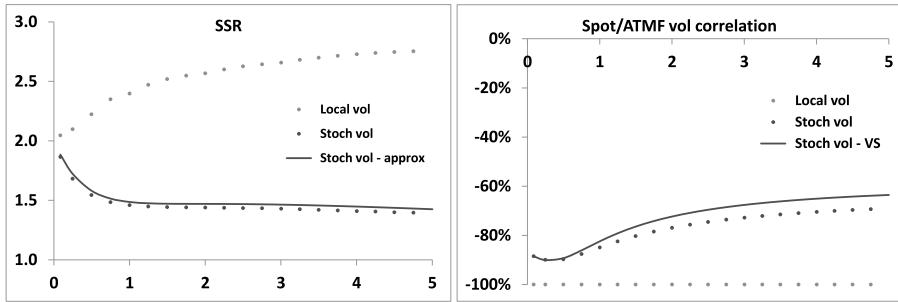


Figure 9.9: Left: the SSR as a function of maturity (years) in the two-factor stochastic volatility model (dark dots) and in the local volatility model (light dots) calibrated on the smile generated by the stochastic volatility model. Parameters in Table 8.2, page 329, have been used. The solid line corresponds to the SSR of the stochastic volatility model calculated using expression (9.21), at order one in volatility of volatility. Right: correlation of spot and ATMF volatilities. The solid line corresponds to the correlation of spot and VS – rather than ATMF – volatilities in the two-factor model, which is easily evaluated exactly.

starts from 2 for short maturities, then *increases* for longer maturities, for typical equity index smiles – see Figure 2.4, page 59.

This is illustrated in Figure 9.9. We have used the parameters in Table 8.2 and a flat term structure of VS volatilities at 20% (same parameters as those used in Figure 9.1). The curves in Figure 9.9 are thus obtained with the same vanilla smile – shown in Figure 8.2, page 330, for select maturities.

In both models the SSR starts from the model-independent value of 2 for short maturities. Equivalently, the implied regression coefficient of the short ATMF volatility on the spot is model-independent and given by the ATMF skew.

Parameters in Table 8.2 are such that the ATMF skew approximately decays with the characteristic exponent $\gamma = \frac{1}{2}$. For the longer maturities in our graph, the SSR tends to 1.5 in the stochastic volatility model (approximately $2 - \gamma$), while for the local volatility model it tends to a value close to 3.

For a smile such that the ATMF skew decays as a power law, the approximate expression (2.81), page 56, of \mathcal{R}_∞ in the local volatility model:

$$\mathcal{R}_\infty = \frac{2 - \gamma}{1 - \gamma}$$

indeed yields $\mathcal{R}_\infty = 3$ for $\gamma = \frac{1}{2}$.

The spot/ATMF volatility covariance

The SSR is a useful indicator as it measures the implied covariance of $\ln S$ and the ATMF volatility $\hat{\sigma}_{F_T T}$ – in units of the ATMF skew – generated by the pricing

model. Spot/volatility cross-gammas are one of the main risks of exotic equity payoffs.⁹ Rather than directly comparing the implied spot/volatility covariance with its realized counterpart, we can convert the *realized* covariance of $\ln S$ and $\widehat{\sigma}_{F_T T}$ into a *realized* SSR.

Model-generated and realized SSR can then be compared to assess whether the pricing model is sufficiently conservative.

The realized SSR for the Euro Stoxx 50 and S&P 500 indexes, for $T = 2$ years appear in Figure 9.3 – the estimator for the realized SSR is given by expression (9.22). For both indexes, the average realized SSR is around 1.5. For other indexes, the realized SSR can be very different. Figure 9.10 shows historical values of the realized 6-month value of \mathcal{R}_T for $T = 1$ year, for the S&P 500 and Nikkei indexes. While the SSR for the S&P 500 is fairly stable around 1.5, for the Nikkei it reaches at times very negative values.

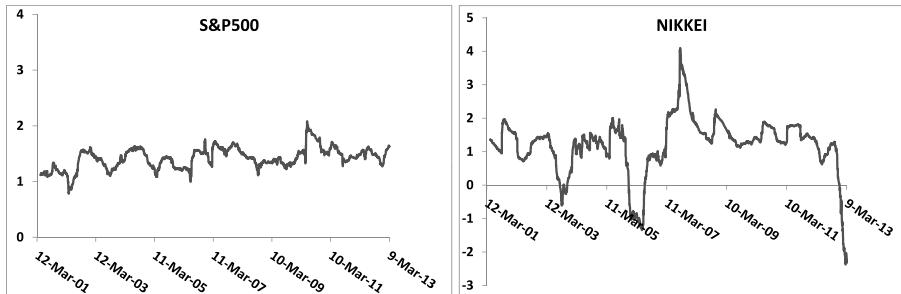


Figure 9.10: The sliding 6-month realized SSR for $T = 1$ year, for the S&P 500 (left) and Nikkei (right) indexes.

The large negative dip of the SSR of the Nikkei at the end of 2012 can be traced to the impact of the vega hedging of autocalls on the liquidity of the Nikkei vanilla option market.

Autocalls¹⁰ provide the buyer a negative volatility exposure which vanishes when the spot price goes above an upper threshold, whereupon the option expires. Upon selling an autocall the dealer hedges his/her vega position by selling vanilla

⁹What matters is the cross-gamma of the hedged position – see the discussion in Section 1.3.

¹⁰Consider an autocall of notional N on an underlying S . At inception the autocall buyer pays N . Periodically – say every 3 months – the buyer receives a fixed coupon proportional to N provided the spot value is below a given threshold, typically 105% of the initial spot value. If, on a coupon date, the spot value is above this threshold, the buyer receives his coupon, gets N back, and the autocall expires. At maturity T the buyer gets back N unless at some point during the autocall's life the spot crosses a lower barrier, typically 60% of the initial spot value, in which case the payoff at maturity is $N \min(\frac{S_T}{S_0}, 1)$.

Autocalls are typically written on an index or on the worst of 3/4 stocks: $S_t \equiv \min_i (S_t^i)$. The (large) coupon of an autocall compensates the buyer for the risk that he may not recover the invested notional, if the underlying drops significantly.

options. The presence of the upper threshold generates a large spot/volatility cross-gamma: as the spot moves up, the dealer needs to buy back vanilla options. In normal circumstances, as the spot moves up, implied volatilities decrease, and the unwinding of the vega hedge should generate on average positive cross-gamma P&L. This cross-gamma P&L is offset by a theta contribution which is quantified by the SSR of the pricing model:

$$\begin{aligned} P\&L &= -S\widehat{\sigma}_T \frac{d^2\Pi}{dSd\widehat{\sigma}_T} \left(\frac{\delta S}{S} \frac{\delta\widehat{\sigma}_T}{\widehat{\sigma}_T} - \mathcal{R}_T \mathcal{S}_T \frac{\sigma^2}{\widehat{\sigma}_T} \delta t \right) \\ &= -SS_T \frac{d^2\Pi}{dSd\widehat{\sigma}_T} (\mathcal{R}_T^r \sigma_r^2 - \mathcal{R}_T \sigma^2) \delta t \end{aligned} \quad (9.33)$$

Figure 9.9 illustrates how different the SSR of the local volatility model and the SSR of a stochastic volatility model are. When carrying a long spot/volatility cross-gamma position, it thus seems preferable to price with the local volatility model, while with a short spot/volatility cross-gamma position, a stochastic volatility model should be used.

We refer the reader to a similar discussion in Section 12.6 of Chapter 12, page 482, in the more general context of mixed local-stochastic volatility models.

9.11.2 Volatilities of volatilities

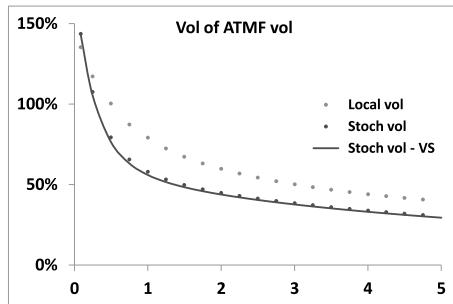


Figure 9.11: Volatility of ATMF volatilities as a function of maturity (years) in the two-factor stochastic volatility model (dark dots) and in the local volatility model (light dots) calibrated on the smile generated by the stochastic volatility two-factor model. Parameters in Table 8.2 have been used. The solid line corresponds to volatilities of VS – rather than ATMF – volatilities in the two-factor model calculated exactly using expression (7.39), page 227.

What about volatilities of volatilities? Volatilities of ATMF volatilities for both models appear in Figure 9.11. The term structure of volatilities of volatilities of the two-factor stochastic volatility model has been examined in Section 7.4. As for the

local volatility model, because $\widehat{\sigma}_{F_T T}$ is a function of S , the volatility of $\widehat{\sigma}_{F_T T}$ can be expressed using the SSR. The instantaneous lognormal volatility of $\widehat{\sigma}_{F_T T}$ is given by (see Section 2.5.5):

$$\text{vol}(\widehat{\sigma}_{F_T T}) = \mathcal{R}_T \mathcal{S}_T \frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}}$$

where the short ATMF volatility $\widehat{\sigma}_{F_0 0}$ is also equal to the instantaneous volatility of S .

For $T \rightarrow 0$ \mathcal{R}_T tends to 2 and the volatility of the short ATMF volatility is twice the ATMF skew – we recover result (2.85) of Section 2.5.5.

For longer maturities, one can use the approximate expression (2.83):

$$\text{vol}(\widehat{\sigma}_{F_T T}) \simeq \left(\mathcal{S}_T + \frac{1}{T} \int_0^T \mathcal{S}_\tau d\tau \right) \frac{\widehat{\sigma}_{F_0 0}}{\widehat{\sigma}_{F_T T}}$$

which only involves the term structure of the ATMF skew.

Again, in contrast with the two-factor model, the local volatility model is not time-homogeneous; volatilities of volatilities depend not only on $T - t$, but also on t . Typically, volatilities of volatilities in the local volatility model are larger for small t /smaller for large t , compared with those of a time-homogeneous stochastic volatility model calibrated on the same smile.

9.11.3 Carry P&L of a partially vega-hedged position

Consider a delta- and vega-hedged option position. Assume that there is only one implied volatility $\widehat{\sigma}$, in addition to S . The gamma/theta P&L of the hedged position reads:¹¹

$$P\&L = -\frac{1}{2} \frac{d^2 P}{dS^2} (\delta S^2 - \bullet \delta t) - \frac{1}{2} \frac{d^2 P}{d\widehat{\sigma}^2} (\delta \widehat{\sigma}^2 - \bullet \delta t) - \frac{d^2 P}{dS d\widehat{\sigma}} (\delta S \delta \widehat{\sigma} - \bullet \delta t)$$

The suitability of the model parametrization is assessed by comparing the model-generated values of the spot/volatility and volatility/volatility covariances – denoted by $\bullet \delta t$ – with their realized counterparts, the latter evaluated using historical data for S and $\widehat{\sigma}$.

In practice delta-hedging is performed daily – *and delta is calculated using current market implied volatilities* – but vega-hedging may not, because of larger bid/offer costs. Vega-hedging and delta-hedging occur asynchronously. Which second order spot gamma, spot/volatility cross-gamma and volatility/volatility cross-gamma P&Ls does one then materialize?¹²

¹¹Typically, either a term structure of implied volatilities $\widehat{\sigma}_T$ – in the case of the two-factor model – or all vanilla implied volatilities $\widehat{\sigma}_{KT}$ – when using the local volatility model or the local-stochastic volatility models of Chapter 12 – are taken as underliers in addition to S , and contribute to the P&L.

¹²This issue calls to mind that of spot/spot cross-gammas for underlyings trading in different time zones – see [12].

Imagine we vega-hedge our option position periodically every n days – say one week ($n = 5$). We make the reasonable assumption that gammas and cross-gammas are constant during this period.

Rather than plodding through the calculation of this P&L, we choose a more enlightening route: start from the P&L generated by hedging once per period, then remove terms that can be offset by daily delta-hedging. We denote by $\delta\hat{\sigma}_i$ and δS_i the respective daily increments over $[t_{i-1}, t_i]$.

Spot gamma

Assuming we delta-hedge only once per period, the spot gamma P&L reads:

$$\begin{aligned} \frac{1}{2} \frac{d^2P}{dS^2} (\Sigma_i \delta S_i)^2 &= \frac{1}{2} \frac{d^2P}{dS^2} (\Sigma_i \delta S_i) (\Sigma_j \delta S_j) \\ &= \frac{1}{2} \frac{d^2P}{dS^2} (\Sigma_i \delta S_i^2) + \frac{d^2P}{dS^2} \Sigma_i (\Sigma_{j < i} \delta S_j) \delta S_i \end{aligned}$$

The prefactor in front of δS_i in the second piece involves spot increments that precede δS_i , thus are known at time t_{i-1} . This P&L linear in δS_i can then be cancelled by a delta strategy $\Delta_{t_{i-1}} \delta S_i$ with $\Delta_{t_{i-1}} = \frac{1}{2} \frac{d^2P}{dS^2} (\Sigma_{j < i} \delta S_j)$. Only the first piece remains and we get the usual result for the spot gamma P&L over one period:

$$\frac{1}{2} \frac{d^2P}{dS^2} (\Sigma_i \delta S_i^2)$$

It involves the spot variance, measured using daily increments – as expected.

Volatility gamma

Since no vega-hedging happens during the n -day period, the volatility gamma P&L is simply:

$$\frac{1}{2} \frac{d^2P}{d\hat{\sigma}^2} (\Sigma_i \delta\hat{\sigma}_i)^2$$

that is, the variance of volatility is sampled according to our weekly schedule – as expected.

Spot/volatility cross-gamma

The spot/volatility cross-gamma P&L reads:

$$\frac{d^2P}{dSd\hat{\sigma}} (\Sigma_i \delta S_i) (\Sigma_j \delta\hat{\sigma}_j) = \frac{d^2P}{dSd\hat{\sigma}} \Sigma_i (\Sigma_{j < i} \delta\hat{\sigma}_j) \delta S_i + \frac{d^2P}{dSd\hat{\sigma}} \Sigma_i (\Sigma_{j \geq i} \delta\hat{\sigma}_j) \delta S_i$$

The first portion of the right-hand side can be offset by a delta position; our final P&L over one period reads:

$$\frac{d^2P}{dSd\hat{\sigma}} \Sigma_i (\Sigma_{j \geq i} \delta\hat{\sigma}_j) \delta S_i \tag{9.34}$$

which involves the product of spot increments and all subsequent volatility increments.

The spot/volatility estimator in (9.34) is what should be used to measure realized spot/volatility covariance, for the sake of comparing realized and model-implied levels. How does it differ from the usual estimator that applies to the situation of synchronous delta and vega-hedging?

Imagine that during the vega-hedging period S and $\hat{\sigma}$ vary by δS , $\delta\hat{\sigma}$ and that both experience a trend: $\delta S_i = \frac{\delta S}{n}$, $\delta\hat{\sigma}_i = \frac{\delta\hat{\sigma}}{n}$.

Then, using the above expressions, the spot gamma P&L is equal to $\frac{1}{n} \left(\frac{1}{2} \frac{d^2 P}{dS^2} \delta S^2 \right)$ – that is delta-hedging has reduced it by a factor $\frac{1}{n}$ – and the volatility gamma P&L is $\frac{1}{2} \frac{d^2 P}{d\hat{\sigma}^2} \delta\hat{\sigma}^2$.

As for the cross-gamma P&L, it is equal to $\frac{n+1}{2n} \left(\frac{d^2 P}{dS d\hat{\sigma}} \delta S \delta\hat{\sigma} \right) \simeq \frac{1}{2} \frac{d^2 P}{dS d\hat{\sigma}} \delta S \delta\hat{\sigma}$, for n large.

Thus, using current implied volatilities in the calculation of the delta – even without trading any intermediate vega hedges – has reduced the realized cross-gamma P&L by a factor $\simeq \frac{1}{2}$.

Chapter's digest

9.1 The ATMF skew

- At order one in volatility of volatility, the ATMF skew is given, as a function of the spot/variance covariance function by:

$$\mathcal{S}_T = \frac{1}{2\sqrt{T}} \frac{1}{\left(\int_0^T \xi_0^\tau d\tau\right)^{3/2}} \int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du$$



9.2 The Skew Stickiness Ratio (SSR)

- The SSR is defined as the instantaneous regression coefficient of the ATMF volatility on $\ln S$, in units of the ATMF skew:

$$\mathcal{R}_T = \frac{1}{\mathcal{S}_T} \frac{E[d \ln S d\hat{\sigma}_{F_T(S)T}]}{E[(d \ln S)^2]}$$

In jump-diffusion models $\mathcal{R}_T = 0, \forall T$. In the local volatility model, since the ATMF volatility is a function of S : $\mathcal{R}_T = \frac{1}{\mathcal{S}_T} \frac{d\hat{\sigma}_{F_T(S)T}}{d \ln S}$.

- At order one in volatility of volatility, \mathcal{R}_T is given, as a function of the spot/variance covariance function, by:

$$\mathcal{R}_T = \frac{\int_0^T \xi_0^\tau d\tau}{T \xi_0^0} \frac{T \int_0^T \mu(0, u, \xi_0) du}{\int_0^T d\tau \int_\tau^T \mu(\tau, u, \xi_0) du}$$



9.3 Short-maturity limit of the ATMF skew and the SSR

- For $T \rightarrow 0$, $\mathcal{S}_0 = \frac{\mu(0,0,\xi_0)}{4(\xi_0^0)^{3/2}}$ and $\mathcal{R}_0 = 2$. The short-maturity limit of the SSR is 2, as in the local volatility model.



9.4 Model-independent range of the SSR

- Assuming the VS term structure is flat and the spot/variance covariance function is time homogeneous – $\mu(\tau, u, \xi_0) \equiv \mu(u - \tau)$ – one derives the following

expressions for ATMF skew and SSR:

$$\begin{aligned}\mathcal{S}_T &= \frac{1}{2\xi_0^{3/2}T^2} \int_0^T (T-t)\mu(t)dt \\ \mathcal{R}_T &= \frac{\int_0^T \mu(t)dt}{\int_0^T (1-\frac{t}{T})\mu(t)dt}\end{aligned}$$

Making the assumption that $\mu(t)$ decays monotonically towards zero as $t \rightarrow \infty$, we get model-independent lower and upper bounds on \mathcal{R}_T :

$$\mathcal{R}_T \in [1, 2]$$

Because of its definition the SSR is inversely proportional to the short VS volatility. One should bear in mind this dependence when assessing SSRs of market smiles.



9.5 Scaling of ATMF skew and SSR – a classification of models

► Depending on the rate of decay of $\mu(t)$ for $t \rightarrow \infty$, two long-maturity regimes for the ATMF skew and SSR can be defined, which lead to the division of stochastic volatility models into two classes. Assuming that $\mu(t) \propto \frac{1}{t^\gamma}$ for $t \rightarrow \infty$, for long maturities:

- (Type I) If $\gamma > 1$:

$$\mathcal{S}_T \propto \frac{1}{T} \quad \text{and} \quad \lim_{T \rightarrow \infty} \mathcal{R}_T = 1 \quad (9.35)$$

- (Type II) If $\gamma < 1$:

$$\mathcal{S}_T \propto \frac{1}{T^\gamma} \quad \text{and} \quad \lim_{T \rightarrow \infty} \mathcal{R}_T = 2 - \gamma \quad (9.36)$$

Exponential decay of $\mu(t)$ produces Type I behavior. In Type I models, in the long-maturity regime, \mathcal{S}_T and \mathcal{R}_T bear no signature of the rate of decay of μ .

For both types of models, the long-maturity ATMF skew and SSR are related through:

$$\mathcal{S}_T \propto \frac{1}{T^{2-\mathcal{R}_\infty}}.$$



9.6 Type I models – the Heston model

► The Heston model produces Type I behavior. Its long-maturity SSR is 1, which can be checked using the order-one expansion of the ATMF skew in Chapter 6.



9.7 Type II models

► Even though the two-factor model is strictly of Type 1, since $\mu(t)$ decays exponentially for large t , we can still achieve Type II scaling of the ATMF skew – and the corresponding value of the SSR – on a range of maturities that is practically relevant.

For a flat term-structure of VS volatilities, at order one in volatility of volatility, both quantities are given by:

$$\begin{aligned} S_T &= \nu \alpha_\theta \left[(1-\theta) \rho_{SX^1} \frac{k_1 T - (1 - e^{-k_1 T})}{(k_1 T)^2} + \theta \rho_{SX^2} \frac{k_2 T - (1 - e^{-k_2 T})}{(k_2 T)^2} \right] \\ R_T &= \frac{(1-\theta) \rho_{SX^1} \frac{1-e^{-k_1 T}}{k_1 T} + \theta \rho_{SX^2} \frac{1-e^{-k_2 T}}{k_2 T}}{(1-\theta) \rho_{SX^1} \frac{k_1 T - (1 - e^{-k_1 T})}{(k_1 T)^2} + \theta \rho_{SX^2} \frac{k_2 T - (1 - e^{-k_2 T})}{(k_2 T)^2}} \end{aligned}$$

► The realized behavior of equity index smiles is consistent with Type II.



9.8 Numerical evaluation of the SSR

► The SSR of the two-factor model is easily evaluated numerically in a Monte Carlo simulation by simply shifting the initial values of processes X_t and Y_t .



9.9 The SSR for short maturities

► The value of the realized SSR of short-maturity equity index smiles is usually substantially lower than the model-independent value of stochastic volatility models. Can this difference be materialized as the P&L of a trading strategy?



9.10 Arbitraging the realized short SSR

► We use a lognormal model for the short ATM volatility. In our two asset-model – spot and short ATM volatility – the difference between the realized SSR and its model-independent value of 2 is materialized as a cross-gamma/theta P&L.

► We backtest a dynamical delta-hedged option strategy on the Euro Stoxx 50 index that consists in maintaining a short-skew position around the money. The resulting P&L approximately captures the cross-gamma/theta P&L corresponding to the difference between implied and realized skew, even though the residual vega of our position impacts the cumulative P&L negatively on large downward moves of the spot.



9.11 Conclusion

► The behavior of \mathcal{R}_T as a function of T is structurally different in stochastic volatility models than in the local volatility model. In the local volatility model, for typical equity index smiles, the SSR starts from 2 for short maturities, then *increases* for longer maturities.

For an ATMF skew that decays algebraically with exponent γ , in the local volatility model $\mathcal{R}_\infty = \frac{2-\gamma}{1-\gamma}$, while in a stochastic volatility model, $\mathcal{R}_\infty = 2 - \gamma$. For the typical value $\gamma = \frac{1}{2}$, the SSR of the local volatility model for long maturities is $\mathcal{R}_\infty = 3$, compared to $\mathcal{R}_\infty = 1.5$ for a stochastic volatility model.

► Instantaneous volatilities of volatilities of long-dated vanilla options are also larger in the local volatility model than in a stochastic volatility model.

► In practice, vega hedging occurs less frequently than delta hedging, thus the spot/volatility cross/gamma-theta P&L is not materialized exactly. Assuming that delta-hedging is performed daily, using actual implied volatilities, while vega-hedging is performed less frequently, we obtain the approximate result that, in case of a trend in both spot and implied volatility in between two vega rehedges, half of the cross-gamma P&L is materialized.

This page intentionally left blank