Chapter 11

Multi-asset stochastic volatility

Multi-asset stochastic volatility models are not mere juxtapositions of single-asset models. Obviously we need to specify an asset/asset correlation matrix, just as in a multi-asset Black-Scholes model.

In addition to the individual spot/volatility and volatility/volatility covariance functions whose role we have examined in Chapter 8, multi-asset stochastic volatility models involve cross spot/volatility and volatility/volatility correlations. What is the effect of these new parameters?

We propose a parametrization of the global correlation matrix based on observable physical quantities, study the ATMF skew of a basket and the correlation swap.

In the course of our discussion we use the example of the multi-asset local volatility model, the most popular multi-asset stochastic volatility model.

11.1 The short ATMF basket skew

This section focuses on the short-maturity ATMF basket skew. Consider a basket B consisting of n assets, with weights α_i:

$$B = \sum_i \alpha_i S_i$$

and let us define relative weights $w_i = \frac{\alpha_i S_i}{B}$: $\Sigma_i w_i = 1$. The variation of B is:

$$\frac{dB}{B} = \sum_i w_i \frac{dS_i}{S_i} \tag{11.1}$$

While the α_i are fixed, the w_i are not; as we will be studying the short-maturity case we will make the approximation that the w_i are constant. From the results of Section 8.5, the short ATMF implied volatility $\widehat{\sigma}_B$ and skew $\mathcal{S}_B$ are given by – see formula (8.36), page 322:

$$\widehat{\sigma}_B = \sigma_B \tag{11.2a}$$

$$\mathcal{S}_B = \frac{1}{2\widehat{\sigma}_B^2} \frac{\langle d\ln B \, d\widehat{\sigma}_B \rangle}{dt} \tag{11.2b}$$

where σ_B is the instantaneous volatility of the basket which, from (11.1) is given by:

$$\sigma_B^2 = \sum_{ij} w_i w_j \rho_{ij} \sigma_i \sigma_j$$

σ_i is the instantaneous volatility of asset i – equal to the short ATMF volatility $\widehat{\sigma}_i$ – and ρ_{ij} is the correlation of S_i and S_j. The relationship between the short ATMF basket volatility and the short ATMF volatilities of the constituents reads:

$$\widehat{\sigma}_B = \sqrt{\sum_{ij} w_i w_j \rho_{ij} \sigma_i \sigma_j} = \sqrt{\sum_{ij} w_i w_j \rho_{ij} \widehat{\sigma}_i \widehat{\sigma}_j} \tag{11.3}$$

$d\widehat{\sigma}_B$ is given by:

$$d\widehat{\sigma}_B = \frac{1}{\widehat{\sigma}_B} \sum_{ij} w_i w_j \rho_{ij} \widehat{\sigma}_i d\widehat{\sigma}_j$$

and we get:

$$\begin{aligned}
\frac{1}{2\widehat{\sigma}_B^2} \frac{\langle d\ln B \, d\widehat{\sigma}_B \rangle}{dt} &= \frac{1}{2\widehat{\sigma}_B^3} \sum_{ijk} w_i w_j w_k \rho_{ij} \widehat{\sigma}_i \frac{\langle d\ln S_k \, d\widehat{\sigma}_j \rangle}{dt} \\
&= \frac{1}{2\widehat{\sigma}_B^3} \sum_{ij} w_i w_j^2 \rho_{ij} \widehat{\sigma}_i \frac{\langle d\ln S_j \, d\widehat{\sigma}_j \rangle}{dt} + \frac{1}{2\widehat{\sigma}_B^3} \sum_{i\; j\neq k} w_i w_j w_k \rho_{ij} \widehat{\sigma}_i \frac{\langle d\ln S_k \, d\widehat{\sigma}_j \rangle}{dt}
\end{aligned}$$

The basket skew is thus given by:

$$\mathcal{S}_B = \frac{1}{\widehat{\sigma}_B^3} \sum_{ij} w_i w_j^2 \rho_{ij} \widehat{\sigma}_i \widehat{\sigma}_j^2 \, \mathcal{S}_j + \frac{1}{2\widehat{\sigma}_B^3} \sum_{i\; j\neq k} w_i w_j w_k \rho_{ij} \widehat{\sigma}_i \frac{\langle d\ln S_k \, d\widehat{\sigma}_j \rangle}{dt} \tag{11.4}$$

where we have separated terms involving the covariance on an asset with its own volatility and have used (11.2b) to relate $\langle d\ln S_j \, d\widehat{\sigma}_j \rangle$ to $\mathcal{S}_j$, the ATMF skew of basket component S_j. Expression (11.4) shows that the basket ATMF skew is generated partly by the ATMF skew of the basket components and partly by the covariance of each component with other components' ATMF volatilities.

11.1.1 The case of a large homogeneous basket

Assume that all volatilities and correlations are equal and that the n components of the basket are equally weighted: $\widehat{\sigma}_i \equiv \widehat{\sigma}$, $\rho_{ij} \equiv \rho_{SS}$, $\mathcal{S}_i = \mathcal{S}$, $w_i = \frac{1}{n}$. Expressions (11.3) and (11.4) for the basket ATMF volatility and skew simplify to:

$$\widehat{\sigma}_B = \sqrt{\frac{1 + (n-1)\rho_{SS}}{n}} \, \widehat{\sigma} \tag{11.5a}$$

$$\mathcal{S}_B = \frac{1 + (n-1)\rho_{SS}}{n} \frac{\widehat{\sigma}^3}{\widehat{\sigma}_B^3} \left[\frac{\mathcal{S}}{n} + \frac{n-1}{n} \frac{1}{2\widehat{\sigma}^2} \frac{\langle d\ln S \, d\widehat{\sigma} \rangle_{\text{cross}}}{dt} \right] \tag{11.5b}$$

where $\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}}$ denotes the instantaneous covariation of a basket component with another component's ATMF volatility. Observe that the portion contributed to $\mathcal{S}_B$ by each component's ATMF skew scales like $\frac{1}{n}$, hence goes to zero for a large basket.

For a large basket ($n \gg 1$) – which is the case of most equity indexes – we get the following simpler formulas:

$$\widehat{\sigma}_B \simeq \sqrt{\rho_{SS}}\,\widehat{\sigma} \tag{11.6a}$$

$$\mathcal{S}_B \simeq \frac{1}{\sqrt{\rho_{SS}}}\frac{1}{2\widehat{\sigma}^2}\frac{\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}}}{dt} \tag{11.6b}$$

Equation (11.6b) makes it clear that the short ATMF skew of a large basket only depends on cross-asset spot-volatility covariances. Consider the following opposite cases:

- $\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}} = 0$: a component's volatility is uncorrelated with other components. In this case $\mathcal{S}_B = 0$.
- $\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}} = \langle d\ln S\, d\widehat{\sigma}\rangle_{\text{diag}}$ where $\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{diag}}$ is the instantaneous covariation of a basket component with its own implied volatility. This situation is obtained for example by assuming that volatilities of all assets are driven by a single Brownian motion: the $\widehat{\sigma}_i$ are 100% correlated. Using:

$$\mathcal{S} = \frac{1}{2\widehat{\sigma}^2}\frac{\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{diag}}}{dt} \tag{11.7}$$

  we get:

$$\mathcal{S}_B \simeq \frac{1}{\sqrt{\rho_{SS}}}\mathcal{S} \tag{11.8}$$

  The basket skew is larger than the component's skew.

Before turning to the real case, we examine the prediction of the local volatility model.

11.1.2 The local volatility model

A multi-asset local volatility model is a peculiar stochastic volatility model in that it only takes as inputs asset-asset correlations. Implied volatilities are *functions* of spot and time: $\widehat{\sigma} = \widehat{\sigma}(t, S)$, thus are driven by the same Brownian motion as the spot process. For a homogeneous basket, such that the local volatility functions of all components are identical, this entails that:

$$\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}} = \rho_{SS}\,\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{diag}}$$

Using again (11.7), expressions (11.5) become:

$$\widehat{\sigma}_B = \sqrt{\frac{1+(n-1)\rho_{SS}}{n}}\,\widehat{\sigma} \simeq \sqrt{\rho_{SS}}\,\widehat{\sigma} \tag{11.9a}$$

$$\mathcal{S}_B = \sqrt{\frac{1+(n-1)\rho_{SS}}{n}}\,\mathcal{S} \simeq \sqrt{\rho_{SS}}\,\mathcal{S} \tag{11.9b}$$

One consequence of (11.9) is that, in the local volatility model, for short maturities, the ratios of ATMF basket skew to ATMF component skew and of ATMF basket volatility to ATMF component volatility are identical and equal to $\sqrt{\rho_{SS}}$.

$$\frac{\mathcal{S}_B}{\mathcal{S}} \simeq \frac{\widehat{\sigma}_B}{\widehat{\sigma}} \simeq \sqrt{\rho_{SS}} \tag{11.10}$$

Figure 11.1 provides an illustration of how well relationship (11.10) is actually obeyed, for a range of maturities. We have taken $n = 10$, $\rho_{SS} = 50\%$, and have used for the 10 components the same smile.[1]

Observe how $\frac{\widehat{\sigma}_B}{\widehat{\sigma}}$ stays remarkably close to its theoretical value $\sqrt{\frac{1+(n-1)\rho_{SS}}{n}} = 74.2\%$, even for long maturities. $\frac{\mathcal{S}_B}{\mathcal{S}}$, on the other hand, is in fact lower than its short-maturity value (11.9b).

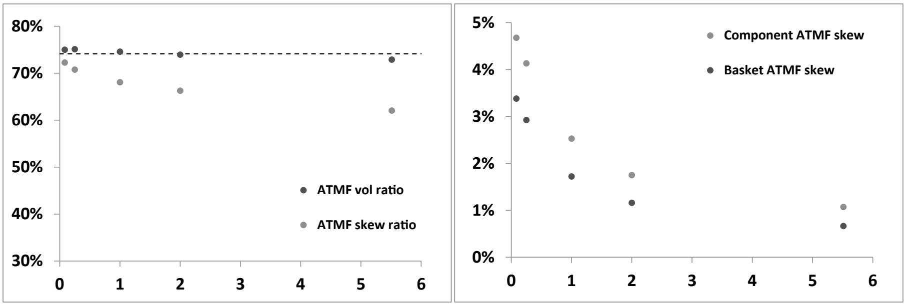

Figure 11.1: Left: ratios $\frac{\widehat{\sigma}_B}{\widehat{\sigma}}$ and $\frac{\mathcal{S}_B}{\mathcal{S}}$ as a function of maturity (years) in the local volatility model for an equally weighted basket of 10 assets with the same smile. The dashed line indicates the theoretical short-maturity value $\sqrt{\frac{1+(n-1)\rho_{SS}}{n}}$. Right: the ATMF skew expressed as the difference of the implied volatilities of the 95% and 105% strikes.

11.1.3 The basket skew in reality

Typically, implied correlations levels for equity indexes – derived from (11.6a) – are $\rho_{SS} \simeq 60\%$: ATMF index implied volatilities are about 25% lower than stock implied volatilities. Market ATMF skews of indexes are stronger than stock skews. They are typically about 25% larger – this is in fact roughly what formula (11.8) for the case of 100% correlated volatilities yields.

This stands in stark contrast with the prediction of the local volatility model: the short-maturity approximate formula (11.10) implies that the index skew should be about 20% *lower* than the stock skew; values of $\frac{\mathcal{S}_B}{\mathcal{S}}$ in Figure 11.1 – obtained with $\rho_{SS} = 50\%$ – show that the actual ratio in the local volatility model is lower still.

[1] We have used for the components the smile of the Euro Stoxx 50 index of June 12th, 2012, which is steeper than typical stock skews – using weaker skews wouldn't alter our conclusions.

11.1.4 Digression – how many stocks are there in an index?

Formulas (11.5a) and (11.5b) for $\widehat{\sigma}_B$ and $\mathcal{S}_B$ apply to the case of an equally weighted basket; the large-basket regime is reached when $n \gg 1$. Equity indexes are not equally weighted baskets. Given a particular index, which value of n should be used in (11.5a) and (11.5b) for assessing whether the asymptotic regime is applicable?

Let us assume that the volatilities of the index components are all equal to σ.[2] σ_B^2 is given by:

$$\sigma_B^2 \;=\; \sigma^2 \sum_{ij} w_i w_j \rho_{ij} \;=\; \sigma^2 \Big(\sum_i w_i^2 + \rho_{SS} \sum_{i \neq j} w_i w_j \Big) \tag{11.11}$$

Comparing this formula with (11.5a), we see that the effective number of components of an index, $n^\star$, should be defined as:[3]

$$n^\star \;=\; \frac{1}{\Sigma_i w_i^2} \tag{11.12}$$

Table 11.1 lists the value of n and $n^\star$ for a few equity indexes; $n^\star$ can be sizeably smaller than n.

	S&P 500	Euro Stoxx 50	NIKKEI	KOSPI	FTSE	SMI	CAC	Russell 2000
n	500	50	225	200	101	20	40	2000
$n^\star$	143	37	47	17	36	8	20	1056

Table 11.1: Actual (n) and effective ($n^\star$) number of components of some equity indexes. In (11.12) values of w_i as of August 6, 2013 have been used.

11.2 Parametrizing multi-asset stochastic volatility models

Given a set of single-asset models, we need to define (a) correlations for the spot processes, (b) a parametrization of cross spot/volatility and volatility/volatility correlations.

Rather than directly parametrizing cross factor/factor and spot/factor correlations, whose significance depends on the specifics of the factor structure, we aim for a parsimonious parametrization of "physical" quantities.

Such parametrization can be used regardless of the particular factor structure of the model at hand.

[2]Or less crudely that there is no systematic correlation between w_i and σ_i.

[3]The same value of $n^\star$ can be extracted from the second component of σ_B^2 since $\Sigma_i w_i = 1$. Interestingly, $n^\star$ is used by economists as a measure of market concentration under the name of Herfindahl-Hirschman Index.

We know from Chapter 8 that, at order two in volatilities of volatilities, the smile is determined by the spot/variance and variance/variance covariance functions $\mu(t,u,\xi) = \frac{\langle d\ln S_t \; d\xi_t^u\rangle}{dt}$ and $\nu(t,u,u',\xi) = \frac{\langle d\xi_t^u \; d\xi_t^{u'}\rangle}{dt}$. We then focus on these objects and define cross spot/factor and factor/factor correlations so that cross-covariance functions are related simply to their diagonal counterparts.

11.2.1 A homogeneous basket

Consider the case a basket of n assets, each driven by the two-factor model of Section 7.4 with identical parameters $\nu, \theta, k_1, k_2, \rho_{X_1X_2}$ (previously noted ρ_{12}), ρ_{SX_1}, ρ_{SX_2}. For the sake of readability, we decide to carry the factor indices as subscripts rather than superscripts, in the present section.

The instantaneous spot/variance and variance/variance covariance functions of an asset S_t and its forward variances ξ_t^u are given by expressions (8.50) and (8.51), page 327:

$$\mu^{\text{diag}}(t,u,\xi) \;=\; 2\nu\alpha_\theta\sqrt{\xi^t}\xi^u\left((1-\theta)\rho^{\text{diag}}_{SX_1}e^{-k_1(u-t)} + \theta\rho^{\text{diag}}_{SX_2}e^{-k_2(u-t)}\right) \tag{11.13}$$

$$\begin{aligned}\nu^{\text{diag}}(t,u,u',\xi) \;=\; & 4\nu^2\xi^u\xi^{u'}\alpha_\theta^2\Big[(1-\theta)^2\rho^{\text{diag}}_{X_1X_1}e^{-k_1(u+u'-2t)} \\ & +\; \theta^2\rho^{\text{diag}}_{X_2X_2}e^{-k_2(u+u'-2t)} \\ & +\; \rho^{\text{diag}}_{X_1X_2}\theta(1-\theta)\left(e^{-k_1(u-t)}e^{-k_2(u'-t)} + e^{-k_2(u-t)}e^{-k_1(u'-t)}\right)\Big]\end{aligned} \tag{11.14}$$

where the *diag* subscript indicates that these covariance functions apply to spot and variances of the same asset. For the diagonal case, we have:

$$\rho^{\text{diag}}_{SX_1} = \rho_{SX_1}, \quad \rho^{\text{diag}}_{SX_2} = \rho_{SX_2}, \quad \rho^{\text{diag}}_{X_1X_1} = 1, \quad \rho^{\text{diag}}_{X_2X_2} = 1, \quad \rho^{\text{diag}}_{X_1X_2} = \rho_{X_1X_2}$$

Upon replacing *diag* with *cross* superscripts, $\mu_{\text{cross}}, \nu_{\text{cross}}$ are given by expressions (11.13) and (11.14) as well, except the *diag* superscript is replaced with *cross*.

Correlations $\rho_{X_1X_2}, \rho_{SX_1}, \rho_{SX_2}$ are set. We need to define $\rho^{\text{cross}}_{SX_1}$, $\rho^{\text{cross}}_{SX_2}$, $\rho^{\text{cross}}_{X_1X_1}$, $\rho^{\text{cross}}_{X_2X_2}$, $\rho^{\text{cross}}_{X_1X_2}$ as well as the asset-asset correlation ρ_{SS}.

Our aim is to parametrize the multi-asset model so that μ^{cross} and ν^{cross} are defined in terms of their diagonal counterparts. Specifically, we require that the term structure of μ^{cross} and ν^{cross} – that is their dependency on $u-t$ and $u'-t$ – match that of μ^{diag} and ν^{diag}, once corrected for the effect of the term structure of forward variances. Note that prefactors $\sqrt{\xi^t}\xi^u$ and $\xi^u\xi^{u'}$ in (11.13) and (11.14) would no longer appear if we worked with instantaneous spot/variance and variance/variance *correlations* rather than *covariances.*

We introduce parameter $\chi_{S\sigma}$ and define cross-correlations $\rho^{\text{cross}}_{SX_1}$, $\rho^{\text{cross}}_{SX_2}$ through:

$$\begin{cases}\rho^{\text{cross}}_{SX_1} &=& \chi_{S\sigma}\,\rho^{\text{diag}}_{SX_1} \\ \rho^{\text{cross}}_{SX_2} &=& \chi_{S\sigma}\,\rho^{\text{diag}}_{SX_2}\end{cases}$$

$\chi_{S\sigma}$ simply measures how much larger or smaller cross spot/volatility correlations are with respect to their diagonal counterparts. Using (11.13) we can check that, for a flat term structure of VS volatilities: $\mu^{\text{cross}}(t,u,\xi) = \chi_{S\sigma}\,\mu^{\text{diag}}(t,u,\xi)$ or, equivalently:

$$\rho^{\text{cross}}(S,\widehat{\sigma}_T) = \chi_{S\sigma}\,\rho^{\text{diag}}(S,\widehat{\sigma}_T)$$

where $\rho^{\text{diag}}(S,\widehat{\sigma}_T)$ is the instantaneous correlation between S and its implied VS volatility of maturity T and $\rho^{\text{cross}}(S,\widehat{\sigma}_T)$ is its cross-counterpart.

Turning now to ν, we similarly introduce coefficient $\chi_{\sigma\sigma}$ and define $\rho^{\text{cross}}_{X_1X_1}, \rho^{\text{cross}}_{X_2X_2}, \rho^{\text{cross}}_{X_1X_2}$ as:

$$\begin{cases} \rho^{\text{cross}}_{X_1X_1} &= \chi_{\sigma\sigma}\,\rho^{\text{diag}}_{X_1X_1} \\ \rho^{\text{cross}}_{X_2X_2} &= \chi_{\sigma\sigma}\,\rho^{\text{diag}}_{X_2X_2} \\ \rho^{\text{cross}}_{X_1X_2} &= \chi_{\sigma\sigma}\,\rho^{\text{diag}}_{X_1X_2} \end{cases}$$

Thus,

$$\rho^{\text{cross}}_{X_1X_1} = \chi_{\sigma\sigma},\ \rho^{\text{cross}}_{X_2X_2} = \chi_{\sigma\sigma},\ \rho^{\text{cross}}_{X_1X_2} = \chi_{\sigma\sigma}\,\rho_{X_1X_2}$$

One can check on (11.14) that with this parametrization, the instantaneous correlation between ξ_i^u and $\xi_j^{u'}$ where i,j denote different assets is equal to $\chi_{\sigma\sigma}$ times the correlation of ξ_i^u and $\xi_i^{u'}$. For a flat term structure of VS volatilities $\nu^{\text{cross}}(t,u,u',\xi) = \chi_{\sigma\sigma}\nu^{\text{diag}}(t,u,u',\xi)$ or, equivalently:

$$\rho^{\text{cross}}(\widehat{\sigma}_T,\widehat{\sigma}_{T'}) = \chi_{\sigma\sigma}\,\rho^{\text{diag}}(\widehat{\sigma}_T,\widehat{\sigma}_{T'})$$

where $\rho^{\text{diag}}(\widehat{\sigma}_T,\widehat{\sigma}_{T'})$ is the instantaneous correlation between VS volatilities of maturities T, T' of the same underlying and $\rho^{\text{cross}}(\widehat{\sigma}_T,\widehat{\sigma}_{T'})$ is a similarly-defined quantity for correlations of VS volatilities of different underlyings.

Finally we introduce the (uniform) correlation between spot processes, ρ_{SS}. Each asset is associated to 3 Brownian motions: W_i^S, W_i^1, W_i^2. The global correlation matrix is thus of dimension $3n\times 3n$. What are the conditions on $\rho_{SS},\ \chi_{\sigma\sigma},\ \chi_{S\sigma}$ so that it is positive?

Conditions on $\rho_{SS},\ \chi_{\sigma\sigma},\ \chi_{S\sigma}$

Let us compute the eigenvalues of matrix Ω, defined by: $\Omega = \frac{1}{dt}\langle dU dU^{\intercal}\rangle$ where $U^{\intercal} = (W_1^S\cdots W_n^S,\ W_1^1\cdots W_n^1,\ W_1^2\cdots W_n^2)$.

Let λ be an eigenvalue of Ω, associated to the eigenvector T, whose components we write $T^{\intercal} = (s_1\cdots s_n,\ x_1\cdots x_n,\ y_1\cdots y_n)$. Expressing that $\Omega T = \lambda T$ we get, for $1\le i\le n$:

$$\begin{cases} \left(s_i + \rho_{SS}\overline{\Sigma}_j s_j\right) + \rho_{SX_1}\!\left(x_i + \chi_{S\sigma}\overline{\Sigma}_j x_j\right) + \rho_{SX_2}\!\left(y_i + \chi_{S\sigma}\overline{\Sigma}_j y_j\right) &= \lambda s_i \\ \rho_{SX_1}\!\left(s_i + \chi_{S\sigma}\overline{\Sigma}_j s_j\right) + \left(x_i + \chi_{\sigma\sigma}\overline{\Sigma}_j x_j\right) + \rho_{X_1X_2}\!\left(y_i + \chi_{\sigma\sigma}\overline{\Sigma}_j y_j\right) &= \lambda x_i \\ \rho_{SX_2}\!\left(s_i + \chi_{S\sigma}\overline{\Sigma}_j s_j\right) + \rho_{X_1X_2}\!\left(x_i + \chi_{\sigma\sigma}\overline{\Sigma}_j x_j\right) + \left(y_i + \chi_{\sigma\sigma}\overline{\Sigma}_j y_j\right) &= \lambda y_i \end{cases}$$

where $\overline{\Sigma}_j$ is a shorthand notation for $\Sigma_{j\neq i}$ and we have removed the *diag* superscripts in $\rho^{\rm diag}_{X_1X_2}$, $\rho^{\rm diag}_{SX_1}$, $\rho^{\rm diag}_{SX_2}$ for notational economy. The correlation structure is invariant under permutation of the i,j indices and in particular under translation: $\rho(W^{\circ}_i, W^{\bullet}_j) = \rho(W^{\circ}_{i+k \bmod n}, W^{\bullet}_{j+k \bmod n})\ \forall k$ where $\circ, \bullet$ stand for $S, 1, 2$.

We thus diagonalize Ω in the basis of eigenvectors of the discrete translation operator. T is parametrized with the 4 numbers θ_k, s, x, y:

$$T^{\mathsf{T}} = (se^{\boldsymbol{i}\theta_k}\cdots se^{\boldsymbol{i}n\theta_k},\ xe^{\boldsymbol{i}\theta_k}\cdots xe^{\boldsymbol{i}n\theta_k}, ye^{\boldsymbol{i}\theta_k}\cdots ye^{\boldsymbol{i}n\theta_k})$$

with $\theta_k = \frac{2k\pi}{n}$ where $k = 0\cdots n-1$ and $\boldsymbol{i} = \sqrt{-1}$. We have:

$$\sum_{j=0, j\neq i}^{n-1} e^{\boldsymbol{i}j\theta_k} = \sum_{j=i+1}^{n-1+i} e^{\boldsymbol{i}j\theta_k} = e^{\boldsymbol{i}i\theta_k}\sum_{j=1}^{n-1} e^{\boldsymbol{i}j\theta_k} = \begin{cases} (n-1)e^{\boldsymbol{i}i\theta_k} & k=0 \\ -e^{\boldsymbol{i}i\theta_k} & k\neq 0 \end{cases}$$

We thus define $A_{SS}(k)$ by:

$$A_{SS}(k) = \left(1 + \rho_{SS}\Sigma_{j=1}^{j=n-1} e^{\boldsymbol{i}j\theta_k}\right) = \begin{cases} 1+(n-1)\rho_{SS} & k=0 \\ 1-\rho_{SS} & k\neq 0 \end{cases} \tag{11.15}$$

$A_{\sigma\sigma}(k),\ A_{S\sigma}(k)$ are defined similarly in terms of $\chi_{\sigma\sigma}, \chi_{S\sigma}$. We have:

$$\begin{cases} A_{\sigma\sigma}(k=0) = 1+(n-1)\chi_{\sigma\sigma} \\ A_{\sigma\sigma}(k\neq 0) = 1-\chi_{\sigma\sigma} \end{cases} \quad \text{and} \quad \begin{cases} A_{S\sigma}(k=0) = 1+(n-1)\chi_{S\sigma} \\ A_{S\sigma}(k\neq 0) = 1-\chi_{S\sigma} \end{cases}$$

The resulting system for s, x, y is:

$$\begin{cases} A_{SS}s + \rho_{SX_1}A_{S\sigma}x + \rho_{SX_2}A_{S\sigma}y = \lambda s \\ \rho_{SX_1}A_{S\sigma}s + A_{\sigma\sigma}x + \rho_{X_1X_2}A_{\sigma\sigma}y = \lambda x \\ \rho_{SX_2}A_{S\sigma}s + \rho_{X_1X_2}A_{\sigma\sigma}x + A_{\sigma\sigma}y = \lambda y \end{cases}$$

The condition is thus that the following symmetric matrix

$$\omega(k) = \begin{pmatrix} A_{SS} & \rho_{SX_1}A_{S\sigma} & \rho_{SX_2}A_{S\sigma} \\ \rho_{SX_1}A_{S\sigma} & A_{\sigma\sigma} & \rho_{X_1X_2}A_{\sigma\sigma} \\ \rho_{SX_2}A_{S\sigma} & \rho_{X_1X_2}A_{\sigma\sigma} & A_{\sigma\sigma} \end{pmatrix}$$

be positive. This implies in particular that $A_{SS}(k) \geq 0$, $A_{\sigma\sigma}(k) \geq 0\ \forall k$, which places the following bounds on $\rho_{SS}, \chi_{\sigma\sigma}$:

$$-\frac{1}{n-1} \leq \rho_{SS}, \chi_{\sigma\sigma} \leq 1$$

These conditions, in the equity context, are not very restrictive.

We now define a "correlation matrix" $\rho(k)$ obtained by rescaling the symmetric matrix $\omega(k)$: $\rho_{ij}(k) = \frac{\omega_{ij}(k)}{\sqrt{\omega_{ii}(k)\omega_{jj}(k)}}$:

$$\rho(k) = \begin{pmatrix} 1 & \zeta\rho_{SX_1} & \zeta\rho_{SX_2} \\ \zeta\rho_{SX_1} & 1 & \rho_{X_1X_2} \\ \zeta\rho_{SX_2} & \rho_{X_1X_2} & 1 \end{pmatrix} \tag{11.16}$$

where $\zeta(k)$ is given by:

$$\zeta(k) = \frac{A_{S\sigma}(k)}{\sqrt{A_{SS}(k)A_{\sigma\sigma}(k)}} = \begin{cases} \frac{1+(n-1)\chi_{S\sigma}}{\sqrt{(1+(n-1)\rho_{SS})(1+(n-1)\chi_{\sigma\sigma})}} & k=0 \\ \frac{1-\chi_{S\sigma}}{\sqrt{(1-\rho_{SS})(1-\chi_{\sigma\sigma})}} & k \neq 0 \end{cases}$$

$\rho(k)$ in (11.16) is in fact the correlation matrix of the single-asset case, but with spot/volatility correlations rescaled by the factor $\zeta(k)$, which may be smaller or larger than 1 depending on the values of $\rho_{SS}, \chi_{\sigma\sigma}, \chi_{S\sigma}$.

$\chi_{S\sigma}$ is allowed to be larger than 1, unlike ρ_{SS} and $\chi_{\sigma\sigma}$. This will be needed in situations when the index skew is much steeper than the component skew – see the discussion in Section 11.1.1.

The case $\zeta(k) \equiv 1$ is obtained by taking $\chi_{S\sigma} = \chi_{\sigma\sigma} = \rho_{SS}$. $\rho(k)$ is then equal to the single-asset correlation matrix thus the positivity of the $3n \times 3n$ global covariance matrix is ensured. In this model, cross-correlations (spot/spot, spot/volatility, volatility/volatility) are all equal to ρ_{SS} times their diagonal counterparts. This is the case of the multi-asset local volatility model.[4]

Conclusion

Given correlations $\rho_{SX_1}, \rho_{SX_1}, \rho_{X_1X_2}$ for the single-asset case, we parametrize the $3n \times 3n$ global correlation matrix of the multi-asset case by introducing three additional parameters: $\rho_{SS}, \chi_{\sigma\sigma}, \chi_{S\sigma}$.

ρ_{SS} is the spot/spot correlation. $\chi_{\sigma\sigma}, \chi_{S\sigma}$ define cross spot/volatility and volatility/volatility covariance functions in terms of their diagonal counterparts.

ρ_{SS} and $\chi_{\sigma\sigma}$ are restricted to the interval $[-\frac{1}{n-1}, 1]$. The necessary and sufficient condition on $\rho_{SS}, \chi_{\sigma\sigma}, \chi_{S\sigma}$ for the global correlation matrix to be positive is that the following effective correlation matrix for the single-asset case:

$$\begin{pmatrix} 1 & \zeta\rho_{SX_1} & \zeta\rho_{SX_2} \\ \zeta\rho_{SX_1} & 1 & \rho_{X_1X_2} \\ \zeta\rho_{SX_2} & \rho_{X_1X_2} & 1 \end{pmatrix}$$

be positive for the two following values of ζ:

$$\zeta = \frac{1+(n-1)\chi_{S\sigma}}{\sqrt{(1+(n-1)\rho_{SS})\,(1+(n-1)\chi_{\sigma\sigma})}} \quad (11.17a)$$

$$\zeta = \frac{1-\chi_{S\sigma}}{\sqrt{(1-\rho_{SS})\,(1-\chi_{\sigma\sigma})}} \quad (11.17b)$$

in addition to the $\zeta = 1$ case. This is the case if $|\zeta\rho_{SX_1}| \leq 1,\ |\zeta\rho_{SX_2}| \leq 1$ and:

$$-1 \leq \frac{\rho_{X_1X_2} - \zeta^2\rho_{SX_1}\rho_{SX_2}}{\sqrt{1-\zeta^2\rho_{SX_1}^2}\sqrt{1-\zeta^2\rho_{SX_2}^2}} \leq 1$$

[4]The local volatility model's assumption of setting volatility/volatility correlations identical to spot/spot correlations is not very realistic. Typically, correlations of implied volatilities of underlyings belonging to the same index are usually larger than the correlations of the underlyings themselves.

It is easy to check, going backwards through the derivation in Section 11.2.1, that these necessary conditions are also sufficient: we have effectively built the $3n$ eigenvectors of Ω.[5]

While in what follows we will be using a homogeneous model, in practice, parameters $\nu, \theta, k_1, k_2, \rho_{X_1X_2}, \rho_{SX_1}, \rho_{SX_2}$ are different for different components. In particular, even though we typically use identical values for $\nu, \theta, k_1, k_2, \rho_{X_1X_2}$ so that volatilities of volatilities are identical, we may use different values for $\rho^{\text{diag}}_{SX_1}, \rho^{\text{diag}}_{SX_2}$. Consider two components S_i, S_j. Should we set $\rho^{\text{cross}}_{S^iX_1^j} = \chi_{S\sigma}\rho^{\text{diag}}_{S^jX_1^j}$ or $\rho^{\text{cross}}_{S^iX_1^j} = \chi_{S\sigma}\rho^{\text{diag}}_{S^iX_1^i}$?

When setting the cross spot/volatility correlations of S_i and $\widehat{\sigma}_j^T$, it is more reasonable to derive the T-dependence of $\rho(S_i, \widehat{\sigma}_j^T)$ from the term-structure of $\rho(S_j, \widehat{\sigma}_j^T)$ rather than that of $\rho(S_i, \widehat{\sigma}_i^T)$. We would thus set:

$$\rho^{\text{cross}}_{S^iX_1^j} \;=\; \chi_{S\sigma}\; \rho^{\text{diag}}_{S^jX_1^j}$$

and likewise for $\rho^{\text{cross}}_{S^iX_2^j}$. Going back to implied volatilities, this is equivalent to setting:

$$\rho(S^i, \widehat{\sigma}_T^j) \;=\; \chi_{S\sigma}\; \rho(S^j, \widehat{\sigma}_T^j)$$

Likewise, with respect to volatility/volatility correlations, we could set

$$\begin{aligned} \text{either } \rho(\widehat{\sigma}_T^i, \widehat{\sigma}_{T'}^j) \;&=\; \chi_{\sigma\sigma}\; \rho(\widehat{\sigma}_T^i, \widehat{\sigma}_{T'}^i) \\ \text{or } \rho(\widehat{\sigma}_T^i, \widehat{\sigma}_{T'}^j) \;&=\; \chi_{\sigma\sigma}\; \rho(\widehat{\sigma}_T^j, \widehat{\sigma}_{T'}^j) \end{aligned}$$

In practice, volatility/volatility correlations are sufficiently similar across underlyings that either choice results in very similar values of $\chi_{\sigma\sigma}$ – see the examples in Table 11.2 below.

The global correlation matrix we thus build for a non-homogeneous basket is not guaranteed to be positive – but is not expected to be very negative either; in case it is we use the algorithm proposed in [61] to generate the closest positive correlation matrix.[6]

11.2.2 Realized values of $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$

Typical realized values of $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ for indexes appear in Table 11.2. For each pair of underlyings (S^1, S^2), two values of $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ have been computed, obtained by averaging respectively either $\rho(\widehat{\sigma}_T^1, S^2)/\rho(\widehat{\sigma}_T^1, S^1)$ or

[5] Ω has 3 non-degenerate eigenvalues (obtained with $k = 0$) and 3 degenerate eigenvalues (obtained with $k \neq 0$, thus each with $(n-1)$ degeneracy).

[6] The algorithm in [61] generates the correlation matrix that is closest to a candidate input matrix. In the formula for the matrix distance, appropriate weights can be used so as to enforce, for example, that diagonal correlations $\rho^{\text{diag}}_{S^iX_1^i}, \rho^{\text{diag}}_{S^iX_2^i}, \rho^{\text{diag}}_{X_1^iX_2^i}$, as well as spot/spot correlations are least altered.

$\rho(\widehat{\sigma}_T^2, S^1)/\rho(\widehat{\sigma}_T^2, S^2)$ over several maturities. Likewise, for the determination of $\chi_{\sigma\sigma}$, two values are obtained by averaging either $\rho(\widehat{\sigma}_T^1, \widehat{\sigma}_{T'}^2)/\rho(\widehat{\sigma}_T^1, \widehat{\sigma}_{T'}^1)$ or $\rho(\widehat{\sigma}_T^1, \widehat{\sigma}_{T'}^2)/\rho(\widehat{\sigma}_T^2, \widehat{\sigma}_{T'}^2)$, for all (T, T') couples.

Note that it is important to use asynchronous estimators – lest we underestimate correlations; see [12].

	June 2008 - June 2013					June 2003 - June 2008				
	S&P500 Stoxx50	Stoxx50 FTSE	S&P500 NIKKEI	Stoxx50 NIKKEI	NIKKEI KOSPI	S&P500 Stoxx50	Stoxx50 FTSE	S&P500 NIKKEI	Stoxx50 NIKKEI	NIKKEI KOSPI
ρ	83%	89%	55%	60%	67%	71%	86%	45%	59%	66%
$\chi_{S\sigma}$	87%	95%	59%	63%	85%	78%	89%	56%	61%	108%
$\chi_{S\sigma}$	86%	93%	71%	78%	78%	75%	91%	53%	77%	65%
$\chi_{\sigma\sigma}$	84%	92%	54%	60%	72%	72%	85%	31%	40%	50%
$\chi_{\sigma\sigma}$	84%	92%	55%	61%	73%	71%	85%	31%	41%	50%

Table 11.2: Historical spot/spot correlations and values of $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ for different pairs of indexes, measured on two 5-year samples. $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ are evaluated by averaging pairwise ratios of spot/volatility and volatility/volatility correlations using implied ATMF volatilities with maturities 3 months, 6 months, 1 year, 2 years. The asynchronous estimator in [12] has been used.

Empirically, diagonal volatility/volatility correlations across indexes are very similar, thus the two ways of estimating $\chi_{\sigma\sigma}$ yield very similar value; we can thus use an average of the two estimates.

This is less the case for $\chi_{S\sigma}$. Indeed, historical regimes of spot/volatility correlations for the Nikkei index, for example, can be quite different from those of the S&P 500 and Euro Stoxx 50 indexes. This is also reflected in the different behavior of their SSRs – see Figure 9.10, page 381.

11.3 The ATMF basket skew

After focusing in Section 11.1 on short maturities, we derive now an approximation of the basket ATMF skew at order one in volatility of volatility by using expression (8.22) which involves the spot/variance covariance function evaluated at order one in volatility of volatility. We use the notation $\hat{\sigma}_T^B$ and $\mathcal{S}_T^B$ for the basket VS volatility and ATMF skew and denote *basket* forward variances by ζ_t^u.

We will assume that the underlying model is time-homogeneous and that the VS term structures of all components are flat and identical.

From equation (8.22), page 314, $\mathcal{S}_T^B$ is given by:

$$\mathcal{S}_T^B = \hat{\sigma}_T^B \frac{C_B^{x\xi}(T)}{2\big(\big(\hat{\sigma}_T^B\big)^2 T\big)^2} \tag{11.18}$$

where $C_B^{x\xi}(T)$ is the doubly-integrated spot/variance covariance function:

$$C_B^{x\xi}(T) = \int_0^T dt \int_t^T du\, \mu_B(t,u) \tag{11.19}$$

$$\mu_B(t,u) = \lim_{dt\to 0} \frac{1}{dt} E_t\Big[d\ln B_t d\zeta_t^u\Big] \tag{11.20}$$

where $B_t = \Sigma_i w_i S_{it}$. $\mu_B(t,u)$ is evaluated at order one in volatility of volatility on the initial basket variance curve.

Let us assume that the basket is homogeneous and equally weighted; spot correlations ρ_{ij}, $i \neq j$ are all identical, equal to ρ and we use the correlation parametrization of Section 11.2.1.

We also make the approximation that weights are frozen, equal to their initial values, which we denote by w_{i0}. Basket forward variances ζ_t^T are given by:

$$\zeta_t^u = E_t[\sum_{ij} w_i w_j \rho_{ij} \sqrt{\xi_{iu}^u}\sqrt{\xi_{ju}^u}] \tag{11.21}$$

$$= \sum_{ij} \rho_{ij} w_{i0} w_{j0} E_t\Big[\sqrt{\xi_{iu}^u}\sqrt{\xi_{ju}^u}\Big] \tag{11.22}$$

Thus

$$d\zeta_t^u = \sum_{ij} w_{i0} w_{j0} \rho_{ij} E_t\Big[\frac{\sqrt{\xi_{ju}^u}}{\sqrt{\xi_{iu}^u}} d\xi_{iu}^u\Big]$$

At order one in volatility of volatility, it is sufficient to replace the prefactor inside the expectation with its value in the unperturbed state, that is using forward variances values at $t=0$. At this order:

$$d\zeta_t^u = \sum_{ij} w_{i0} w_{j0} \rho_{ij} \frac{\sqrt{\xi_{j0}^u}}{\sqrt{\xi_{i0}^u}} E_t\Big[d\xi_{iu}^u\Big] = \sum_{ij} w_{i0} w_{j0} \rho_{ij} \frac{\sqrt{\xi_{j0}^u}}{\sqrt{\xi_{i0}^u}} d\xi_{it}^u$$

Let us now compute $\mu_B(t,u)$. We have:

$$\frac{dB_t}{B_t} = \sum_i w_{k0} \frac{dS_{k,t}}{S_{k,t}}$$

We then get the expression of $\mu_B(t,u)$ as a function of the diagonal spot/variance covariance function, which we denote compactly by $\mu(u-t) = \frac{1}{dt}\langle d\ln S_{it} d\xi_{it}^u\rangle$, as we have assumed a time-homogeneous model:

$$\begin{aligned}
\mu_B(t,u) &= \frac{1}{dt}\langle d\ln B_t d\zeta_t^u\rangle = \sum_{ijk} w_{i0}w_{j0}w_{k0}\rho_{ij}\frac{\sqrt{\xi_{j0}^u}}{\sqrt{\xi_{i0}^u}}\frac{\langle d\ln S_{kt} d\xi_{it}^u\rangle}{dt} \\
&= \sum_{ij} w_{i0}^2 w_{j0}\rho_{ij}\frac{\sqrt{\xi_{j0}^u}}{\sqrt{\xi_{i0}^u}}\mu(u-t) + \chi_{S\sigma}\sum_{i\neq k,j} w_{i0}w_{j0}w_{k0}\rho_{ij}\frac{\sqrt{\xi_{j0}^u}}{\sqrt{\xi_{i0}^u}}\mu(u-t)
\end{aligned}$$

For a homogeneous basket with flat and identical term structures of VS volatilities, this simplifies to:

$$\mu_B(t,u) = \frac{1+(n-1)\chi_{S\sigma}}{n}\frac{1+(n-1)\rho_{SS}}{n}\mu(u-t)$$

Using expression (11.19) for $C_B^{x\xi}(T)$:

$$C_B^{x\xi}(T) = \left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]\left[\frac{1+(n-1)\rho_{SS}}{n}\right]\int_0^T dt\int_0^T du\,\mu(u-t)$$

Finally, the ATMF skew of a homogeneous basket is given, at order one in volatility of volatility, by:

$$\mathcal{S}_T^B = \frac{\hat\sigma_{0T}^B}{2\big((\hat\sigma_{0T}^B)^2 T\big)^2}\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]\left[\frac{1+(n-1)\rho_{SS}}{n}\right]\int_0^T dt\int_t^T du\,\mu(u-t) \tag{11.23}$$

where $\hat\sigma_{0T}^B$ is the basket VS volatility at order zero in volatility of volatility, that is with zero volatility of volatility. Assuming that the VS term structures of all components are flat, and equal to $\widehat{\sigma}^2$, and that the basket is equally weighted, $\hat\sigma_{0T}^B$ is given by:

$$\left(\hat\sigma_{0T}^B\right)^2 = \frac{1+(n-1)\rho_{SS}}{n}\widehat{\sigma}^2$$

Using this expression of $\hat\sigma_{0T}^B$ in (11.23) and remembering that the single-asset skew is given by expression (8.22), page 314, with $C^{x\xi}$ given by (8.13):

$$\mathcal{S}_T = \frac{\widehat{\sigma}}{2(\widehat{\sigma}^2 T)^2}\int_0^T dt\int_t^T du\,\mu(u-t)$$

we get our final expression for the basket skew, as a function of the components' skew:

$$\mathcal{S}_T^B = \mathcal{S}_T\sqrt{\frac{n}{1+(n-1)\rho_{SS}}}\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right] \tag{11.24}$$

Thus an implied value for $\chi_{S\sigma}$ can be backed out of the ratio of basket skew to component skew.

Let us check that in the limit $T \to 0$, (11.24) yields back the short-maturity result (11.5b) for $\mathcal{S}_T^B$. Using the fact that $\int_0^T dt \int_t^T du = \frac{T^2}{2}$ and that $\mathcal{S}_{T=0} = \frac{\mu(0)}{4\widehat{\sigma}_{T=0}^3}$, expression (11.23) for $\mathcal{S}_T^B$ yields:

$$
\begin{aligned}
\mathcal{S}_{T=0}^B &= \frac{1}{4\left(\hat{\sigma}_{T=0}^B\right)^3}\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]\left[\frac{1+(n-1)\rho_{SS}}{n}\right]\mu(0) \\
&= \frac{1}{4\left(\hat{\sigma}_{T=0}^B\right)^3}\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]\left[\frac{1+(n-1)\rho_{SS}}{n}\right]4\widehat{\sigma}_{T=0}^3\mathcal{S}_{T=0} \\
&= \mathcal{S}_{T=0}\left(\frac{\widehat{\sigma}_{T=0}}{\hat{\sigma}_{T=0}^B}\right)^3\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]\left[\frac{1+(n-1)\rho_{SS}}{n}\right] \qquad (11.25)
\end{aligned}
$$

where we have omitted the 0 subscript in $\hat{\sigma}_{0T}^B$, as, for $T=0$, $\hat{\sigma}_{0T}^B = \hat{\sigma}_T^B$. (11.25) is identical to (11.5b) for $T=0$.

Again, expression (11.24) for $\mathcal{S}_T^B$ involves the factor $\frac{1+(n-1)\chi_{S\sigma}}{n}$: the portion contributed to the basket skew by the components' skew scales like $\frac{1}{n}$ and the bulk of the basket skew is generated by cross spot/volatility correlations.

11.3.1 Application to the two-factor model

In the two-factor model the dynamics of each component reads:

$$
\begin{cases}
dS_t &= \sqrt{\xi_t^t}S_t dW_t^S \\
d\xi_t^T &= 2\nu\xi_t^T\,\alpha_\theta\left((1-\theta)e^{-k_1(T-t)}dW_t^1 + \theta e^{-k_2(T-t)}dW_t^2\right)
\end{cases}
$$

where ν is the lognormal volatility of a very short-dated implied VS volatility, and $\alpha_\theta = \left((1-\theta)^2+\theta^2+2\rho_{X_1X_2}\theta(1-\theta)\right)^{-1/2}$. The cross spot/variance and variance/variance correlations are parametrized using $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ defined in Section 11.2.1 and the correlations between the S_i are equal to ρ.

For a flat term structure of VS volatilities equal to $\widehat{\sigma}$, the spot/variance covariance function $\mu(\tau)$ is given by:

$$
\mu(\tau) = (2\nu)\,\hat{\sigma}^3\alpha_\theta\left((1-\theta)\rho_{SX_1}e^{-k_1\tau} + \theta\rho_{SX_2}e^{-k_2\tau}\right) \qquad (11.26)
$$

and the component's ATMF skew at order one in volatility of volatility is given by expression (8.55):

$$
\mathcal{S}_T = \nu\alpha_\theta\left[(1-\theta)\,\rho_{SX^1}\frac{k_1T-(1-e^{-k_1T})}{(k_1T)^2} + \theta\rho_{SX^2}\frac{k_2T-(1-e^{-k_2T})}{(k_2T)^2}\right] \qquad (11.27)
$$

Using (11.24) the basket skew – at order one in volatility of volatility – is given by:

$$
\begin{aligned}
\mathcal{S}_T^B &= \nu\alpha_\theta\sqrt{\frac{n}{1+(n-1)\rho_{SS}}}\left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right] \\
&\times\left[(1-\theta)\,\rho_{SX^1}\frac{k_1T-(1-e^{-k_1T})}{(k_1T)^2} + \theta\rho_{SX^2}\frac{k_2T-(1-e^{-k_2T})}{(k_2T)^2}\right] \qquad (11.28)
\end{aligned}
$$

While, for the sake of calculating $\mathcal{S}_T^B$ we have needed $\hat{\sigma}_T^B$ at order zero in volatility of volatility only, it can be calculated exactly – with the assumption of frozen weights:

$$\begin{aligned}\left(\hat{\sigma}_T^B\right)^2 &= \frac{1}{T}\int_0^T \sum_{ij} w_{i0} w_{j0} \rho_{ij} E\left[\sqrt{\xi_{it}^t}\sqrt{\xi_{jt}^t}\right] dt \\ &= \frac{1}{T}\int_0^T \sum_{ij} w_{i0} w_{j0} \rho_{ij} \sqrt{\xi_{i0}^t}\sqrt{\xi_{j0}^t} h_{ij}(t) dt\end{aligned}$$

with $h_{ij}(\tau)$ given by: $h_{ii}(\tau) = 1$, $h_{i \neq j}(\tau) = h(\tau)$, where $h(\tau)$ reads:

$$h(\tau) = e^{-\nu^2 \alpha_\theta^2 (1-\chi_{\sigma\sigma})\left((1-\theta)^2 \frac{1-e^{-2k_1\tau}}{2k_1} + \theta^2 \frac{1-e^{-2k_2\tau}}{2k_2} + 2\theta(1-\theta)\rho_{X_1X_2} \frac{1-e^{-(k_1+k_2)\tau}}{k_1+k_2}\right)} \tag{11.29}$$

For flat and identical term-structures of VS volatilities:

$$\left(\hat{\sigma}_T^B\right)^2 = \widehat{\sigma}^2 \, \frac{1+(n-1)\rho_{SS} \, \frac{1}{T}\int_0^T h\left(\tau\right) d\tau}{n} \tag{11.30}$$

11.3.2 Numerical examples

We now test the accuracy of formulas (11.24), (11.30) for $\mathcal{S}_T^B$, $\hat{\sigma}_T^B$ with the two-factor model.

To generate stock-like parameters, we start from the index-like parameters in Table 8.2, page 329, and reduce ρ_{SX^1}, ρ_{SX^2} by about 30% so as to reduce the ATMF skew by the same relative amount. The other parameters are left unchanged. The basket component's parameters we obtain appear in Table 11.3.

ν	θ	k_1	k_2	$\rho_{X^1X^2}$	ρ_{SX^1}	ρ_{SX^2}
174%	0.245	5.35	0.28	0%	-53.0%	-33.9%

Table 11.3: Basket component's parameters in the two-factor model.

The component's term-structure of volatilities of volatilities is that corresponding to Set II in Figure 7.2, page 229. The VS term structure of volatilities is flat at 30%. We take $n = 10$ components and will use $\rho_{SS} = 60\%$ for the component's correlations. We use vanishing rate and repo.

Let us first set $\chi_{S\sigma} = 0$. The basket smile – with $\chi_{\sigma\sigma} = 0$ – is shown in Figure 11.2, along with the component's smile for comparison. The basket ATMF skew almost vanishes – a reflection of the fact that the basket ATMF skew is mostly generated by cross spot/volatility correlations.

We now use more reasonable values for $\chi_{S\sigma}$. Figure 11.3 shows basket smiles generated with ($\chi_{S\sigma} = 80\%,\ \chi_{\sigma\sigma} = 80\%$) and ($\chi_{S\sigma} = 115\%,\ \chi_{\sigma\sigma} = 80\%$).

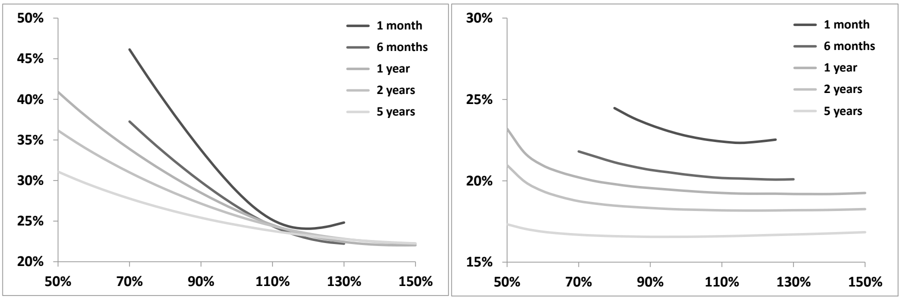

Figure 11.2: Left: component's smile with parameters of Table 11.3. Right: basket smile with $\rho_{SS} = 60\%,\ \chi_{S\sigma} = 0$ and $\chi_{\sigma\sigma} = 0$.

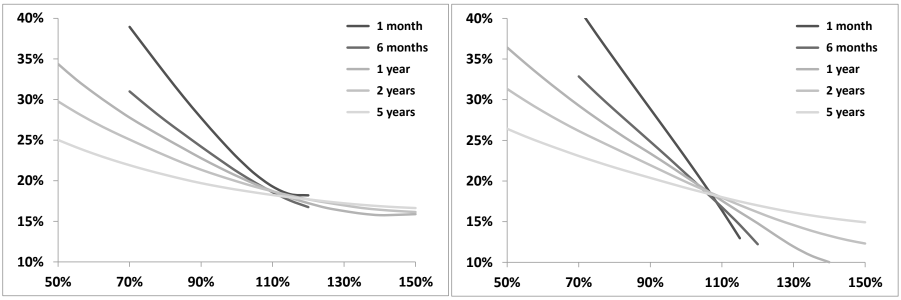

Figure 11.3: Basket smile with: $\rho_{SS} = 60\%,\ \chi_{\sigma\sigma} = 80\%$. Left: $\chi_{S\sigma} = 80\%$. Right: $\chi_{S\sigma} = 115\%$.

Comparing with the left-hand graph in Figure 11.2 it is apparent that the basket ATMF skew with ($\chi_{S\sigma} = 115\%,\ \chi_{\sigma\sigma} = 80\%$) is now steeper than the component's ATMF skew.

How accurate are formulas for (11.24),(11.30) for $\mathcal{S}_T^B, \hat{\sigma}_T^B$?

Basket ATMF skew

Figure (11.4) shows the 95/105 ATMF basket skew ($\hat{\sigma}^B_{K=0.95S_0,T} - \hat{\sigma}^B_{K=1.05S_0,T}$), either calculated in a Monte Carlo simulation, or evaluated using (11.24), with $\mathcal{S}_T$ given either by the actual component's ATMF skew, or given by (11.27), for two different values of $\chi_{S\sigma}$. The left-hand graph confirms that expression (11.24) is very accurate. The small overestimation of $\mathcal{S}_T^B$ in the right-hand graph is due to the slight overestimation of the component's skew in (11.27), evidenced in Figure 8.3, page 331.

Figure (11.5) provides another confirmation that the basket ATMF skew is controlled by $\chi_{S\sigma}$. Here we have varied $\chi_{\sigma\sigma}$ while keeping $\rho_{SS} = 60\%$ and $\chi_{S\sigma} = 80\%$ fixed. The ATMF skew hardly changes when $\chi_{\sigma\sigma}$ is varied. For the three values

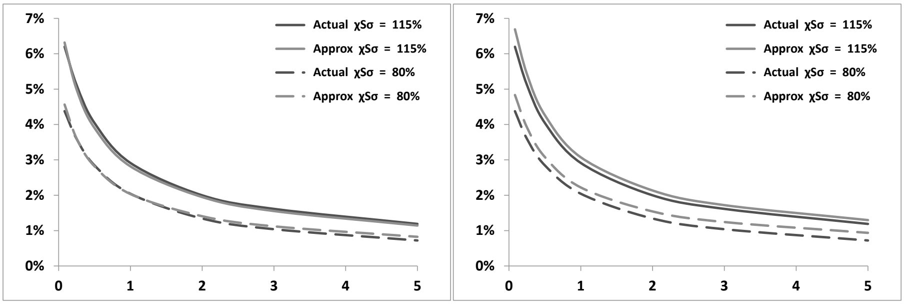

Figure 11.4: Left: basket 95/105 skew in volatility points (Actual) compared to formula (11.24) (Approx) where the actual component's 95/105 skew has been used. Right: basket 95/105 skew (Actual) compared to formula (11.28) (Approx). Maturities are in years. The two values of $\chi_{S\sigma} = 80\%,\ 115\%$ have been used. All other parameters are kept constant, including $\chi_{\sigma\sigma} = 80\%$.

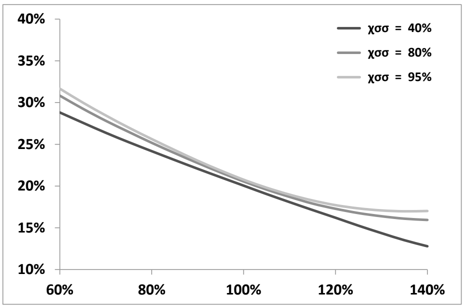

Figure 11.5: One-year basket smiles for different values of $\chi_{\sigma\sigma}$. All other parameters are kept constant: $\rho_{SS} = 60\%$ and $\chi_{S\sigma} = 80\%$.

of $\chi_{\sigma\sigma}$ used: $40\%,\ 80\%,\ 95\%$, the 95/105 one-year skew values are respectively: $2.01\%, 2.05\%, 2.06\%$.

Using identical and flat term structures of VS volatilities for the basket components, as well as identical sets of parameters, leads to the particularly simple formulas (11.24) and (11.28) but the derivation of $\mathcal{S}_T^B$ in the general case presents no particular difficulty. One only needs to express $\mu_B(t,u)$ as a function of the component's spot/volatility covariance functions $\mu(t,u,\xi)$, which are given by expression (8.50) in the two-factor model.

Basket VS volatility

$\hat{\sigma}_T^B$, either evaluated in a Monte Carlo simulation or given by (11.30), is graphed in Figure 11.6 for different values of $\chi_{S\sigma}$. In expression (11.30), with the approximation

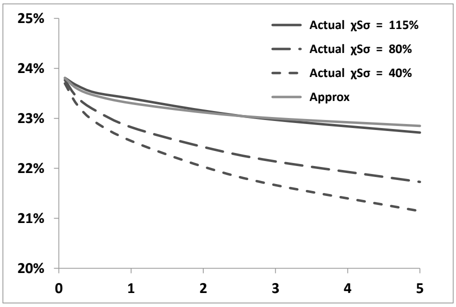

Figure 11.6: $\hat{\sigma}_T^B$, either evaluated in a Monte Carlo simulation of the two-factor model, for three values of $\chi_{S\sigma}$: $40\%,\ 80\%,\ 115\%$, or given by expression (11.30). Maturities are in years. The component's parameters are listed in Table 11.3 and $\rho_{SS} = 60\%,\ \chi_{\sigma\sigma} = 80\%$.

of frozen weights, only spot/spot (ρ_{SS}) and volatility/volatility ($\chi_{\sigma\sigma}$) correlations appear.

Figure 11.6 makes it clear that this assumption is not adequate: $\hat{\sigma}_T^B$ does depend on cross spot/volatility correlations.

11.3.3 Mimicking the local volatility model

Can our two-factor stochastic volatility model mimic a multi-asset local volatility model? In the local volatility model, $\rho^{\mathrm{cross}}(S, \widehat{\sigma}_T) = -\rho_{SS}$ and $\rho^{\mathrm{cross}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'}) = \rho_{SS}$. We have: $\rho^{\mathrm{diag}}(S, \widehat{\sigma}_T) = -1$ and $\rho^{\mathrm{diag}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'}) = 1$.[7] Thus, in the local volatility model:

$$\begin{aligned} \rho^{\mathrm{cross}}(S, \widehat{\sigma}_T) &= \rho_{SS}\ \rho^{\mathrm{diag}}(S, \widehat{\sigma}_T) \\ \rho^{\mathrm{cross}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'}) &= \rho_{SS}\ \rho^{\mathrm{diag}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'}) \end{aligned}$$

Thus, with regard to spot/volatility and volatility/volatility correlations, the local volatility model can be viewed as a particular breed of multi-asset stochastic volatility with $\chi_{S\sigma},\ \chi_{\sigma\sigma}$ given by:

$$\chi_{S\sigma} = \rho_{SS} \tag{11.31a}$$

$$\chi_{\sigma\sigma} = \rho_{SS} \tag{11.31b}$$

With this choice of cross-parametrization the values of ζ in (11.17) are both equal to 1, thus positivity conditions are satisfied.

[7] We use here $\rho^{\mathrm{diag}}(S, \widehat{\sigma}_T) = -1$ as we are considering the typical case of negatively sloping equity smiles. We could equivalently have written $\rho^{\mathrm{diag}}(S, \widehat{\sigma}_T) = 1$.

It is easy to check that in case spot/spot correlations are not all equal, the correlation structure of the multi-asset local volatility model is still given by (11.31) with ρ_{SS}, $\chi_{S\sigma}$, $\chi_{\sigma\sigma}$ adorned with ij superscripts, as $\rho^{\rm cross}(S^i, \widehat{\sigma}^j_T) = -\rho^{ij}_{SS}$ and $\rho^{\rm cross}(\widehat{\sigma}^i_T, \widehat{\sigma}^j_{T'}) = \rho^{ij}_{SS}$.

We now test this mapping with the same parameter values used in Section 11.3.2; $\rho_{SS} = 60\%$. We use two-factor-model parameters in Table 11.3 to generate a vanilla smile. Using this as the component's vanilla smile, we calibrate the component's local volatility function and price the basket smile using $\rho_{SS} = 60\%$.

We then compare the resulting basket smile with that produced by our multi-asset two-factor model with $\chi_{\sigma\sigma} = \chi_{S\sigma} = \rho_{SS} = 60\%$. Results appear in Figure 11.7.

It is apparent that the mapping (11.31) is accurate for short maturities; for longer maturities implied volatilities in the stochastic volatility model are lower than in the local volatility model. This is not surprising: because volatilities – especially forward volatilities – are more volatile in the stochastic volatility model than in the local volatility model: the level of “effective” spot/spot correlation is lower than ρ_{SS}. This affects all multi-asset products, including the correlation swap which we now study.

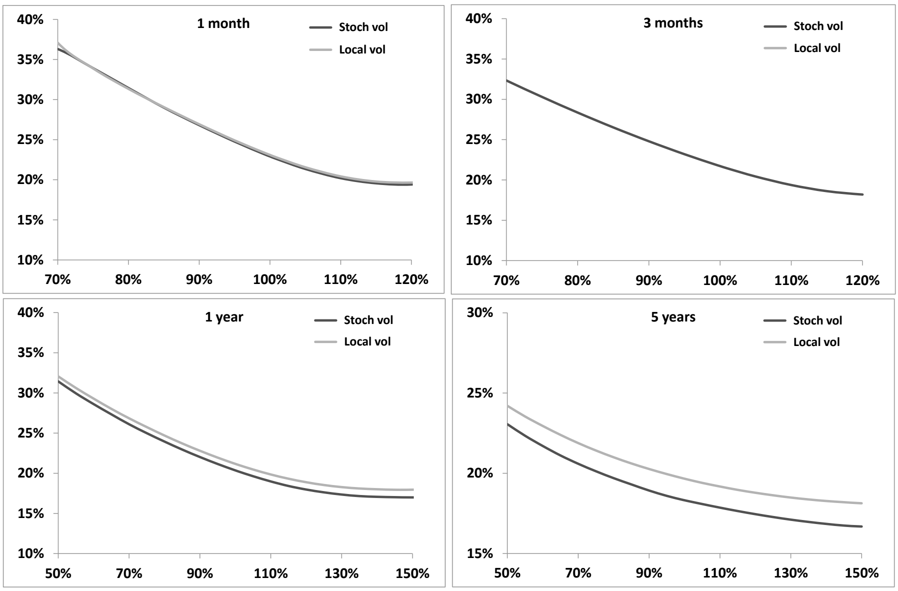

Figure 11.7: Comparison of basket smiles in local volatility and stochastic volatility models, for several maturities. The mapping of local to stochastic volatility is obtained with $\chi_{\sigma\sigma} = \chi_{S\sigma} = \rho_{SS} = 60\%$. The basket consists of $n = 10$ components, with identical smiles given by parameters in Table 11.3.

11.4 The correlation swap

Consider a basket of stocks or indexes. A correlation swap of maturity T pays at T the average pairwise realized correlation of the basket components minus a fixed strike $\widehat{\rho}$

$$\frac{1}{n(n-1)}\sum_{i\neq j}\rho_{ij} - \widehat{\rho} \tag{11.32}$$

where $\widehat{\rho}$ is set so that the swap's initial value vanishes. Correlation ρ_{ij} of S_i,S_j is defined with the standard realized correlation estimator, using daily log-returns:

$$\rho_{ij} = \frac{\Sigma r_k^i r_k^j}{\sqrt{\Sigma\, {r_k^i}^2}\sqrt{\Sigma\, {r_k^j}^2}} \tag{11.33}$$

where $r_k^i = \ln(S_k^i/S_{k-1}^i)$ and the sums runs from $k=1$ to $k=N$, where N is the number of returns used for estimating covariances and variances in (11.33).

n is the number of securities: 2 or 3 when the components are indexes and up to 50 for a correlation swap on the constituents of the Euro Stoxx 50 index. We have used in (11.32) the typical equal weighting.

Strike $\widehat{\rho}$ is also called the implied correlation of the swap. Indeed, in a constant volatility model, for daily returns, with all spot/spot correlations ρ_{SS} equal and a large number of returns – thus a long maturity – $\widehat{\rho} = \rho_{SS}$.

For shorter maturities, (11.33) is biased. Let us make the assumption of centered log-returns: $r_k^i = \sigma^i\sqrt{\Delta t}Z_k^i$, where σ^i is the (constant) volatility of asset i, Δt the interval between two spot observations – here one day, and the Z_k^i are iid standard normal random variables.

The bias and standard deviation of the correlation estimator are derived in Appendix A, at lowest order in $\frac{1}{N}$:

$$E[\rho_{ij}] = \rho_{SS}\left(1 - \frac{1-\rho_{SS}^2}{2N}\right) \tag{11.34}$$

$$\text{Stdev}\,(\rho_{ij}) = \frac{1}{\sqrt{N}}\left(1-\rho_{SS}^2\right) \tag{11.35}$$

where N is the number of returns in the historical sample. Typically the bias in $E[\rho_{ij}]$ is about one point of correlation for $\rho_{SS} = 60\%$ and $T = 1$ month.

Correlation swaps were introduced as a means of trading correlation and making implied correlation an observable parameter.

As a measure of correlation, however, averaging pairwise correlations results in a poorly defined estimator.

One would expect of an adequately defined estimator that its standard deviation vanishes either in the limit of a large sample size ($N \to \infty$) or as the number of basket constituents is increased ($n \to \infty$). This is not the case for the correlation swap estimator.

Making the same assumption of constant volatility centered log-returns as above, the bias and standard deviation of the correlation swap estimator in the limit $n \to \infty$ are given, at lowest order in $\frac{1}{N}$, by:

$$\begin{aligned}
\lim_{n\to\infty} E\left[\frac{1}{n(n-1)}\sum_{i\neq j}\rho_{ij}\right] &= \rho_{SS}\left(1-\frac{1-\rho_{SS}^2}{2N}\right) \\
\lim_{n\to\infty} \text{Stdev}\left(\frac{1}{n(n-1)}\sum_{i\neq j}\rho_{ij}\right) &= \frac{1}{\sqrt{N}}\left(1-\rho_{SS}^2\right)\frac{\sqrt{2}\,|\rho_{SS}|}{1+\rho_{SS}}
\end{aligned} \tag{11.36}$$

Observe that the right-hand side of (11.36) ($n = \infty$) is only marginally smaller than that of (11.35) ($n = 2$). For the sake of creating an estimator of average realized correlation, it would have been more judicious to average covariances and divide them by the average of variances – the standard deviation of the resulting correlation estimator tends to zero as the number of constituents is increased.

One would hope that the correlation swap is the perfect instrument for calibrating spot/spot correlation levels, but it is not so. Let us assume that our basket is homogeneous and let σ_t^i be the instantaneous (lognormal) volatility of S_t^i. In the limit of short returns, $\widehat{\rho}$ is given by:

$$\widehat{\rho}(T) = E[\rho_{ij}] = \rho_{SS}\, E\left[\frac{\int_0^T \sigma_t^i\sigma_t^j\,dt}{\sqrt{\int_0^T {\sigma_t^i}^2 dt}\,\sqrt{\int_0^T {\sigma_t^j}^2 dt}}\right] \tag{11.37}$$

where we explicitly keep track of the maturity dependence of $\widehat{\rho}$.

In case instantaneous volatilities σ_t^i and σ_t^j are collinear – $\sigma_t^j = \lambda\sigma_t^i$ – then $\widehat{\rho}$ is indeed equal to ρ_{SS}. In all other cases, $\widehat{\rho} \leq \rho_{SS}$, and the amount by which $\widehat{\rho}$ differs from ρ_{SS} depends on volatilities and correlations of instantaneous volatilities. Instantaneous volatilities are unobservable quantities – how does this manifest itself practically?

Risk-managing the correlation swap in the Black-Scholes model

Consider risk-managing a correlation swap on two underlyings S^1, S^2 with the Black-Scholes model. Since the dollar gammas and vegas of the correlation swap do not depend on S^1, S^2, it is natural to use variance swaps as hedge instruments.

Denote by $\sigma_{1t},\ \sigma_{2t},\ \rho_t$ the realized volatilities and correlation of log returns over $[0,\ t]$ and $\widehat{\sigma}_{1t}, \widehat{\sigma}_{2t}$ implied VS volatilities at time t for maturity T, which we are using as implied volatilities in our Black-Scholes model. Denote by $\widehat{\rho}_0$ the (implied) correlation level at which we are risk-managing the correlation swap, which we keep constant.

The value at t of the correlation swap in the Black-Scholes model is given by:[8]

$$P = \frac{t\rho_t\sigma_{1t}\sigma_{2t} + (T-t)\widehat{\rho}_0\widehat{\sigma}_{1t}\widehat{\sigma}_{2t}}{\sqrt{t\sigma_{1t}^2 + (T-t)\widehat{\sigma}_{1t}^2}\sqrt{t\sigma_{2t}^2 + (T-t)\widehat{\sigma}_{2t}^2}}$$

At $t = 0$, $P = \widehat{\rho}_0$ and $\frac{dP}{d(\widehat{\sigma}_{1t}^2)} = \frac{dP}{d(\widehat{\sigma}_{2t})} = 0$: the correlation swap has no sensitivity to implied volatilities at inception. We use variances rather than volatilities as state variables, since variance swaps provide perfect delta hedges with respect to variances. Let us now examine the correlation swap's sensitivities at later times.

Consider a short position in a correlation swap vega-hedged with variance swaps of maturity T. Let us calculate the $P\&L$ over $[t,\ t+\delta t]$ generated by variations $\delta(\widehat{\sigma}_{1t}^2),\ \delta(\widehat{\sigma}_{2t}^2)$ of implied VS variances, at order two. Computing second derivatives of P with respect to $\widehat{\sigma}_{1t}^2,\ \widehat{\sigma}_{2t}^2$ is straightforward.

Consider the particular situation when realized volatilities and correlation over $[0,\ t]$ match the implied values at time t: $\sigma_{1t} = \widehat{\sigma}_{1t},\ \sigma_{2t} = \widehat{\sigma}_{2t},\ \rho_t = \widehat{\rho}_0$ – this allows for a more compact expression or the P&L. The P&L at order two in $\widehat{\sigma}_{1t}^2,\ \widehat{\sigma}_{2t}^2$ – for a short position – reads:

$$\begin{aligned} P\&L &= \frac{\widehat{\rho}_0}{8}\frac{t(T-t)}{T^2}\left(\frac{\left(\delta(\widehat{\sigma}_{1t}^2)\right)^2}{(\widehat{\sigma}_{1t}^2)^2} + \frac{\left(\delta(\widehat{\sigma}_{2t}^2)\right)^2}{(\widehat{\sigma}_{2t}^2)^2} - 2\frac{\delta(\widehat{\sigma}_{1t}^2)\delta(\widehat{\sigma}_{2t}^2)}{\widehat{\sigma}_{1t}^2\widehat{\sigma}_{2t}^2}\right) \\ &= \frac{\widehat{\rho}_0}{8}\frac{t(T-t)}{T^2}\left(\frac{\delta(\widehat{\sigma}_{1t}^2)}{\widehat{\sigma}_{1t}^2} - \frac{\delta(\widehat{\sigma}_{2t}^2)}{\widehat{\sigma}_{2t}^2}\right)^2 \end{aligned} \tag{11.38}$$

Had we used different values for $\sigma_{1t},\ \sigma_{2t},\ \rho_t$, (11.38) would have been replaced with a more complicated quadratic form of $\delta(\widehat{\sigma}_{1t}^2),\ \delta(\widehat{\sigma}_{2t}^2)$, but the prefactor $\frac{t(T-t)}{T^2}$ would have remained.

(11.38) confirms – for the special case of realized volatilities and correlations at time t matching their implied values – that if $\frac{\delta\widehat{\sigma}_{1t}}{\widehat{\sigma}_{1t}} = \frac{\delta\widehat{\sigma}_{2t}}{\widehat{\sigma}_{2t}}$ no P&L is generated by the variation of VS volatilities.[9]

Correlation swaps are thus exotic volatility instruments that depend on the dynamics of forward variances. Practically, this manifests itself through volatility/volatility cross-gamma P&Ls.

11.4.1 Approximate formula in the two-factor model

We now compute $\widehat{\rho}(T)$ in (11.37) at order two in volatility of volatility in the two-factor model, for two underlyings with the same parameters, with a flat VS term

[8] We have made the assumption of very short returns, so that the contribution from the risk-neutral drift of $\ln S_t$ is negligible – this is adequate for daily returns of equities.

[9] In contrast with options on realized variance, there is no way that the vega hedge of the correlation swap may also function as gamma hedge since VSs on S^1 and S^2 generate gamma P&Ls proportional to $(\delta S^1)^2$ and $(\delta S^2)^2$ while the gamma P&L of the correlation swap also includes a $\delta S^1\delta S^2$ term. In addition to the P&L in (11.38), our P&L thus comprises a spreaded position of diagonal spot gammas against spot cross-gammas, accompanied by their respective thetas.

structure, in the limit of short returns. The instantaneous variance ξ_t^t is given by:

$$\xi_t^t = \xi_0^t e^{2\nu x_t^t - 2\nu^2 E\left[\left(x_t^t\right)^2\right]}$$

where x_t^T has been defined in (7.30), page 226. For $T = t$:

$$x_t^t = \alpha_\theta \left[(1-\theta)\, X_t^1 + \theta X_t^2\right] \tag{11.39}$$

and $\alpha_\theta = \left((1-\theta)^2 + \theta^2 + 2\rho_{X_1X_2}\theta(1-\theta)\right)^{-1/2}$. Remember ν is the volatility of a very short volatility. Expanding at order two in ν:

$$\xi_t^t = \xi_0^t \left(1 + 2\nu x_t^t + 2\nu^2 \left[(x_t^t)^2 - E[(x_t^t)^2]\right]\right) \tag{11.40}$$

In an expansion of $\widehat{\rho}(T)$ at order two in ν, the ν^2 term in (11.40) is multiplied by a constant: when evaluating its expectation its contribution vanishes. For the sake of calculating $\widehat{\rho}(T)$ at order two in ν, the expansion of ξ_t^t at order one is thus sufficient:

$$\xi_t^t = \xi_0^t \left(1 + 2\nu x_t^t\right)$$

We consider two underlyings each with a flat VS term structure and denote by x_t, y_t the x_t^t processes for, respectively, the first and second underlying. At second order in ν, $\widehat{\rho}(T)$ reads:

$$\widehat{\rho}(T) = \rho_{SS}\, E\left[\frac{\int_0^T \sqrt{1+2\nu x_t}\sqrt{1+2\nu y_t}\,dt}{\sqrt{\int_0^T (1+2\nu x_t)\,dt}\,\sqrt{\int_0^T (1+2\nu y_t)\,dt}}\right]$$

Working out the expansion at order two in ν we get:

$$\begin{aligned}
\widehat{\rho}(T) &= \rho_{SS} E\left[\frac{1+\nu\left(\overline{x}+\overline{y}\right) - \frac{\nu^2}{2}\overline{(x-y)^2}}{1+\nu\left(\overline{x}+\overline{y}\right) - \frac{\nu^2}{2}\left(\overline{x}-\overline{y}\right)^2}\right] \\
&= \rho_{SS} E\left[1 - \frac{\nu^2}{2}\left(\overline{(x-y)^2} - \overline{x-y}^2\right)\right]
\end{aligned}$$

which at order two in ν can be rewritten as:

$$\widehat{\rho}(T) = \rho_{SS}\, e^{-\frac{\nu^2}{2} E\left[\overline{(x-y)^2} - \overline{x-y}^2\right]} \tag{11.41}$$

where:

$$\begin{aligned}
\overline{x} &= \frac{1}{T}\int_0^T x_t dt \\
\overline{(x-y)^2} &= \frac{1}{T}\int_0^T (x_t - y_t)^2\, dt \\
\overline{x-y}^2 &= \left(\frac{1}{T}\int_0^T (x_t - y_t)\, dt\right)^2
\end{aligned}$$

Using the exponential in (11.41) ensures that for large values of ν, $\widehat{\rho}$ at most vanishes, but does not become negative. From the definitions above we have the pathwise inequality $\overline{(x-y)^2} \geq \overline{x-y}^2$ thus $E\left[\overline{(x-y)^2}\right] \geq E\left[\overline{x-y}^2\right]$.

(11.41) thus also ensures that the property $\widehat{\rho} \leq \rho_{SS}$ that expression (11.37) for $\widehat{\rho}(T)$ implies, is obeyed in our expansion at second order in ν. If forward variances for the two underlyings are collinear processes – $x_t = y_t$ – we recover that $\widehat{\rho} = \rho_{SS}$.

Carrying out the computation for the two-factor model with x_t^t given by (11.39) yields:

$$\widehat{\rho}(T) = \rho_{SS}\, e^{-(1-\chi_{\sigma\sigma})\left(\nu^2 T\right)\frac{1}{T}\int_0^T f(t)dt} \tag{11.42}$$

with the dimensionless $f(t)$ given by:

$$\begin{aligned} f(t) = \alpha_\theta^2 \Bigg[& (1-\theta)^2 \frac{1-e^{-2k_1 t}}{2k_1 T}\left(1 - 2\frac{1-e^{-k_1(T-t)}}{k_1 T}\right) \\ & + \theta^2 \frac{1-e^{-2k_2 t}}{2k_2 T}\left(1 - 2\frac{1-e^{-k_2(T-t)}}{k_2 T}\right) \\ & + 2\rho_{X_1X_2}\theta(1-\theta)\frac{1-e^{-(k_1+k_2)t}}{(k_1+k_2)T}\left(1 - \frac{1-e^{-k_1(T-t)}}{k_1 T} - \frac{1-e^{-k_2(T-t)}}{k_2 T}\right)\Bigg] \end{aligned}$$

11.4.2 Examples

We use parameter values in Table 11.3, page 435, for each of our two underlyings, which correspond to the smiles in Figure 11.2. We take $\rho_{SS} = 60\%$ and use three different values for $\chi_{\sigma\sigma}$: $40\%, 60\%, 80\%$.

$\widehat{\rho}(T)$ appears in Figure 11.8, both evaluated numerically in a Monte Carlo simulation, and as given by formula (11.42). The right-hand side of Figure 11.8 shows the same curves for $\chi_{\sigma\sigma} = 60\%$, together with the values generated by a local volatility model parametrized as in the previous section, that is calibrated on the smile generated by the two-factor model, and with $\rho_{SS} = 60\%$.

Observe first that the accuracy of the second-order expansion in volatility of volatility is satisfactory, even though it is systematically biased low, except for short maturities. For short maturities $\widehat{\rho}(T)$ in (11.42) surges above the exact value: this is due to the fact that we have carried out our expansion in the continuous-time version of the correlation swap.

Thus $\widehat{\rho}(T)$ in (11.42) is not subject to the bias in (11.34), which instead affects the actual Monte Carlo estimation – as it uses discrete daily returns – and is material for small values of N, that is short maturities.

The right-hand graph is a compelling illustration of how different a local volatility and a stochastic volatility model can be, even though they are calibrated on the same smile.

In the local volatility model, $\widehat{\rho}(T)$ hardly depends on T: this is likely due to the fact that forward variances – especially long-dated ones – have very little volatility.

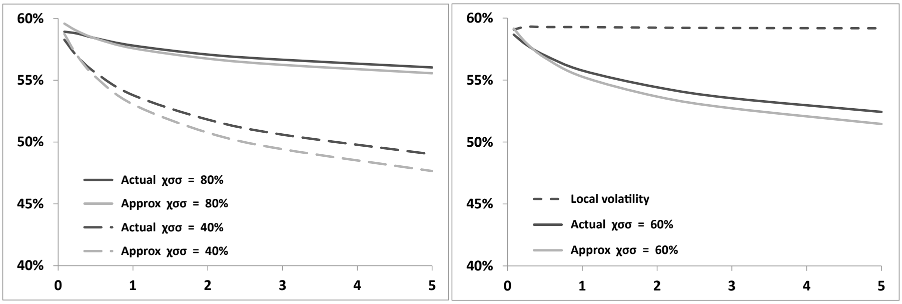

Figure 11.8: Left: $\widehat{\rho}(T)$ for maturities $T = 1$ month to 5 years, evaluated in a Monte Carlo simulation (Actual) and with formula (11.42) at second order in volatility of volatility (Approx), for $\chi_{\sigma\sigma} = 80\%$ and $\chi_{\sigma\sigma} = 40\%$. Right: same curves for $\chi_{\sigma\sigma} = 60\%$, along with $\widehat{\rho}(T)$ given by the local volatility model, calibrated on the smile generated by the two-factor model with parameters in Table 11.3.

Contrast this with the graphs in Figure 11.7. While for the basket smile the mapping $\chi_{\sigma\sigma} = \chi_{S\sigma} = \rho_{SS} = 60\%$ makes both models equivalent, they are not with respect to the correlation swap.

Figure 11.8 shows that the longer the maturity, the larger the impact of volatility of volatility, thus the lower $\widehat{\rho}(T)$. This deviates from market practice: typically the term structure of implied correlation swap correlations rises, rather than decreasing.

Finally, we have taken $\chi_{S\sigma} = 80\%$ in our numerical tests; using different values results, as expected, in no change of the actual and approximate values of $\widehat{\rho}(T)$.

11.5 Conclusion

- Life in the multi-asset local volatility framework used to be simple. Say we chose $\rho_{SS} = 60\%$ as correlation for the driving Brownian motions, priced multi-asset options and backed out the implied correlation in a multi-asset Black-Scholes model. Depending on the payoff at hand – say, an ATM basket call option, a VS on the basket, a correlation swap, a forward on the worst-performing asset – we would get different numbers, but clustered within a small range below 60%.

- This is not so with multi-asset stochastic volatility: exotic payoffs are sensitive to spot/volatility and volatility/volatility correlation levels, which, unlike in the local volatility model, can now be set separately. The richness of correlation risks manifests itself in the fact that implied spot/spot correlations backed out in a multi-asset Black-Scholes model for different payoffs will now vary widely.

- In our simple parametrization, in addition to ρ_{SS}, we introduce two quantities: $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$. By choosing $\chi_{\sigma\sigma} = \chi_{S\sigma} = \rho_{SS}$ we are able to mimic the local volatility model, for basket European payoffs, for short maturities.

- Do $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ have implied counterparts? We have shown that only a small fraction of the basket skew is contributed by the components' skew; the bulk of it is determined by $\chi_{S\sigma}$. Thus, the ratio of basket-to-component skew is a measure of the implied value of $\chi_{S\sigma}$, once ρ_{SS} has been set. Similarly $\chi_{\sigma\sigma}$ determines the strike of the correlation swap, again once ρ_{SS} has been set.

It would be useful to have good approximations for the basket VS and ATMF volatilities. It doesn't help that baskets are typically of the arithmetic – rather than geometric – type: the weights of each asset's return in the basket's return are not constant. Thus even simple objects such as basket VS and ATMF volatilities depend in a complicated way on $\rho_{SS}, \chi_{S\sigma}, \chi_{\sigma\sigma}$, let alone payoffs on worst-ofs.

Appendix A – bias/standard deviation of the correlation estimator

We present here the derivation of formulas (11.34) and (11.35), for the case of Gaussian-distributed returns with constant volatility. Since volatility is constant, we can normalize returns by their standard deviation and evaluate correlations with normalized returns. Consider the i-th and j-th underlyings and denote by x_i^τ, x_j^τ their τ-th daily return. x_i^τ, x_j^τ are assumed to be iid standard normal random variables with $\left\langle x_i^\tau x_j^\tau \right\rangle = \rho$ if $i \neq j$ and $\langle x_i^\tau x_i^\tau \rangle = 1$ where ρ is the correlation between the two underlyings.

The correlation estimator for the (i,j) couple is:

$$\widehat{\rho}_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}$$

with

$$C_{ij} = x_i . x_j \equiv \frac{1}{N} \Sigma_\tau x_i^\tau x_j^\tau$$

where N is the number of returns in our historical sample. C_{ij} is the average of independent random variables $x_i^\tau x_j^\tau$. As $N \to \infty$, fluctuations around its mean $\langle C_{ij} \rangle = \rho$ tend to zero, so let us write

$$\begin{aligned} C_{i \neq j} &= \rho + \varepsilon_{ij} \quad \varepsilon_{ij} = x_i . x_j - \rho \\ C_{ii} &= 1 + \varepsilon_{ii} \quad \varepsilon_{ij} = x_i . x_i - 1 \end{aligned}$$

We have:

$$\widehat{\rho}_{ij} \;=\; \frac{\rho + \varepsilon_{ij}}{\sqrt{1+\varepsilon_{ii}}\sqrt{1+\varepsilon_{jj}}}$$

We now expand $\widehat{\rho}_{ij}$ in powers of fluctuations ε_{ij}. By construction the ε_{ij} are centered: $\langle \varepsilon_{ij} \rangle = 0$. The first non-trivial contribution to the bias of $\widehat{\rho}_{ij}$ is thus generated by second-order terms. At second order in the ε_{ij} we have:

$$\widehat{\rho}_{ij} \;=\; \rho \;-\; \left(\rho\frac{\varepsilon_{ii}+\varepsilon_{jj}}{2} - \varepsilon_{ij}\right) \;-\; \left(\varepsilon_{ij}\frac{\varepsilon_{ii}+\varepsilon_{jj}}{2} - \rho\Big(\frac{3}{8}(\varepsilon_{ii}^2+\varepsilon_{jj}^2) + \frac{1}{4}\varepsilon_{ii}\varepsilon_{jj}\Big)\right) \tag{11.43}$$

At this order, the expectation of the correlation estimator is given by:

$$\langle \widehat{\rho}_{ij} \rangle \;=\; \rho \;-\; \left\langle \varepsilon_{ij}\frac{\varepsilon_{ii}+\varepsilon_{jj}}{2} - \rho\Big(\frac{3}{8}(\varepsilon_{ii}^2+\varepsilon_{jj}^2) + \frac{1}{4}\varepsilon_{ii}\varepsilon_{jj}\Big)\right\rangle$$

This formula involves moments of Gaussian variables x_i^τ, x_j^τ. The Wick theorem yields the following expressions for 4th order moments:

$$\begin{array}{llll}
\left\langle \varepsilon_{ii}^2 \right\rangle = \frac{2}{N} & \left\langle \varepsilon_{ii}\varepsilon_{jj} \right\rangle = \frac{2\rho^2}{N} & \left\langle \varepsilon_{ii}\varepsilon_{ij} \right\rangle = \frac{2\rho}{N} & \left\langle \varepsilon_{ij}^2 \right\rangle = \frac{1+\rho^2}{N} \\
\left\langle \varepsilon_{ij}\varepsilon_{kk} \right\rangle = \frac{2\rho^2}{N} & \left\langle \varepsilon_{ij}\varepsilon_{kl} \right\rangle = \frac{2\rho^2}{N} & \left\langle \varepsilon_{ij}\varepsilon_{il} \right\rangle = \frac{\rho+\rho^2}{N} &
\end{array}$$

where indices i, j, k, l are all different.

We then get, from (11.43):

$$\begin{aligned}
\langle \widehat{\rho}_{ij} \rangle \;&=\; \rho - \left(\frac{2\rho}{N} - \rho\Big(\frac{3}{8}\frac{4}{N} + \frac{1}{4}\frac{2\rho^2}{N}\Big)\right) \\
&=\; \rho\Big(1 - \frac{1-\rho^2}{2N}\Big)
\end{aligned}$$

which is the result in (11.34).

We now turn to the variance of $\widehat{\rho}_{ij}$. The first non-trivial contribution is generated by the square of the order-one correction to ρ in (11.43), and is of order $\frac{1}{N}$. For the sake of calculating the variance of $\widehat{\rho}_{ij}$ we can then ignore the bias just derived, which generates a contribution of order $\frac{1}{N^2}$. Using the expressions of 4th order moments above, we have:

$$\left\langle \left(\rho\frac{\varepsilon_{ii}+\varepsilon_{jj}}{2} - \varepsilon_{ij}\right)^2 \right\rangle \;=\; \frac{1}{N}\left(1-\rho^2\right)^2$$

from which (11.35) follows.

Consider now the correlation estimator used in the definition of the correlation swap, namely the average of all pairwise correlation estimators:

$$\widehat{\rho} \;=\; \frac{1}{n(n-1)}\sum_{i\neq j}\widehat{\rho}_{ij}$$

Assuming all pairwise correlations are equal to ρ, at order one in the fluctuations ε_{ij}:

$$\widehat{\rho} = \rho - \frac{1}{n(n-1)} \sum_{i \neq j} \left(\rho \frac{\varepsilon_{ii} + \varepsilon_{jj}}{2} - \varepsilon_{ij} \right) \tag{11.44}$$

Calculating now the expectation of the square of the right-hand side of (11.44) and taking the limit $n \to \infty$ yields expression (11.36) for the standard deviation of $\widehat{\rho}$.

Chapter's digest

11.1 The short ATMF basket skew

▶ We consider the case of an equally weighted basket of n components. Assuming identical correlations ρ_{SS} among components, the ATM implied volatility and skew of the basket are given, for short maturities, by:

$$\begin{aligned}\widehat{\sigma}_B &= \sqrt{\frac{1+(n-1)\rho_{SS}}{n}}\,\widehat{\sigma}\\ \mathcal{S}_B &= \frac{1+(n-1)\rho_{SS}}{n}\frac{\widehat{\sigma}^3}{\widehat{\sigma}_B^3}\left[\frac{\mathcal{S}}{n}+\frac{n-1}{n}\frac{1}{2\widehat{\sigma}^2}\frac{\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}}}{dt}\right]\end{aligned}$$

where $\widehat{\sigma}$ and $\mathcal{S}$ are the components' ATMF volatility and skew, assumed to be identical for all components, and $\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}}$ is the cross spot/volatility covariance.

Only a fraction $\frac{1}{n}$ of the basket ATMF skew is contributed by the ATMF skew of the components. The bulk of the basket skew is generated by cross spot/volatility correlations.

▶ Specializing to the case of a large homogeneous basket:

$$\begin{aligned}\widehat{\sigma}_B &\simeq \sqrt{\rho_{SS}}\,\widehat{\sigma}\\ \mathcal{S}_B &\simeq \frac{1}{\sqrt{\rho_{SS}}}\frac{1}{2\widehat{\sigma}^2}\frac{\langle d\ln S\, d\widehat{\sigma}\rangle_{\text{cross}}}{dt}\end{aligned}$$

▶ For vanishing cross spot/volatility correlations, the basket skew vanishes.

▶ If cross spot/volatility covariances are identical to their diagonal counterparts, $\mathcal{S}_B \simeq \frac{1}{\sqrt{\rho_{SS}}}\mathcal{S}$: the basket skew is larger than the component's.

▶ In the local volatility model:

$$\begin{aligned}\widehat{\sigma}_B &\simeq \sqrt{\rho_{SS}}\,\widehat{\sigma}\\ \mathcal{S}_B &\simeq \sqrt{\rho_{SS}}\mathcal{S}\end{aligned}$$

The basket skew is smaller than the component's skew.

❧ ❧ ❧ ❧ ❧

11.2 Parametrizing multi-asset stochastic volatility models

▶ We wish to parametrize a multi-asset stochastic volatility model (a) with few additional parameters, (b) in a manner that does not depend on the model's factor structure. We introduce dimensionless numbers $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ to specify the cross spot/volatility and volatility/volatility covariance functions, in terms of their diagonal counterparts:

$$\begin{aligned}\rho^{\text{cross}}(S, \widehat{\sigma}_T) &= \chi_{S\sigma}\, \rho^{\text{diag}}(S, \widehat{\sigma}_T)\\ \rho^{\text{cross}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'}) &= \chi_{\sigma\sigma}\, \rho^{\text{diag}}(\widehat{\sigma}_T, \widehat{\sigma}_{T'})\end{aligned}$$

For a homogeneous basket, the conditions on $\chi_{S\sigma}, \chi_{\sigma\sigma}$ so that the global correlation matrix is positive are easily expressed. While $\chi_{\sigma\sigma} \leq 1$, $\chi_{S\sigma}$ can go above 1.

▶ Realized values of $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ can be measured using historical data for spot and implied volatilities.

❧❧❧❧❧

11.3 The ATMF basket skew

▶ At order one in volatility of volatility, for a homogeneous basket, the ATMF skew is given by:

$$\mathcal{S}_T^B = \mathcal{S}_T \sqrt{\frac{n}{1+(n-1)\rho_{SS}}} \left[\frac{1+(n-1)\chi_{S\sigma}}{n}\right]$$

where ρ_{SS} is the correlation among basket components and $\mathcal{S}_T$ their ATMF skew.

▶ Numerical experiment show that this formula for the basket skew is accurate, and that the latter hardly depends on $\chi_{\sigma\sigma}$.

▶ The approximate expression for the basket VS volatility derived, as is the case for $\mathcal{S}_T^B$, with the assumption of frozen weights, is, on the other hand, not accurate. In reality it does depend on $\chi_{S\sigma}$.

▶ The multi-asset local volatility model is mimicked by setting $\chi_{S\sigma}$ and $\chi_{\sigma\sigma}$ equal to the spot/spot correlation.

In practice, while the shape of the resulting smile indeed matches that of the multi-asset local volatility model, the level of basket ATMF volatility is shifted downwards, especially for long maturities.

❧❧❧❧❧

11.4 The correlation swap

▶ The correlation swap is an exotic volatility instrument. It is very sensitive to the correlation of the volatility processes of the basket's components, thus to parameter $\chi_{\sigma\sigma}$.

At lowest order in volatility of volatility, the fair strike of the correlation swap, for the case of a homogeneous pair of underlyings, is given by:

$$\widehat{\rho}(T) = \rho_{SS}\, e^{-(1-\chi_{\sigma\sigma})(\nu^2 T)\frac{1}{T}\int_0^T f(t)dt}$$

where $f(t)$ is a simple function that involves model parameters.

► Numerical experiments show that the accuracy of our approximate formula is acceptable. For long maturities, the correlation swap's fair strike lies much lower than the correlation of spot processes. This contrasts with the behavior in the local volatility model, which yields a fair strike almost equal to the correlation of spot processes.

This page intentionally left blank