Smile Dynamics IV

Lorenzo Bergomi
Société Générale
lorenzo.bergomi@sgcib.com

June 2009

Abstract

In this paper we address the relationship between the smile that stochastic volatility models produce and the dynamics they generate for implied volatilities. We introduce a new quantity, which we call the Skew Stickiness Ratio and show how, at order one in the volatility of volatility, it is linked to the rate at which the at-the-money-forward skew decays with maturity. We then focus on short maturity skews and (a) show that the difference between realized and implied SSR can be materialized as the P&L of an option strategy, (b) introduce the notion of *realized skew.*

1. Introduction

In previous work (Bergomi, 2004, 2005, 2008) we studied the dynamical properties of popular smile models and proposed a new framework for specifying stochastic volatility models with the objective of controlling some of their dynamical properties such as (a) the term-structure of the volatilities of volatilities, (b) the level of short forward skew, (c) the smile of volatility of volatility.

While these issues are mostly relevant for pricing and risk-managing exotic options, the subject of the joint dynamics of spot and implied volatilites has wider relevance, both for managing exotic and vanilla books.

Stochastic volatility models can be assessed either synchronically, by looking at the strike and maturity-dependence of the smile they produce or, diachronically, by studying the dynamics of volatilities they generate. How are these two aspects of a model related ? Is one a reflection of the other and is this connection quantifiable ? If so, in case a violation of this relationship is observed on market smiles, can it be arbitraged ?

These are the issues we address in this article, for general stochastic volatility models based on diffusion processes. In the first section of this paper we derive a relationship, at first order in the volatility of volatility, linking two features – one static, one dynamic – of stochastic volatility models:

- the rate at which the ATMF skew decays with maturity
- the rate at which the ATMF volatility moves when the spot moves

which will prompt us to introduce a new quantity: the Skew Stickiness Ratio (SSR).

In the second section we address the issue of practically materializing the P&L resulting from a difference between implied and realized SSR, focusing on short maturities.This leads us to the introduction of an estimator for the ATMF skew, the *realized* skew.

2. The skew stickiness ratio (SSR)

2.1. The vanilla ATM skew

Let us assume a general stochastic volatility model driven by Brownian motion. Quite generally, the dynamics in a stochastic volatility model can be formulated using as basic objects forward variances ξ^T: ξ_t^T is the instantaneous variance for date T, observed at time t: $\xi_t^T = \frac{d}{dT}\left((T-t)\,\widehat{\sigma}_{tT}^2\right)$ where $\widehat{\sigma}_{tT}$ is the implied variance swap (VS) volatility for maturity T, observed at time t. The ξ_t^T have no drift.

The dynamics of the ξ^T may or may not have a low-dimensional Markov representation. While, for example, the Heston model allows for a one-dimensional representation built on the instantaneous variance, we can imagine an extreme case where the dynamics of the variance curve is driven by a Brownian sheet: each ξ^T is driven by its own Brownian motion.

Let us write a general stochastic volatility model driven by a diffusion process as:

$$\begin{aligned} dS_t^\omega &= (r-q)\, S_t^\omega dt + \sqrt{\xi_t^t} S_t^\omega dZ_t \\ d\xi_t^T &= \omega \sum_{i=1}^{n} \xi_t^T \lambda_{it}^T dW_t^i \end{aligned} \tag{2.1}$$

where Z_t is a Brownian motion, W_t is a vector of n Brownian motions – all possibly correlated – and ω is a common scale factor for volatilities of volatilities λ_{it}^T which may depend very generally on the curve ξ_t and time, but not on S_t. Without loss of generality we factor ξ_t^T out of λ_{it}^T . When $\omega = 0$, volatilities are not stochastic anymore: $\xi_t^T = \xi_0^T$, where ξ_0 is the variance curve calibrated at $t = 0$. Let us expand ξ_t^T at first order in ω. We have

$$\xi_t^T = \xi_0^T \left(1 + \omega \int_0^t \sum_{i=1}^{n} \left(\lambda_{i\tau}^T\right)^0 dW_\tau^i \right)$$

where $\left(\lambda_{i\tau}^T\right)^0$ is evaluated in the unperturbed state ($\omega = 0$), in which forward variances are frozen and S_t^0 is lognormal. We now derive an expression of the ATMF skew for maturity T, $\mathcal{S}_T$, at first order in ω by evaluating the skewness of $x_T = \ln\left(\frac{S_T}{F_T}\right)$ and using the well-known approximation relating the ATMF skew to the skewness s_T of x_T[1]:

$$\mathcal{S}_T = \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_F = \frac{s_T}{6\sqrt{T}}$$

For the sake of analytical tractability, we assume that $\left(\lambda_{i\tau}^T\right)^0$ does not depend on S_t: this restricts our analysis to pure stochastic volatility models with no local volatility component.

s_T is given by: $s_T = M_3^T \,/\, \left(M_2^T\right)^{\frac{3}{2}}$ where $M_i^T = \left\langle \left(x_T - \langle x_T \rangle\right)^i \right\rangle$ and $\langle X \rangle$ denotes $E[X]$. Let us denote by $\delta\xi_t$ the perturbation of the instantaneous variance at time t at order 1 in ω:

$$\begin{aligned} \xi_t^t &= \xi_0^t + \delta\xi_t \\ \delta\xi_t &= \omega \xi_0^t \int_0^t \sum_{i=1}^{n} \left(\lambda_{i\tau}^t\right)_0 dW_\tau^i \end{aligned}$$

For $\omega = 0$, $M_3 = 0$. At lowest order, M_3 is thus of order 1 in ω. We then need to compute M_3 at order 1 and M_2 at order 0 in ω:

$$\begin{aligned} x_T - \langle x_T \rangle &= \int_0^T \sqrt{\xi_0^t + \delta\xi_t} dZ_t - \frac{1}{2}\int_0^T \left(\xi_0^t + \delta\xi_t\right) dt + \frac{1}{2}\int_0^T \xi_0^t dt \\ &= \int_0^T \sqrt{\xi_0^t} dZ_t + \frac{1}{2}\int_0^T \frac{\delta\xi_t}{\sqrt{\xi_0^t}} dZ_t - \frac{1}{2}\int_0^T \delta\xi_t dt \\ M_2 &= \int_0^T \xi_0^t dt \\ M_3 &= \frac{3}{2} E\left[\left(\int_0^T \sqrt{\xi_0^t} dZ_t \right)^2 \left(-\int_0^T \delta\xi_t dt + \int_0^T \frac{\delta\xi_t}{\sqrt{\xi_0^t}} dZ_t \right) \right] \end{aligned}$$

Evaluating the expectation for M_3, we get:

$$M_3 = 3\omega \int_0^T dt\, \xi_0^t \int_0^t \sqrt{\xi_0^\tau} \sum_{i=1}^{n} \rho_{iS} \left(\lambda_{i\tau}^t\right)_0 d\tau$$

where ρ_{iS} is the correlation between Z and W^i. This expression can be rewritten as:

$$M_3 = 3\int_0^T dt \left(\int_0^t E\left[\frac{dS_\tau^0}{S_\tau^0} \delta\xi_t \right] \right) \tag{2.2}$$

Equation 2.2 shows that M_3 is given at first order in the volatility of volatility by the double integral of the spot/volatility covariance function. The expression $E\left[\frac{dS_\tau^0}{S_\tau^0}\delta\xi_t\right]$ quantifies how much

[1] This approach allows for an economical derivation of the ATMF skew at order 1 in ω. We would get the same result by perturbing the pricing equation at order 1 in ω.

a move of the (unperturbed) spot at time τ is correlated with the fluctuation of the instantaneous variance at a later time t. Let us define the spot/volatility covariance function f as :

$$f(\tau,t) = \frac{1}{d\tau} E\left[\frac{dS_\tau^0}{S_\tau^0}\delta\xi_t\right] \tag{2.3}$$

At first order in the volatility of volatility, the ATMF skew is then given by:

$$\mathcal{S}_T = \frac{1}{2\sqrt{T}} \frac{\int_0^T dt \int_0^t f(\tau,t)\, d\tau}{\left(\int_0^T \xi_0^t dt\right)^{\frac{3}{2}}} \tag{2.4}$$

For an illustration of the accuracy of formula 2.4, see figure 3.1 for the case of a two-factor lognormal model for forward variances.

2.2. The skew stickiness ratio

Different models generate different deltas for vanilla options as they imply different scenarios for implied volatilities, conditional on a move of the spot. Market-makers on index options empirically adjust their deltas by making an assumption for the following ratio:

$$r_T = \frac{\Delta\widehat{\sigma}_F^T}{\left.\frac{\widehat{d\sigma_{KT}}}{d\ln K}\right|_F \Delta\ln S}$$

which quantifies how much the ATMF volatility $\widehat{\sigma}_F^T$ moves conditional on a move of S. They have coined names for two types of market regimes: sticky-strike ($r=1$) and sticky-delta ($r=0$). While they may be correlated with S, volatilities are not functions of S. Let us introduce the Skew Stickiness Ratio $\mathcal{R}_T$ which we define as:

$$\mathcal{R}_T = \frac{E\left[d\widehat{\sigma}_F^T\, d\ln(S)\right]}{\left.\frac{\widehat{d\sigma_{KT}}}{d\ln K}\right|_F E\left[(d\ln S)^2\right]}$$

$\mathcal{R}_T$ is then the regression coefficient of $\Delta\widehat{\sigma}_F^T$ on $\Delta\ln S$ in units of the ATMF skew. The values of $\mathcal{R}$ for some classes of models are well known:

- in models built with Jump or Lévy processes, $\mathcal{R}=0$
- in local volatility models, for weak skews, $\mathcal{R}=2$
- in stochastic volatility models, for short maturities and weak skews, $\mathcal{R}=2$.

As we are working at order 1 in ω and the numerator of $\mathcal{R}_T$ is of order 1 in ω, using either the VS or the ATMF volatility is indifferent, as their difference is of order 1 in ω. For the purpose of computing the numerator we then use the VS volatility whose variation at order 1 in ω at lowest order in dt is given by:

$$d\widehat{\sigma}_{tT} = \frac{1}{2\widehat{\sigma}_{tT}(T-t)} \int_t^T d\xi_t^u du$$

where $d\xi_t^u$ is given by eq. 2.1. Taking now the expectation $E\left[d\widehat{\sigma}_{tT}\, d\ln S\right]$ and keeping only terms at order 1 in ω we get:

$$E\left[d\widehat{\sigma}_{tT}\, d\ln S_t\right] = \frac{1}{2\widehat{\sigma}_{tT}(T-t)} \int_t^T E\left[d\xi_t^u\, d\ln S_t\right] du = \frac{1}{2\widehat{\sigma}_{tT}(T-t)} \int_t^T E\left[\frac{dS_t^0}{S_t^0}\delta\xi_u\right] du$$

We now divide by $\left\langle (d\ln S)^2\right\rangle$ and evaluate expectations at $t=0$, making use of the definition of f in eq. 2.3 and the expression of the ATMF skew 2.4 to get:

$$\mathcal{R}_T = \frac{\int_0^T \xi_0^t dt}{\xi_0^0 T} \frac{T\int_0^T f(0,u)\, du}{\int_0^T dt \int_0^t f(\tau,t)\, d\tau} \tag{2.5}$$

Consider how, except for some dependence on the term-structure of the variance curve, this expression for $\mathcal{R}_T$, as well as expression 2.4 for $\mathcal{S}_T$, involve the same ingredient: the spot/volatility covariance function. The common dependence of $\mathcal{R}_T$ and $\mathcal{S}_T$ on f supplies the connection between a static feature of the smile – the term-structure of the ATMF skew – and a dynamic property – the SSR .

We now study the limit of $\mathcal{R}_T$ and $\mathcal{S}_T$ when $T\to 0$ then characterize further the relationship between the SSR and the ATMF skew for the case of a time-homogeneous model and a flat variance curve.

2.3. Short-maturity limit of the ATMF skew and the SSR

Let us take the limit $T \to 0$. Using expression 2.4:

$$\mathcal{S}_0 = \lim_{T\to 0} \frac{1}{2\sqrt{T}} \frac{\int_0^T dt \int_0^t f(\tau,t)\, d\tau}{\left(\int_0^T \xi_0^t dt\right)^{\frac{3}{2}}} = \frac{f(0,0)}{4\left(\xi_0^0\right)^{\frac{3}{2}}} \tag{2.6}$$

The short skew has a finite limit which directly measures the covariance function at the origin. Let us now turn to $\mathcal{R}$. The prefactor in eq. 2.5 tends to 1 and we get:

$$\mathcal{R}_0 = \lim_{T\to 0} \frac{T\int_0^T du}{\int_0^T dt \int_0^t d\tau} = 2 \tag{2.7}$$

We recover for short maturities, at first order in volatility of volatility, the same value for stochastic volatility as for local volatility models. We had pointed out this general property in (Bergomi, 2004) explaining why, for short maturities and low skews, the dynamics – and hence the deltas – in stochastic volatility and local volatility models calibrated on the same ATMF skew were identical. Our calculation at order 1 in the volatility of volatility is also of order 1 in the spot/volatility correlations ρ_{iS}: stochastic and local volatility models behave differently if the smile near the money is dominated by curvature ($\rho_{iS} = 0$) rather than skew.

3. Scaling behaviour of $\mathcal{S}_T$ and $\mathcal{R}_T$ for a time-homogeneous model and a flat term-structure of variance

Let us now assume that the term-structure of variance is flat and that the underlying model is time-homogeneous, so that the covariance function is a function of $t-\tau$ only: $f(\tau,t) \equiv f(t-\tau)$. We now get simpler expressions for $\mathcal{S}_T$ and $\mathcal{R}_T$:

$$\mathcal{S}_T = \frac{\int_0^T (T-t) f(t)\, dt}{2(\xi_0)^{\frac{3}{2}} T^2}, \quad \mathcal{R}_T = \frac{\int_0^T f(t)\, dt}{\int_0^T \left(1-\frac{t}{T}\right) f(t)\, dt}$$

3.1. Admissible range for $\mathcal{R}_T$

The expression for $\mathcal{R}_T$ can be rewritten as

$$\mathcal{R}_T = \frac{g(T)}{\frac{1}{T}\int_0^T g(t)\, dt}$$

where $g(t) = \int_0^t f(u)\, du$. Let us make the natural assumption that $f(u)$ decays monotonically towards 0 as $u \to \infty$. $\mathcal{R}_T$ is the ratio of $g(T)$ – either positive increasing concave or negative decreasing convex, depending on the sign of f – to its average value over $[0,T]$. Thus $\mathcal{R}_T \geq 1$. Using the fact that $\frac{g(t)}{g(T)} \geq \frac{t}{T}$ yields a higher bound for $\mathcal{R}_T$: $\mathcal{R}_T \leq 2$.

We then have the following model-independent range for $\mathcal{R}_T$:

$$1 \leq \mathcal{R}_T \leq 2$$

3.2. Scaling of $\mathcal{S}_T$ and $\mathcal{R}_T$

Let us investigate the scaling behaviour of $\mathcal{S}_T$ and $\mathcal{R}_T$ by assuming that for large time separations f decays algebraically with exponent γ: $f(u) \propto u^{-\gamma}$.

As we take the limit $x \to \infty$, the integral $\int_0^x f(u)\, du$ either scales like $T^{1-\gamma}$ if $\gamma < 1$ or tends to a constant if $\gamma > 1$. Working out the limiting regimes for $\mathcal{S}_T$ and $\mathcal{R}_T$ we get:

- (Type I) If $\gamma > 1$: $\mathcal{S}_T \propto \frac{1}{T}$ and $\lim_{T\to\infty} \mathcal{R}_T = 1$
- (Type II) If $\gamma < 1$: $\mathcal{S}_T \propto \frac{1}{T^\gamma}$ and $\lim_{T\to\infty} \mathcal{R}_T = 2-\gamma$

It is easy to check that exponential decay falls into the type I category. Let us comment on these results:

- If $f(u)$ decays faster than $\frac{1}{u}$ or exponentially, the ATMF skew decays like $\frac{1}{T}$ and the long-maturity limit of the SSR is 1. We had already reported this property for the specific case of the Heston model (Bergomi, 2004). The fact that $\mathcal{S}_T$ decays like $\frac{1}{T}$ can be understood by realizing that, if the spot/volatility covariance function decays too rapidly, increments of $\ln(S_t)$ become independent. This leads to the $\frac{1}{T}$ scaling for the ATMF skew, a feature shared with Jump and Lévy models. Unlike Jump and Lévy models, however, for which $\mathcal{R} = 0$, whe have $\lim_{T\to\infty} \mathcal{R}_T = 1$.
- If $f(u)$ decays more slowly than $\frac{1}{u}$, $\mathcal{S}_T$ decays with the same exponent as f and $\mathcal{R}_T$ tends to the non-trivial limit $2-\gamma$.

The connection between the decay of the ATMF skew and the long-maturity limit of the SSR can be summarized compactly by the following formula : if the spot/volatility covariance function has either algebraic or exponential decay, then for long maturities:

$$\mathcal{S}_T \propto \frac{1}{T^{2-\mathcal{R}_*}} \tag{3.1}$$

with $\mathcal{R}_* = \lim_{T\to\infty} \mathcal{R}_T$.

3.3. Type-II behaviour in a two-factor model

Of the two types of scaling listed above for $\mathcal{S}_T$ and $\mathcal{R}_T$ type II is the most interesting. Can we build a model that generates such behaviour? Consider a model of the following type (Bergomi, 2008):

$$d\xi_t^T = \xi_t^T \omega \sum_i w_i e^{-k_i(T-t)} dW_t^i$$

where w_i are positive weights. Assuming a flat initial VS term-structure, f has the following form:

$$f(\tau) = \omega(\xi_0)^{\frac{3}{2}} \sum_n w_i \rho_{Si} e^{-k_i\tau}$$

Plugging this expression in equations 2.4 and 2.5 yields :

$$\begin{aligned} \mathcal{S}_T &= \frac{\omega}{2} \sum_i w_i \rho_{Si} \frac{k_i T - \left(1-e^{-k_i T}\right)}{(k_i T)^2} \\ \mathcal{R}_T &= \frac{\sum_i w_i \rho_{Si} \frac{1-e^{-k_i T}}{k_i T}}{\sum_i w_i \rho_{Si} \frac{k_i T - \left(1-e^{-k_i T}\right)}{(k_i T)^2}} \end{aligned} \tag{3.2}$$

f is a linear combination of exponentials: as $T\to\infty$, $f(u) \propto e^{-\min_i(k_i)u}$. Thus when $T\to\infty$, $\mathcal{S}_T \propto \frac{1}{T}$ and $\mathcal{R}_T \to 1$: the model eventually behaves like type I. However, by suitably choosing parameters, it is possible to generate a power law-like behaviour for f over a sufficiently wide range of maturities.

Let us take two factors and use the following parameters – these are values we used in (Bergomi, 2008): $k_1 = 8.0, k_2 = 0.35, w_1 = 72\%, w_2 = 28\%, \rho_{S1} = -70\%, \rho_{S2} = -35.7, \omega = 3.36$, which are typical of equity index skews and vols-of-vols. Figure 3.1 shows a comparison of the 95/105 skew in volatility points computed either by direct Monte-Carlo simulation of the model, or using equation 3.2.

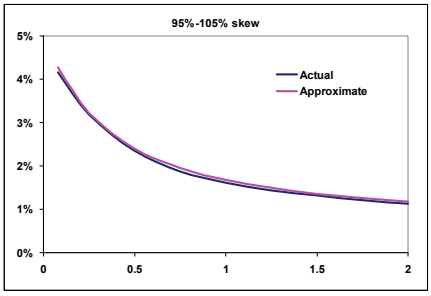

Figure 3.1: Approximate and actual 95/105 skew in a two-factor model, as a function of T.

Figure 3.2 shows on the left the 95/105 skew defined as $\mathcal{S}_T \,.\ln(95/105)$ as a function of T in log/log plot, for maturities from 3 months to 5 years. It is almost a straight line with slope about $\frac{1}{2}$ (-0.51). This exponent is a well-known typical feature of equity smiles. The graph on the right shows $\mathcal{R}_T$ as a function of T.

As expected, $\mathcal{R}_T$ starts from 2 and tends for long maturities to 1. Notice though the shoulder around 1.5 for intermediate maturities. This can be traced to the scaling of $\mathcal{S}_T$: initially $\mathcal{S}_T$ decays approximately algebraically with power $\frac{1}{2}$, consequently, according to equation 3.1 $\mathcal{R}_T$ initially stabilizes to a value equal to $2 - \frac{1}{2} = 1.5$. Eventually, for longer maturities, the exponential decay of f kicks in ($f \sim e^{-k_2 T}$), so that $\mathcal{S}_T$ decays like $\frac{1}{T}$ and $\mathcal{R}_T$ tends to its long-maturity limit of 1.

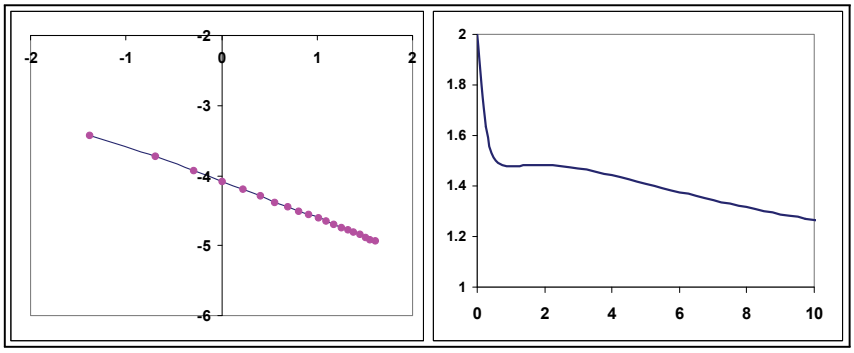

Figure 3.2: Left: $\ln(\mathcal{S}_T \,.\ln(95/105))$ as a function of $\ln(T)$. Right: $\mathcal{R}(T)$

Even though the model becomes of type I when $T \to \infty$, we are able to get type II behaviour over a range of maturities that is wide enough for practical purposes.

3.4. Type-II behaviour with the Eurostoxx 50 index

As mentioned above, the ATMF skew of equity smiles typically decays like $\frac{1}{T^{1/2}}$. This would suggest type-II behaviour, but is this confirmed by the value of $\mathcal{R}_T$? Let us here look at the realized SSR of the Eurostoxx50 index for different maturities, measured using the ATM volatility. Figure 3.3 displays the three-month running average of $\mathcal{R}_T$ for maturities 1 month, 6 months, 2 years, since May 2002, computed as:

$$\mathcal{R}_T = \frac{\sum_i (\widehat{\sigma}_{i+1} - \widehat{\sigma}_i) \ln\left(\frac{S_{i+1}}{S_i}\right)}{\sum_i \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_S^i \ln\left(\frac{S_{i+1}}{S_i}\right)^2}$$

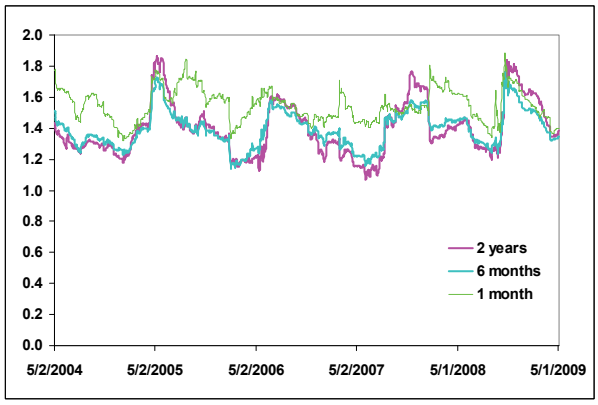

Figure 3.3: 3-month running average of $\mathcal{R}_T$ for maturities: 1 month, 6 months, 2 years, for the Eurostoxx50 index

We observe that:

- The SSR usually lies in the interval $[1, 2]$
- The SSR is very noisy and is impacted by sudden and simultaneous changes in spot and ATM volatility. However, over the last 5 years, the average value of the SSR for long-dated options has been notably larger than one and hovers around 1.4. This number is – in our framework – compatible with a power law decay of the skew with an exponent around $\frac{1}{2}$.

Equity volatility markets thus seem to behave like type II. Note, however, that the short-maturity limit of the SSR (equation 2.7) is 2 whether the model be of type I or type II. This value is very different than the historical value of the SSR for one-month options on the Eurostoxx50. As figure 3.3 shows, its value is always lower than 2, at times markedly.

It is then natural to ask whether this can be arbitraged. It could be that the joint historical dynamics of spot and implied volatilities is not compatible with that generated by a stochastic volatility model – this would be the case, for example, if the "real" dynamics was generated by a Jump model, in which case the difference between 2 and $\mathcal{R}_0$ would stem from the misspecification of our hedging model and would not give rise to arbitrage opportunities. This issue is best tackled by answering the following question : is it possible to implement an option strategy whose P&L is $2 - \mathcal{R}_0$?

4. Arbitraging the SSR for short maturities

We consider short maturities for which the distinction between ATMF and ATM strikes is not relevant: let us use the ATM volatility, which we denote by σ_0. We also neglect interest rate and repo effects.The SSR relates the level of the market ATMF skew to the spot/volatility covariance. Arbitraging the SSR entails being able to materialize this covariance as a P&L, hence requires a hedging model in which at least both S and σ_0 can move.

We now develop a model for the joint dynamics of S and the smile, using as dynamical variables S and σ_0. Our goal is to be able to write the profit and loss (P&L) of a delta-hedged, σ_0-hedged vanilla option as:

$$P\&L = \frac{1}{2}S^2\frac{d^2Q}{dS^2}\left(\left(\frac{\delta S}{S}\right)^2 - \sigma_S^2\delta t\right) + \frac{1}{2}\sigma_0^2\frac{d^2Q}{d\sigma_0^2}\left(\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2 - \nu^2\delta t\right) + S\sigma_0\frac{d^2Q}{dSd\sigma_0}\left(\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0} - \rho\sigma_S\nu\delta t\right) \tag{4.1}$$

with the crucial condition that the break-even levels σ_S, ν, ρ be strike-independent – unlike the Black-Scholes implied volatility $\widehat{\sigma}_K$ – and such that the market smile is recovered.

4.1. A model for short near-the-money options

Let us consider short-maturity vanilla options. Let us introduce moneyness $x = \ln\left(\frac{K}{S}\right)$ and parameterize the smile near the money as:

$$\widehat{\sigma}(x) = \sigma_0\left(1 + \alpha(\sigma_0)\,x + \frac{\beta(\sigma_0)}{2}x^2\right) \tag{4.2}$$

The smile is characterized by three quantities: σ_0, the skew $\sigma_0\alpha(\sigma_0)$ and the curvature $\sigma_0\beta(\sigma_0)$. α and β are functions of σ_0 and $\widehat{\sigma}(x)$ has no explicit dependence on $T - t$.

The price of an option is then: $Q(S, K, \sigma_0, \alpha, \beta, T) = P_{BS}(S, K, \sigma(x), T)$ where $P_{BS}(S, \widehat{\sigma}, T)$ is the Black-Scholes formula. The P&L of a delta-hedged, σ_0-hedged vanilla option reads, at order δt:

$$P\&L = \frac{dQ}{dt}\delta t + \frac{1}{2}S^2\frac{d^2Q}{dS^2}\left(\frac{\delta S}{S}\right)^2 + \frac{1}{2}\sigma_0^2\frac{d^2Q}{d\sigma_0^2}\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2 + S\sigma_0\frac{d^2Q}{dSd\sigma_0}\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0} \tag{4.3}$$

Our parametrization for the smile 4.2, together with the assumption that S, σ_0 are the only dynamical quantities in our model, is consistent only if we are able to find break-even levels σ_S, ν, ρ that make this P&L vanish on average, irrespective of the strike considered. Since $\widehat{\sigma}$ has no explicit time-dependence, the theta in our model is the same as the Black-Scholes theta: $\frac{dQ}{dt} = \frac{dP}{dt}$. Our consistency requirement can be stated as:

$$-\frac{dQ}{dt} = \frac{1}{2}S^2\frac{d^2P_{BS}^K}{dS^2}\widehat{\sigma}_K^2 = \frac{1}{2}S^2\frac{d^2Q^K}{dS^2}\sigma_S^2 + \frac{1}{2}\sigma_0^2\frac{d^2Q^K}{d\sigma_0^2}\nu^2 + S\sigma_0\frac{d^2Q^K}{dSd\sigma_0}\rho\sigma_S\nu \tag{4.4}$$

In other words, we need to be able to split the Black-Scholes theta into 3 pieces – matching our 3 gammas.

Inspection of the derivatives $\frac{d^2P_{BS}}{dS^2}$, $\frac{d^2P_{BS}}{dSd\sigma}$, $\frac{d^2P_{BS}}{d\sigma^2}$ shows that they all have the same prefactor $\frac{SN'(d)}{\sigma\sqrt{T}}$, where $d = \frac{-x+\frac{\sigma^2T}{2}}{\sigma\sqrt{T}}$, which encapsulates the singularity at $T = 0$. Factoring this prefactor

out and keeping terms at order 2 in x and order 0 in T yields the following expressions for the greeks in equation 4.3:

$$\begin{aligned}
\frac{dQ}{dt} &= -\frac{1}{2}\frac{SN'(d)}{\sqrt{T}}\sigma_0\left(1+\alpha x+\frac{\beta^2}{2}x^2\right)\\
\frac{1}{2}S^2\frac{d^2Q}{dS^2} &= \frac{1}{2}\frac{SN'(d)}{\sigma_0\sqrt{T}}\left(1-3\alpha x+\left(6\alpha^2-\frac{5}{2}\beta\right)x^2\right)\\
\frac{1}{2}\sigma_0^2\frac{d^2Q}{d\sigma_0^2} &= \frac{1}{2}\frac{SN'(d)}{\sigma_0\sqrt{T}}x^2\\
S\sigma_0\frac{d^2Q}{dSd\sigma_0} &= \frac{SN'(d)}{\sigma_0\sqrt{T}}\left(x-\left(2\alpha-\sigma_0\alpha'\right)x^2\right)
\end{aligned}$$

where $\alpha' = \frac{d\alpha}{d\sigma_0}$. Plugging these expressions in equation 4.1 yields:

$$P\&L = \frac{1}{2}S\frac{N'(d)}{\sigma_0\sqrt{T}}\left[\begin{array}{c}-\sigma_0^2\delta t\left(1+\alpha x+\frac{\beta}{2}x^2\right)+\left(1-3\alpha x+\left(6\alpha^2-\frac{5}{2}\beta\right)x^2\right)\left(\frac{\delta S}{S}\right)^2\\ +x^2\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2+2\left(x-\left(2\alpha-\sigma_0\alpha'\right)x^2\right)\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0}\end{array}\right]$$

Let us now find breakeven levels: $\left\langle\left(\frac{\delta S}{S}\right)^2\right\rangle = \sigma_S^2\delta t$, $\left\langle\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2\right\rangle = \nu^2\delta t$, $\left\langle\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0}\right\rangle = \rho\sigma_0\nu\delta t$ that make the P&L vanish for all x. Grouping terms by powers of x we get:

$$P\&L = \frac{1}{2}S\frac{N'(d)}{\sigma_0\sqrt{T}}\left[\begin{array}{c}\left(-\sigma_0^2+\sigma_S^2\right)+\left(-\sigma_0^2\alpha-3\alpha\sigma_S^2+2\rho\sigma_S\nu\right)x\\ +\left(-\sigma_0^2\frac{\beta}{2}+\left(6\alpha^2-\frac{5}{2}\beta\right)\sigma_S^2+\nu^2-2\left(2\alpha-\sigma_0\alpha'\right)\rho\sigma_S\nu\right)x^2\end{array}\right]\delta t$$

This yields the following equations for σ_S, ν, ρ:

$$\sigma_S = \sigma_0 \tag{4.5}$$
$$\rho\nu = 2\alpha\sigma_0 \tag{4.6}$$
$$\nu^2 = \sigma_0^2\left(3\beta+2\alpha^2-4\sigma_0\alpha\alpha'\right) \tag{4.7}$$

- The first equation expresses that the break-even volatility for the spot is the ATM volatility
- The 2nd equation relates the ATM skew to the covariance of S and σ_0: using the fact that, in our parameterization $\frac{d\sigma}{d\ln K}\Big|_S = \alpha\sigma_0$, it can be rewritten as:

$$\frac{\frac{1}{\delta t}\left\langle\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0}\right\rangle}{\frac{d\sigma}{d\ln K}\Big|_S\sigma_0} = 2$$

  which recovers the result that $\mathcal{R}_0 = 2$. A discrepancy between the realized value of $\mathcal{R}$ and 2 is then materialized as a spot/volatility cross-gamma/theta P&L: using condition 4.6 the third piece in equation 4.1 now reads:

$$S\sigma_0\frac{d^2Q}{dSd\sigma_0}\left(\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0}-2\sigma_0\left.\frac{d\sigma}{d\ln K}\right|_S\delta t\right) \tag{4.8}$$

- Most importantly, these two properties are model-independent: they do not depend on the functions $\alpha(\sigma_0)$ and $\beta(\sigma_0)$. The third equation, which relates the curvature parameter β to the volatility of volatility ν and α, is model-dependent as it involves $\alpha'(\sigma_0)$.

We are free to specify any functional from for $\alpha(\sigma_0)$ and $\beta(\sigma_0)$ provided conditions 4.5, 4.6, 4.7 hold.

4.1.1. Consistency conditions

S, σ_0 are allowed to move while functions α, β stay constant. If ρ is assumed to be constant, then equations 4.6, 4.7 show that the dependence of the ATM skew to σ_0 is related to the dependence of ν on σ_0.

Lognormal dynamics for σ_0: Let us assume that ν is a constant: eq. 4.6 implies that $\sigma_0\alpha(\sigma_0)$ is constant, thus α is proportional to $\frac{1}{\sigma_0}$. Equation 4.7 then implies that β is proportional to $\frac{1}{\sigma_0^2}$. Let us write $\alpha = \frac{a}{\sigma_0}$, $\beta = \frac{b}{\sigma_0^2}$. We get the following expression for the smile:

$$\sigma(x) = \sigma_0\left(1 + \frac{a}{\sigma_0}x + \frac{b}{2\sigma_0^2}x^2\right) \tag{4.9}$$

where a, b are constant, and ν and ρ are given by:

$$\rho\nu = 2a \tag{4.10}$$
$$\nu = \sqrt{3b + 6a^2} \tag{4.11}$$

Thus assuming a lognormal dynamics for the ATM volatility implies on one hand that the skew is constant, on the other hand that the curvature is inversely proportional to the ATM volatility.

Normal dynamics for σ_0: Assume now that ν is inversely proportional to σ_0: $\nu = \frac{\mu}{\sigma_0}$ where μ is the constant normal volatility of volatility. Equations 4.6 and 4.7 then imply that α is inversely proportional to σ_0^2 and β is inversely proportional to σ_0^4: $\alpha = \frac{a}{\sigma_0^2}$, $\beta = \frac{b}{\sigma_0^4}$where a, b are constant. We get the following expression for the smile:

$$\sigma(x) = \sigma_0\left(1 + \frac{a}{\sigma_0^2}x + \frac{b}{2\sigma_0^4}x^2\right)$$

where a, b are related to μ and ρ by:

$$\rho\mu = 2a$$
$$\mu = \sqrt{3b + 10a^2}$$

Assuming a normal dynamics for the ATM volatility then generates a skew which is inversely proportional to the ATM volatility and a curvature inversely proportional to the cube of the ATM vol.

These scaling properties for skew and curvature as a function of the ATM volatility induced by the dependence of the volatility of the ATM volatility on the ATM volatility itself, are in agreement with more general results derived for short maturities by directly specifying a dynamics for the implied volatilities and imposing that the discounted option price be a martingale – see for example (Balland, 2006) and (Durrleman, 2004).

One can check that the lognormal (resp. normal) case is the short-maturity limit of the SABR (resp. Heston) model.

4.2. Conclusion and application to one-month-maturity options - lognormal dynamics for σ_0

In what follows we use a lognormal dynamics for σ_0. The parameterization of the smile near the money is given by equation 4.9 and the P&L during δt of a delta-hedged, σ_0-hedged vanilla option reads:

$$P\&L = \frac{1}{2}S^2\frac{d^2Q}{dS^2}\left(\left(\frac{\delta S}{S}\right)^2 - \sigma_0^2\delta t\right) + \frac{1}{2}\sigma_0^2\frac{d^2Q}{d\sigma_0^2}\left(\left(\frac{\delta\sigma_0}{\sigma_0}\right)^2 - \left(3b + 6a^2\right)\delta t\right) + S\sigma_0\frac{d^2Q}{dSd\sigma_0}\left(\frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0} - 2a\sigma_0\delta t\right) \tag{4.12}$$

This was derived in the limit of a short maturity and for strikes near the money. How reliable are our approximations, for practical trading purposes? Let us check how accurately equation 4.4 holds: how well do our three thetas add up to the the Black-Scholes theta and what are their relative magnitudes?

We consider a one-month maturity and a typical index short-maturity smile, shown in figure 4.1. The corresponding values of σ_0,a, b are: $\sigma_0 = 20\%$, $a = -10\%$, $b = 0.4\%$.

Figure 4.2 shows, on the left: the spot theta: $\frac{\sigma_0^2}{2}S^2\frac{d^2Q}{dS^2}$, the volatility theta: $\frac{3b+6a^2}{2}\sigma_0^2\frac{d^2Q}{d\sigma_0^2}$, and the cross spot/volatility theta: $2a\sigma_0 S\frac{d^2Q}{dSd\sigma_0}$. The right-hand graph shows both the Black-Scholes theta: $\frac{\sigma_K^2}{2}S^2\frac{d^2P_{BS}}{dS^2}$ and the sum of the three thetas in the left-hand graph . The x-axis in both graphs is the option's strike. The right-hand graph shows good agreement: the model is usable in practice.

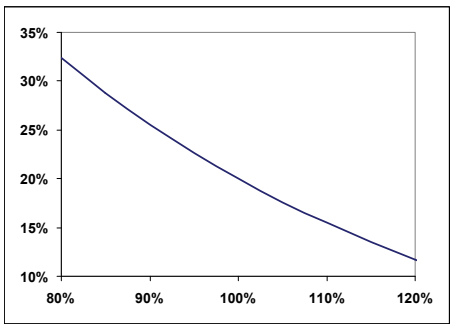

Figure 4.1: Implied volatilities for $\sigma_0 = 20\%$, $a = -10\%$, $b = 0.4\%$

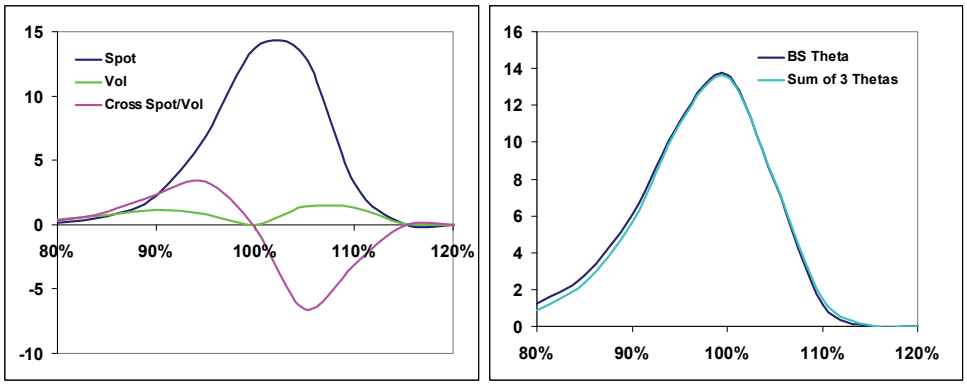

Figure 4.2: Left: the spot, volatility and cross - theta. Right: the sum of the 3 thetas compared to the Black-Scholes theta.

4.3. Arbitraging the 95/105 1-month skew on the Eurostoxx50

The discrepancy between the realized value of $\mathcal{R}_0$ and its model-independent value of 2 pictured in figure 3.3 can equivalently be expressed as a discrepancy between the market ATM skew and the "realized" skew whose definition is suggested by expression 4.8:

$$\left.\frac{d\sigma}{d\ln K}\right|_S^{\text{Realized}} = \frac{1}{2\delta t}\left\langle \frac{\delta S}{S}\frac{\delta\sigma_0}{\sigma_0^2}\right\rangle \tag{4.13}$$

which materializes into a non-vanishing cross gamma/theta P&L. This is shown on figure 4.3, where the realized and implied skew have been multiplied by 10 to correspond approximately to a 95/105 skew.

Note that the "fair" estimator of the realized skew in equation 4.13 involves the covariance of the spot and the implied ATM volatility, rather than the covariance of the spot and its realized *volatility*.

Equation 4.12 shows that to isolate as a P&L the difference between realized and implied skew we need to: a) cancel the spot and volatility gamma, b) ensure that it is not dwarfed by the P&L generated by remarking to market a and b.

We now backtest a strategy that involves selling every day 1 one-month option of strike 95 and buying the appropriate number – typically around 0.7 – of options of strike 105 so as to cancel the spot gamma. This position is delta-hedged and unwound the next day, then started again. On top of the third piece in equation 4.12, our total P&L comprises (a) a volatility gamma/theta P&L – the second piece in expression 4.12, (b) a vega P&L as our position has some small residual sensitivity to σ_0, (c) additional P&L created by remarking a, b to market on the next day. Since $\frac{dP}{d\sigma_0}$ and $\frac{d^2P}{d\sigma_0^2}$ are approximately symmetric around the money, P&Ls (a) and (b) are expected to be small.

The results of our backtest are illustrated in figure 4.4 – for information the vega of the 95-strike option is around 0.1. The left-hand scatter plot shows the daily P&L without portions (b) and (c) as a function of the P&L computed using equation 4.12 while the scatter plot on the right shows the real total daily P&L as a function of the cross-gamma/theta P&L 4.8.

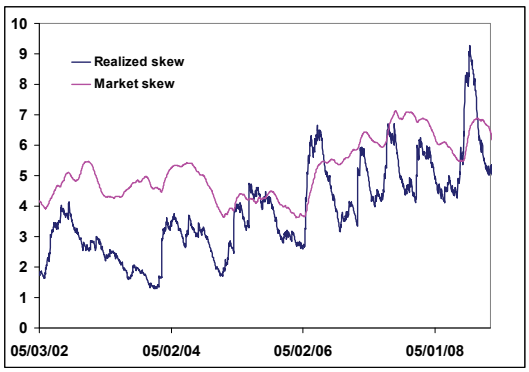

Figure 4.3: 3-month running averages of realized and market skew (difference in volatility points between the 95 and the 105 strikes) for 1-month options on the Eurostoxx50

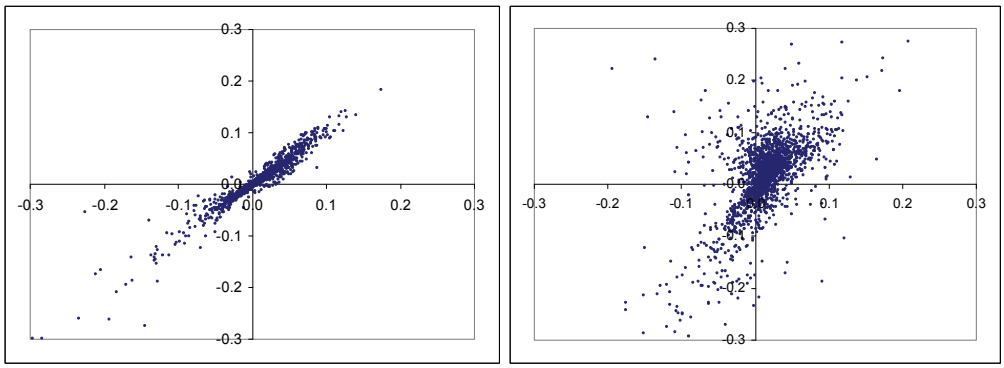

Figure 4.4: P&L of a dynamic option strategy - see text

The dispersion of points around a straight line in the left-hand graph is another measure of how well the P&L accounting in equation 4.12 holds: note that it is only correct at order 1 in δt and at order 2 in δS and $\delta\sigma_0$ and is has been derived with the assumption of short maturities and strikes near the money – still the agreement is satisfactory. More interestingly, the right-hand graph shows that, although some noise is contributed by P&Ls (b) and (c), the total P&L of our strategy is still well correlated with the cross-gamma/theta P&L, that is: the difference between realized and implied skew.

This is also manifested in figure 4.5 which shows, for the proposed strategy, on one hand the total cumulative P&L of the strategy, on the other hand the cumulative cross-gamma/theta P&L: on average the two P&Ls track each other rather well.

Our conclusion is that the difference between the realized value of $\mathcal{R}_0$ and 2 can be materialized as a P&L and hence arbitraged. Running the strategy we have outlined entails unreasonable bid/offer costs; one would rather use an automaton to optimally rebalance an option position near the money so as to maintain vanishing spot gamma, small vega and volatility gamma.

5. Conclusion

In time-homogeneous stochastic volatility models at order 1 in the volatility of volatility, the Skew Stickiness Ratio and the rate at which the ATMF skew decays with maturity are structurally related through the spot/volatility covariance function. Assuming time-homogeneity and a flat VS volatility term-structure implies that the SSR is restricted to the interval [1, 2]. Inspection of the historical behaviour of Eurostoxx50 implied volatilities shows that while for longer maturities the relationship between the SSR and the ATMF skew holds approximately, for short maturities the SSR is at times significantly smaller than its model-independent value of 2, indicating that the market skew is too steep. We show that this discrepancy is equivalent to the difference between the market implied at-the-money skew and its "realized" counterpart, which we define, and that it can be materialized

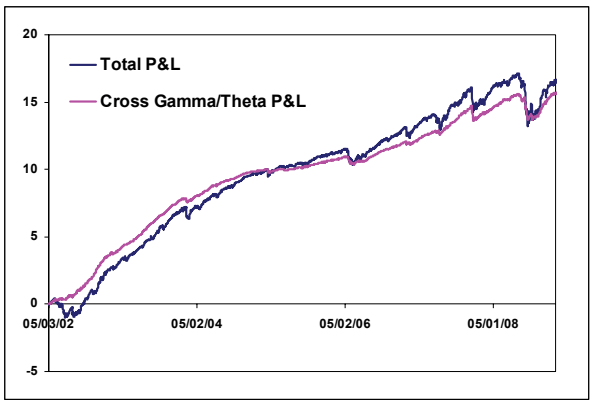

Figure 4.5: P&Ls of arbitrage strategy

as the cross-gamma/theta P&L of a hedged near-the-money vanilla option position. We provide an example of a dynamic option strategy whose P&L approximately captures this effect. The estimator for the "realized" skew involves the realized covariance of the spot and the ATM volatility.

References

Balland P., 2006
Forward smile
Presentation at Global Derivatives, Paris, 2006

Bergomi L., 2004
Smile dynamics
Risk September, pages 117-123

Bergomi L., 2005
Smile dynamics II
Risk October, pages 67-73

Bergomi L., 2008
Smile dynamics III
Risk October, pages 90-96

Durrleman V., 2004
From implied to spot volatilities
http://math.stanford.edu/~valdo/papers/dissertation.pdf