# Chapter 6

# An example of one-factor dynamics: the Heston model

In the next chapter we cover models built on a specification of the dynamics of forward VS variances that can be calibrated to a term structure of VS volatilities, or a term-structure of vanilla implied volatilities of an arbitrary moneyness.

We now briefly pause to consider the Heston model, which is not a forward variance model, since in its native form it can be calibrated to the VS volatility of one single maturity.

It is instructive to assess the Heston model and its capabilities in the framework of forward variances. This exercise will help us introduce suitability criteria which we then apply to the forward variance models covered in Chapter 7. Moreover, the Heston model is an archetypal example among first-generation stochastic volatility models, that is models written in terms of the instantaneous variance $V_t$.

Unless stated otherwise, forward VS variances $\xi_t^T$ will henceforth be called simply "forward variances" and we use $\widehat{\sigma}_T$ interchangeably for VS or log-contract implied volatilities.

## 6.1 The Heston model

The Heston model [60] is a first-generation model; it owes its popularity to the fact that, being an affine model, the Laplace transform of its moment-generating function for $\ln S$ is analytically known. Numerical inversion of this transform then yields vanilla option prices – when there are no dividends in fixed cash amounts.

The analytics of the Heston model is abundantly covered in the literature – see for example [48], [69]. Rather, we concentrate on the joint spot/volatility dynamics that this model generates. We assume for simplicity zero interest rate and repo.

The Heston model is a diffusive one-factor model – instead of forward variances, the instantaneous variance is modeled, according to the following SDEs:

$$\begin{cases} dS_t &= \sqrt{V_t} S_t dW_t \\ dV_t &= -k(V_t - V^0)dt + \sigma\sqrt{V_t} dZ_t \end{cases} \tag{6.1}$$

$V_t$ is the instantaneous variance – with respect to our previous notation:

$$V_t \;=\; \xi_t^t \;=\; \overline{\sigma}_t^2$$

$\sigma$ is commonly called "volatility of volatility", though it is not a lognormal volatility as it has dimension time$^{-1}$. The Brownian motions $W_t, Z_t$ are correlated, with correlation $\rho$. $V^0$ is a constant.

---

## 6.2 Forward variances in the Heston model

Forward variances $\xi_t^T$ are defined by:

$$\xi_t^T \;=\; E_t[V_T]$$

Taking the expectation of the second equation in (6.1) and using the compact notation $\overline{V}_u \;=\; E_t[V_u]$ yields:

$$d\overline{V}_T = -k(\overline{V}_T - V^0)dT \tag{6.2}$$

whose solution is

$$\overline{V}_T \;= V^0 \;+\; e^{-k(T-t)}(V_t - V^0)$$

which, using our notation reads:

$$\xi_t^T \;=\; V^0 \;+\; e^{-k(T-t)}\left(\xi_t^t - V^0\right) \tag{6.3}$$

VS volatilities $\widehat{\sigma}_t^T$ are given by:

$$\widehat{\sigma}_T^2\,(t) \;=\; \frac{1}{T-t}\int_t^T \xi_t^\tau d\tau \;=\; V^0 + \frac{1-e^{-k(T-t)}}{k\,(T-t)}(V_t - V^0) \tag{6.4}$$

Differentiating (6.3) gives:

$$d\xi_t^T \;=\; \sigma e^{-k(T-t)}\sqrt{\xi_t^t}\,dZ_t$$

$\xi_t^T$ is driftless – as it should. In the framework of forward variances the Heston dynamics (6.1) reads:

$$\begin{cases} dS_t \;=\; \sqrt{\xi_t^t}S_t dW_t \\ d\xi_t^T \;=\; \sigma e^{-k(T-t)}\sqrt{\xi_t^t}dZ_t \end{cases} \tag{6.5}$$

The Heston model is thus a one-factor model for forward variances where the instantaneous volatility of all forward variances $\xi_t^T$ is proportional to the instantaneous volatility $\overline{\sigma}_t = \sqrt{\xi_t^t}$. It is a Markov-functional model for forward variances, as $\xi_t^T$ is a *function* of $\xi_t^t$, given by (6.3).

Still, the Heston model should not be considered as a particular version of the forward variance models of Chapter 7; a one-dimensional Markov representation exists *only* if the initial values $\xi_{t=0}^T$ of forward variances satisfy condition (6.2):

$$\frac{d\xi_0^T}{dT} = -k\left(\xi_0^T - V^0\right)dT$$

It is not able to accommodate general term structure of VS volatilities.

Before analyzing further the dynamics of the Heston model, let us discuss the issue of the drift of $V_t$ in first-generation models.

---

## 6.3 Drift of $V_t$ in first-generation stochastic volatility models

The traditional approach to these models typically found in papers and textbooks can be summarized as follows:

- Start with historical dynamics of the instantaneous variance:

  $$dV_t = \mu\left(t, S, V, p\right)dt + \alpha dZ_t$$

  where $p$ are model parameters – such as $k, V^0$ in the Heston model.

- In risk-neutral dynamics, drift of $V_t$ is altered by “market price of risk” $\lambda$, which is an arbitrary function of $t, S, V$:

  $$dV_t = \left(\mu\left(t, S, V, p\right) + \lambda\left(t, S, V\right)\right)dt + \alpha dZ_t$$

- A few lines down the road, jettison “market price of risk” and conveniently decide that risk-neutral drift has same functional form as historical drift, except parameters now have stars:

  $$dV_t = \mu\left(t, S, V, p^\star\right)dt + \alpha dZ_t$$

- Calibrate (starred) parameters on vanilla smile.

Discussions surrounding the “market price of risk” and the uneasiness generated by its *a priori* arbitrary form and hasty disposal are pointless – the “market price of risk” is a nonentity.

$V_t$ itself is an artificial object: for different $t$, $V_t$ represents a different forward variance. Its drift is then only a reflection of the term-structure of forward variances. Differentiating the identity $V_t = \xi_t^t$ and using the following dynamics for $\xi_t^T$:

$$d\xi_t^T = \lambda_t^T dZ_t^T$$

yields

$$dV_t = \left.\frac{d\xi_t^T}{dT}\right|_{T=t} dt + \lambda_t^t dZ_t^t$$

The drift of the instantaneous variance is thus simply the slope at time $t$ of the short end of the variance curve. One can check that taking the derivative at $\xi_t^T$ with respect to $T$ in (6.3) indeed yields the drift of $V_t$ in the Heston model.

---

## 6.4 Term structure of volatilities of volatilities in the Heston model

Using (6.4) we get the following SDE for $\widehat{\sigma}_T$:

$$d\widehat{\sigma}_T = \bullet\, dt + \frac{\sigma}{2}\frac{1-e^{-k(T-t)}}{k\,(T-t)}\frac{\widehat{\sigma}_t}{\widehat{\sigma}_T} dZ_t \tag{6.6}$$

where $\widehat{\sigma}_t = \sqrt{V_t}$. Let us examine the term structure of the volatility of VS volatilities, that is the dependence of the volatility of $\widehat{\sigma}_T$ on $T$. We have the two following limiting regimes:

$$T-t \ll \frac{1}{k} \qquad d\widehat{\sigma}_T \simeq \bullet\, dt + \frac{\sigma}{2}\left(1-\frac{k\,(T-t)}{2}\right)\frac{\widehat{\sigma}_t}{\widehat{\sigma}_T} dZ_t \tag{6.7}$$

$$T-t \gg \frac{1}{k} \qquad d\widehat{\sigma}_T \simeq \bullet\, dt + \frac{\sigma}{2}\frac{1}{k\,(T-t)}\frac{\widehat{\sigma}_t}{\widehat{\sigma}_T} dZ_t \tag{6.8}$$

Thus, for long maturities, the instantaneous volatility of $\widehat{\sigma}_T$ decays like $\frac{1}{T-t}$. For a flat term structure of VS volatilities, (6.6) implies that the instantaneous lognormal volatility of $\widehat{\sigma}_T$ is:

$$\mathrm{vol}(\widehat{\sigma}_T) \propto \frac{1-e^{-k(T-t)}}{k\,(T-t)} \tag{6.9}$$

Figure 6.1 shows the lognormal volatilities of VS implied volatilities of the Euro Stoxx 50 index, compared to expression (6.9), suitably rescaled, along with a power-law fit.

As Figure 6.1 shows, volatilities of VS volatilities are typically larger than volatilities of the underlying itself. This is not due to the fact that we are using VS volatilities – volatilities of ATM volatilities are similar.

As is apparent, while realized levels of volatilities of VS volatilities are consistent with a power law dependence on maturity, they cannot be captured by the dynamics of the Heston model over a wide range of maturities. The two values of $k$ used in Figure 6.1, 2 and $0.4$, have been chosen so as to best match the short and long end of the term structure of the volatilities of VS volatilities in the graph.

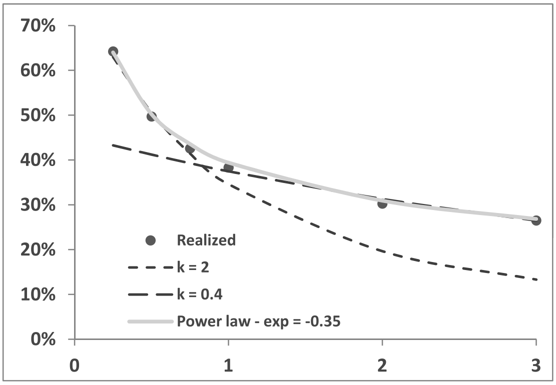


**Figure 6.1**: Volatility of VS volatilities of the Euro Stoxx 50 index as a function of maturity (years), evaluated on the period [2005, 2010] (dots), along with (a) volatilities of VS volatilities in the Heston model given by (6.9) for two different values of $k$ (dotted lines), (b) a power-law fit $\propto T^{-0.35}$.

---

## 6.5 Smile of volatility of volatility

In diffusive models, for short maturities, the ATMF implied volatility is approximately equal to the instantaneous volatility, which is equal to $\sqrt{V}$. From (6.1) we get:

$$d\sqrt{V} \;=\; \bullet\, dt + \frac{\sigma}{2} dZ_t$$

Thus, in the Heston model, short ATMF implied volatilities are approximately normal, rather than lognormal, with a normal volatility equal to $\frac{\sigma}{2}$. Figure 6.2 shows the 3-month ATMF implied volatility as well as its realized (lognormal) volatility computed over a 6-month sliding window.

Inspection of the scales of the left-hand and right-hand axes again confirms that volatilities of short-dated volatilities are larger than volatilities themselves. Also, Figure 6.2 shows that high levels of volatility tend to coincide with high levels of volatility of volatility. In this respect, implied volatilities are in fact more than lognormal: their dynamics seems to be of the type:

$$d\widehat{\sigma}_{\mathrm{ATM}} \;=\; \bullet\, dt \;+ \widehat{\sigma}_{\mathrm{ATM}}^{\gamma} dZ_t$$

with $\gamma > 1$. This should be compared with the value $\gamma = 0$ that the Heston model generates for short maturities.

What about implied volatilities for longer maturities? Equation (6.8) shows that the instantaneous volatility of long-dated VS volatilities decays like $\frac{1}{T-t}$. Equation

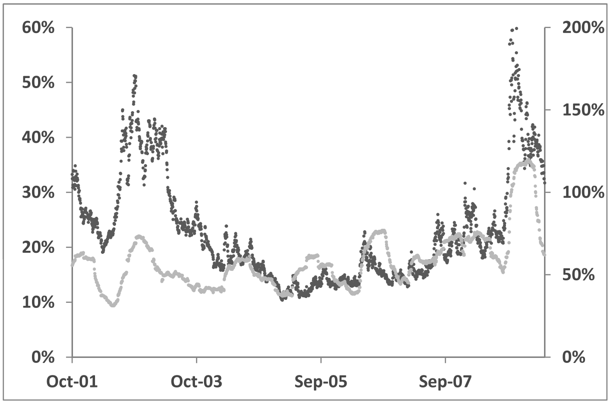


**Figure 6.2**: The 3-month ATMF implied volatility of the Euro Stoxx 50 index (darker dots, left-hand axis) together with its (lognormal) volatility, evaluated with a six-month sliding window (lighter dots, right-hand axis) from October 2001 to May 2009.

(6.4) implies that, since $V_t$ is positive, $\widehat{\sigma}_T\left(t\right)$ has a floor:

$$\widehat{\sigma}_T\left(t\right) \;\geq\; \widehat{\sigma}_T^{\min}\left(t\right) = \;\sqrt{V^0}\sqrt{1-\frac{1-e^{-k(T-t)}}{k\left(T-t\right)}}$$

Figure 6.3 shows $\frac{\widehat{\sigma}_T^{\min}(t)}{\sqrt{V^0}}$ as a function of $T-t$; note that, from (6.4), $\sqrt{V^0}$ is the level of VS volatility for long-dated maturities, calibrated at $t=0$. Forward VS volatilities are thus floored at a fraction of the initial long-run VS volatility level: The consequence for the smile of volatility of volatility is that volatilities of VS volatilities vanish as VS volatilities come near the floor. As is apparent in Figure 6.3, this floor on $\widehat{\sigma}_T\left(t\right)$ is not a minor effect.

## 6.6 ATMF skew in the Heston model

We now turn our attention to the dependence of the ATMF skew on maturity and volatility level. We obtain an approximation of the ATMF skew at order one in the volatility of volatility $\sigma$.

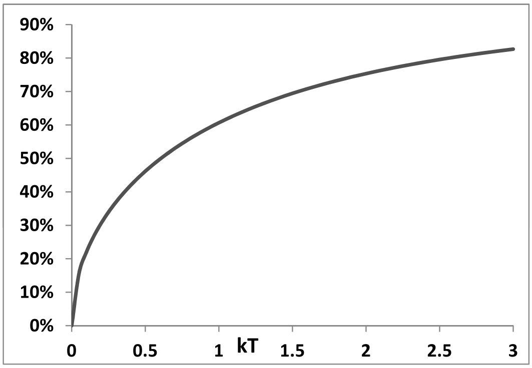


Figure 6.3: $\frac{\widehat{\sigma}_T^{\min}(t)}{\sqrt{V^0}}$ as a function of $k(T-t)$.

### 6.6.1 The smile at order one in volatility of volatility

While we provide in Chapter 8 a general expression of the ATMF skew at order one in the volatility of volatility for general stochastic volatility models, we now carry out a derivation for the specific case of the Heston model.

The Heston model is homogeneous, thus vanilla implied volatilities are not a separate function of spot and strike, but of the ratio of the strike to the forward of the option's maturity. To reduce bookkeeping we now set interest rate and repo to zero and reinstate them once we have our final formulas.

From SDEs (6.1) we get the following equation for the price $P(t,S,V)$ of a European option:

$$\frac{dP}{dt} - k\left(V - V^0\right)\frac{dP}{dV} + \frac{V}{2}S^2\frac{d^2P}{dS^2} + \frac{\sigma^2}{2}V\frac{d^2P}{dV^2} + \rho\sigma VS\frac{d^2P}{dSdV} = 0 \quad (6.10)$$

with the terminal condition at maturity $P(T,S,V) = f(S)$, where $f$ is the option's payoff. Denote by $P^0(t,S,V)$ the solution of (6.10) with $\sigma = 0$:

$$\frac{dP^0}{dt} - k\left(V - V^0\right)\frac{dP^0}{dV} + \frac{V}{2}S^2\frac{d^2P^0}{dS^2} = 0 \quad (6.11)$$

with terminal condition $P^0(T,S,V) = f(S)$, where $f$ is the option's payoff.

To gain some intuition on the solution of (6.11) let us go back to the stochastic representation (6.1). For $\sigma = 0$, $V$ is deterministic: the SDE for $V_t$ in (6.1) becomes an ODE, identical to (6.2). Given the value $V_t$ at time $t$, the value of $V_\tau$ at a later time $\tau$ is given by

$$V_\tau(V_t) = V^0 + (V_t - V^0)e^{-k(\tau-t)} \quad (6.12)$$

For $\sigma = 0$, the Heston model becomes a lognormal model with deterministic time-dependent volatility $\sigma(\tau)$ given by:

$$\sigma^2(\tau) = V_\tau(V_t)$$

The implied volatility at time $t$ for maturity $T$ is given by:

$$\widehat{\sigma}_T^2(t,V) = \frac{1}{T-t}\int_t^T V_\tau(V)\,d\tau = V^0 + (V-V^0)\frac{1-e^{-k(T-t)}}{k(T-t)} \tag{6.13}$$

where $V$ is the instantaneous variance at time $t$. $P^0$ is thus simply given by the Black-Scholes formula evaluated with implied volatility $\widehat{\sigma}_T(t,V)$:

$$P^0(t,S,V) = P_{BS}(t,S,\widehat{\sigma}_T(t,V))$$

Let us expand $P$:

$$P = P^0 + \delta P$$

where $\delta P$ is of order one in $\sigma$ and let us insert this expression in equation (6.10), keeping terms up to order one in $\sigma$. Using the fact that $P^0$ obeys (6.11) we are left with the following equation for $\delta P$:

$$\frac{d\delta P}{dt} - k(V-V^0)\frac{d\delta P}{dV} + \frac{V}{2}S^2\frac{d^2\delta P}{dS^2} = -\rho\sigma VS\frac{d^2P^0}{dSdV} \tag{6.14}$$

with the terminal condition $\delta P(T,S,V)=0$. $\delta P$ is generated by the source term in the right-hand side. The second derivative of $P_0$ with respect to $V$ does not appear, as it is multiplied by $\sigma^2$ – at order 1 in $\sigma$ only the mixed derivative of $P_0$ with respect to $S,V$ contributes.

The solution of (6.14) at time 0 is given by:

$$\delta P(t,S,V) = \underset{\substack{S_t=S\\ V_t=V}}{E_t^0}\left[\int_t^T \rho\sigma V_\tau S_\tau \left.\frac{d^2P^0}{dSdV}\right|_{\tau,S_\tau,V_\tau} d\tau\right]$$

where the subscript 0 indicates that the expectation is taken with respect to the dynamics (6.1) with $\sigma = 0$, that is in a Black-Scholes model with deterministic time-dependent volatility $\sigma(\tau)$.

In the Black-Scholes model the vega of a European option is related to its dollar gamma – see (5.66), page 181:

$$\frac{dP_{BS}}{d\widehat{\sigma}} = S^2\frac{d^2P_{BS}}{dS^2}\widehat{\sigma}(T-\tau)$$

We then have:

$$\begin{aligned}\frac{dP_0}{dV} &= S^2\frac{d^2P^0}{dS^2}\widehat{\sigma}_T(\tau,V)(T-\tau)\,\frac{d\widehat{\sigma}_T(\tau,V)}{dV}\\ &= \frac{1-e^{-k(T-\tau)}}{2k}S^2\frac{d^2P^0}{dS^2}\end{aligned}$$

where we have used expression (6.13) for $\widehat{\sigma}_T(\tau,V)$. $\delta P$ now reads:

$$\delta P(t,S,V) = \underset{\substack{S_t=S\\ V_t=V}}{E_t^0}\left[\int_t^T \frac{\rho\sigma}{2}V_\tau(V_t)\,\frac{1-e^{-k(T-\tau)}}{k}\,S_\tau\frac{d}{dS}S^2\left.\frac{d^2P^0}{dS^2}\right|_{\tau,S_\tau,V_\tau} d\tau\right]$$

where $V_\tau$ is given by expression (6.12) as a function of $V_t$, the instantaneous variance at time 0.

We now take $t = 0$ and simply denote by $V$ the instantaneous variance at $t = 0$. We have shown in Appendix A of Chapter 5 that, in the Black-Scholes model with deterministic time-dependent volatility, $\mathrm{e}^{-r(\tau-t)} \frac{d^n P^0}{d\ln S^n}(\tau, S_\tau)$ is a martingale. Using:

$$S\frac{d}{dS}S^2\frac{d^2}{dS^2} = \frac{d^3}{d\ln S^3} - \frac{d^2}{d\ln S^2}$$

we get:

$$\delta P = \frac{\rho\sigma}{2}\left[\int_0^T V_\tau\left(V\right)\frac{1-e^{-k(T-\tau)}}{k}\,d\tau\right]\left(\frac{d^3P^0}{d\ln S^3} - \frac{d^2P^0}{d\ln S^2}\right)_{t=0,S,V} \tag{6.15}$$

where $V_\tau\left(V\right)$ is given by:

$$V_\tau\left(V\right) = V^0 + (V - V^0)e^{-k\tau}$$

Compare (6.15) with expression (5.88), page 193: $\delta P$ has the same expression as a function of the second and third order derivatives of $P^0$ with respect to $\ln S$: perturbing at order one in the volatility of volatility amounts to perturbing at order one in the third-order cumulant with fixed forward variances.[1]

This is not surprising. Indeed, the ODE for $E[V_t]$ in the Heston model – see equation (6.2) – does not involve $\sigma$. Consequently, the perturbation in powers of $\sigma$ leaves forward variances unchanged at all orders.

The interpretation of the integral over $\tau$ in the prefactor is not straightforward at this stage. It will become clear when we carry out the derivation for general stochastic volatility models – see Section 8.6.

The expansion of the implied volatility $\widehat{\sigma}_{KT}$ at order one in $\sigma$ is: $\widehat{\sigma}_{KT} = \widehat{\sigma}_T\left(0,V\right) + \delta\widehat{\sigma}_{KT}$ where $\delta\widehat{\sigma}_{KT}$ is given by:

$$\begin{aligned}
\delta\widehat{\sigma}_{KT} &= \left(\frac{dP_{BS}}{d\widehat{\sigma}}\right)^{-1}\delta P \\
&= \frac{1}{\widehat{\sigma}_T\left(0,V\right)T}\left[\int_0^T \frac{\rho\sigma}{2}V_\tau\left(V\right)\frac{1-e^{-k(T-\tau)}}{k}\,d\tau\right]\frac{\frac{d^3P^0}{d\ln S^3} - \frac{d^2P^0}{d\ln S^2}}{\frac{d^2P^0}{d\ln S^2} - \frac{dP^0}{d\ln S}}
\end{aligned}$$

$$= \frac{1}{\widehat{\sigma}_T\left(0,V\right)T}\left[\int_0^T \frac{\rho\sigma}{2}V_\tau\left(V\right)\frac{1-e^{-k(T-\tau)}}{k}\,d\tau\right]\frac{d}{d\ln S}\ln\left(\frac{d^2P^0}{d\ln S^2} - \frac{dP^0}{d\ln S}\right)$$

[1]The mistrustful reader is encouraged to compute $\kappa_3$ at order one in $\sigma$ – an easy task as the characteristic function of $\ln S$ is analytically known in the Heston model – to verify that one indeed recovers (6.15) from (5.88).

where we have expressed vega in terms of gamma using relationship (5.66). Using the formula for the Black-Scholes dollar gamma, we get:

$$\delta\widehat{\sigma}_{KT} = \frac{1}{\widehat{\sigma}_T^3 T^2}\left[\frac{\rho\sigma}{2}\int_0^T V_\tau(V)\,\frac{1-e^{-k(T-\tau)}}{k}\,d\tau\right]\left(\frac{\widehat{\sigma}_T^2 T}{2}+\ln\frac{K}{S}\right)$$

We now reinstate interest rate and repo. As mentioned above, we only need to replace $\frac{K}{S}$ with $\frac{K}{F_T}$, where $F_T$ is the forward for maturity $T$:

$$\delta\widehat{\sigma}_{KT} = \frac{1}{\widehat{\sigma}_T^3 T^2}\left[\frac{\rho\sigma}{2}\int_0^T V_\tau(V)\,\frac{1-e^{-k(T-\tau)}}{k}\,d\tau\right]\left(\frac{\widehat{\sigma}_T^2 T}{2}+\ln\frac{K}{F_T}\right) \tag{6.16}$$

We have used the shorthand notation $\widehat{\sigma}_T = \widehat{\sigma}_T(0,V) = \sqrt{V^0+(V-V^0)\frac{1-e^{-kT}}{kT}}$. The notation $\widehat{\sigma}_T$ is appropriate, as $\widehat{\sigma}_T$ is the VS volatility both in the order-zero and order-one expansion in $\sigma$. Denoting the ATMF skew by $\mathcal{S}_T$ and the ATMF volatility by $\widehat{\sigma}_{F_T T}$, (6.16) gives:

$$\widehat{\sigma}_{F_T T} = \widehat{\sigma}_T\left(1+\frac{\widehat{\sigma}_T T}{2}\mathcal{S}_T\right) \tag{6.17a}$$

$$\mathcal{S}_T = \left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T} = \frac{1}{\widehat{\sigma}_T^3 T^2}\,\frac{\rho\sigma}{2}\int_0^T V_\tau\,\frac{1-e^{-k(T-\tau)}}{k}\,d\tau \tag{6.17b}$$

**Short maturities**

Let us take the limit $kT \ll 1$. Formula (6.16) translates at lowest order in $kT$ in the following expressions for $\widehat{\sigma}_{F_T T}$ and $\mathcal{S}_T$:

$$\widehat{\sigma}_{F_T T} = \sqrt{V}\left(1+\frac{\rho\sigma T}{8}\right) \tag{6.18a}$$

$$\mathcal{S}_T = \frac{\rho\sigma}{4\sqrt{V}} = \frac{\rho\sigma}{4\widehat{\sigma}_{F_T T}} \tag{6.18b}$$

where the second equality in (6.18b) is correct at order 1 in $\sigma$.

**Long maturities**

We now take the limit $kT \gg 1$. Keeping only terms of order 1 in $1/kT$ we get:

$$\widehat{\sigma}_{F_T T} = \sqrt{V^0}\left(1+\frac{\rho\sigma}{4k}\right)+\frac{1}{2kT}\left(\frac{V-V^0}{\sqrt{V^0}}+\frac{\rho\sigma}{4k}\frac{V-3V^0}{\sqrt{V^0}}\right) \tag{6.19a}$$

$$\mathcal{S}_T = \frac{\rho\sigma}{2\sqrt{V^0}}\frac{1}{kT} \tag{6.19b}$$

### 6.6.2 Example

Consider the example of a smile for a three-month maturity generated with the following parameters, whose values are typical of index smiles: $V^0 = 0.04$, $k = 1$, $\sigma = 0.6$, $\rho = -80\%$.

Figure 6.4 shows $\widehat{\sigma}_{KT}$ as a function of $\ln\left(\frac{K}{F_T}\right)$ for three different values of $V$: 1%, 4%, 16%, corresponding approximately to ATMF volatilities around 10%, 20%, and 40%, respectively.

The reason why we choose to vary $V$ while keeping other parameters constant is that $V$ is the only state variable of the Heston model. Figure 6.4 provides an illustration of future smiles that the Heston model generates at time $T_1$ for maturity $T_2 = T_1 + 3$ months – as a function of $V_{T_1}$.

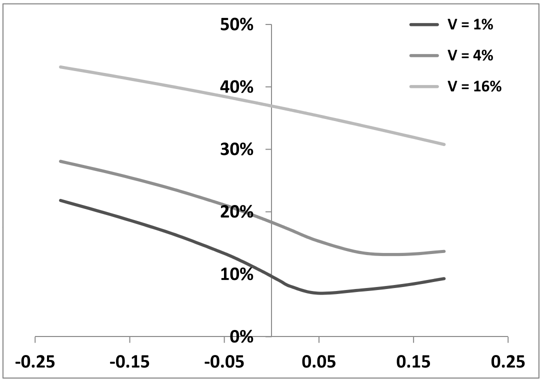


**Figure 6.4**: Implied volatilities $\widehat{\sigma}_{KT}$ as a function of $\ln\left(\frac{K}{F_T}\right)$ for a maturity $T = 3$ months, in the Heston model, generated with the following parameters: $V^0 = 0.04$, $k = 1$, $\sigma = 0.6$, $\rho = -80\%$, for three different values of $V$.

The quality of approximations (6.17a, 6.17b) for these parameter values is assessed in Tables 6.1 and 6.2.

| $V$ | $1\%$ | $4\%$ | $16\%$ |
|---|---|---|---|
| $\mathcal{S}_T^{\text{real}}$ | $-8.1$ | $-6.0$ | $-3.1$ |
| $\mathcal{S}_T^{\text{approx}}$ | $-8.8$ | $-5.5$ | $-3.0$ |

**Table 6.1**: ATMF skew in the Heston model for a three-month maturity, for different values of $V$ with $V^0 = 0.04$, $k = 1$, $\sigma = 0.6$, $\rho = -80\%$, computed either exactly: $\mathcal{S}_T^{\text{real}}$ or using approximation (6.17b): $\mathcal{S}_T^{\text{approx}}$. Results have been multiplied by 10 to represent approximately the difference in volatility points of implied volatilities for strikes $0.95F_T$ and $1.05F_T$.

| $V$ | 1% | 4% | 16% |
|---|---|---|---|
| $\widehat{\sigma}_T$ | 11.6 | 20 | 38.2 |
| $(\widehat{\sigma}_{F_T T} - \widehat{\sigma}_T)_{\text{real}}$ | −1.9 | −1.7 | −1.2 |
| $(\widehat{\sigma}_{F_T T} - \widehat{\sigma}_T)_{\text{approx}}$ | −0.1 | −0.3 | −0.5 |

**Table 6.2**: $\widehat{\sigma}_T$ and the difference $\widehat{\sigma}_{F_T T} - \widehat{\sigma}_T$, both in volatility points, in the Heston model for a three-month maturity, for different values of $V$ with $V^0 = 0.04$, $k = 1$, $\sigma = 0.6$, $\rho = -80\%$, computed either exactly (real) or using approximation (6.17a) (approx).

The ATMF skew is acceptably captured by an approximation at order one in $\sigma$: the maximum relative error is about 10%.

In contrast, the difference $\widehat{\sigma}_{F_T T} - \widehat{\sigma}_T$, is poorly estimated by (6.17a). Though both $\mathcal{S}_T$ and $\widehat{\sigma}_{F_T T} - \widehat{\sigma}_T$ are of order one in $\sigma$, the approximation for $\mathcal{S}_T$ is more robust.

This is generally observed in stochastic volatility models for equities: one typically needs to carry out the expansion of the ATMF volatility at order two in volatility of volatility to reach acceptable accuracy – more on this in Section 8.2.

### 6.6.3 Term structure of the ATMF skew

Equations (6.18b), (6.19b) show that while the ATMF skew tends to a constant for $T \to 0$, it decays like $\frac{1}{T}$ for long maturities, at order one in $\sigma$. The $\frac{1}{T}$ decay for long maturities is expected: because $V$ is mean-reverting, for maturities $T$ such that $T \gg \frac{1}{k}$, the distribution of returns of $\ln S$ over long periods becomes independent on the initial value of $V$, hence returns of $\ln S$ over long time scales become independent. The third-order cumulant $\kappa_3$ of $\ln S$ then scales like $T$, the skewness $s$ of $\ln S$ scales like $\frac{1}{\sqrt{T}}$ and (5.92) then implies that $\mathcal{S}_T \propto \frac{1}{T}$.

In the special case when $V = V^0$, the term structure of VS volatilities is flat, $V_t$ is constant and (6.16) takes the following simple form:

$$\left.\frac{d\widehat{\sigma}_{KT}}{d\ln K}\right|_{F_T} = \frac{\rho\sigma}{2\sqrt{V^0}}\frac{kT + e^{-kT} - 1}{(kT)^2} \tag{6.20}$$

Figure 6.5 shows an example of the term structure of the ATMF skew for the Euro Stoxx 50 and S&P 500 indexes, together with a power-law fit and a best fit using formula (6.20). The maturity dependence of the market skew indeed exhibits a power-law-like behavior, which cannot be captured by the Heston model for both short and long maturities: the Heston model is a one-factor model, with an embedded time scale $1/k$.

The issue here is not only about whether we are or aren't able to fit the vanilla smile. Rather, when risk-managing cliquets we may need to carry a naked forward smile position: it is then necessary to assess whether the model is able to generate *future* smiles – that is vanilla smiles at future dates – that are comparable to historically observed vanilla smiles – see the discussion below.

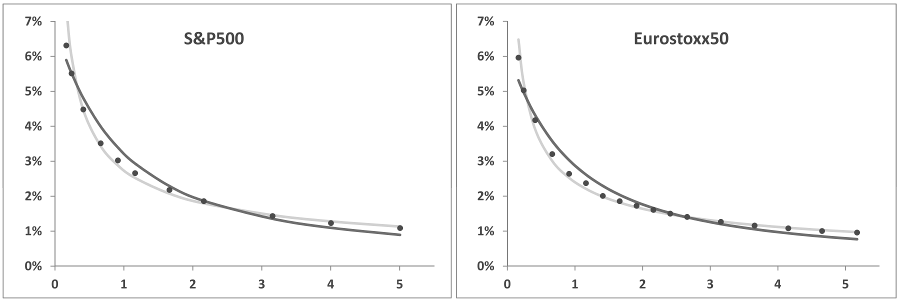


**Figure 6.5**: The ATMF skew as a function of maturity for the Euro Stoxx 50 and S&P 500 indexes (dots), observed on October 22, 2010, expressed as the difference in volatility points of the implied volatilities of strikes $0.95F_T$ and $1.05F_T$. A best fit using a power law with exponent $0.55$ (lighter line) as well as a best fit using formula (6.20) ($k = 2.9$) (darker line) are shown as well.

### 6.6.4 Relationship between ATMF volatility and skew

For short maturities, equation (6.18b) shows that, in the Heston model the ATMF skew is inversely proportional to the ATMF implied volatility – this is also evidenced in Figure 6.6. Is this dependence observed in reality?

Figure 6.6 displays the ATM volatility together with the ATM skew for a 3-month maturity, for the Euro Stoxx 50 index.[2] We can see that while skew and volatility seem to behave fairly independently, they are, if anything, positively correlated rather than negatively. It seems unreasonable to hard-wire an inverse dependence of the ATMF skew to the ATMF volatility in our model.

---

## 6.7 Discussion

The above analysis has highlighted some discrepancies between the spot/volatility dynamics generated by the Heston model on one hand, and observed in reality on the other hand – see also [8]. What makes the Heston model unsuitable for handling exotics however, rather than its inability to reproduce *exactly* the observed historical dynamics, is the lack of flexibility it affords.

[2] We have used the ATM volatility and skew for simplicity – using ATMF data would have yielded a similar graph.

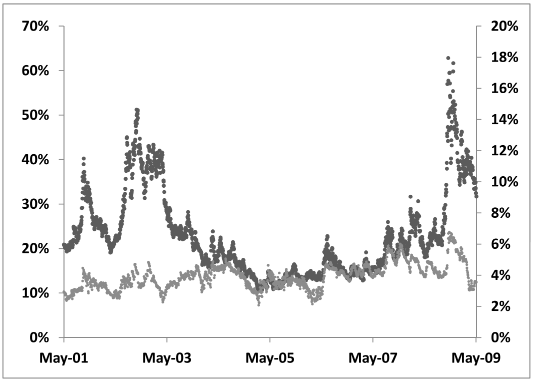


**Figure 6.6**: The ATM skew (lighter dots, right-hand axis) as the difference of the implied volatilities of the 95% and 105% strikes, and the ATM volatility (darker dots, left-hand axis) of the Euro Stoxx 50 index, for a 3-month maturity.

Indeed, from a trading perspective, one may choose to use parameter levels different than their historical averages and structural dependencies other than what is historically observed, even for parameters that have no implied counterpart. For example, even though the historical realized correlation between two quantities may be negligibly small, we will use bid/offer levels that are different than zero, depending on the size and sign of the option's sensitivity to correlation.

Likewise, imagine that in reality the short ATMF skew was indeed approximately inversely proportional to the short ATMF volatility, thus in line with the behavior generated in the Heston model. Still, when selling an exotic option that has positive forward ATMF volatility/skew cross-gamma: $\frac{d^2P}{d\widehat{\sigma}_{F_T T}\, d\mathcal{S}_T} > 0$, we may choose to conservatively price with a model that generates positive or vanishing covariance between ATMF volatility and skew – rather than negative.

Making the level of the short ATMF skew independent on the level of short ATMF volatility is easy: we only need to replace the SDE for $V_t$ in (6.1) with:

$$dV_t = \bullet\, dt + \nu V_t dZ_t$$

This would also have the advantage of making the short ATMF volatility lognormal rather than normal.

Other deficiencies of the Heston model are structural. For example, the scaling with maturity of the volatility of volatility in (6.6) and of the ATMF skew in (6.20) is intimately related to the fact that the Heston model is a one-factor model. Some have advocated making the parameters of the Heston model time-dependent so as to alter the dependence of $\mathcal{S}_T$ on $T$ and achieve accurate calibration of the vanilla smile.

Making $V^0$ time-dependent so as to best match the term structure of VS volatilities is appropriate, as forward variances can be hedged by trading variance swaps.[3] Making $\sigma$ and $\rho$ time-dependent is more questionable.

Indeed, consider the example of a call-spread cliquet paying at time $T_2$ the payoff $\left(\frac{S_{T_2}}{S_{T_1}} - 95\%\right)^+ - \left(\frac{S_{T_2}}{S_{T_1}} - 105\%\right)^+$, where $T_2 - T_1 \ll \frac{1}{k}$. This call spread has negligible sensitivity to volatility but is very sensitive to the ATMF skew observed at $T_1$ for maturity $T_2$. The order-one expansion in $\sigma$ in (6.18b) yields the following expression for the ATMF skew at $T_1$ for maturity $T_2$:

$$\mathcal{S}_{T_2}(T_1) = \frac{\rho(T_1)\,\sigma}{4\sqrt{V_{T_1}}}$$

where we have allowed $\rho$ to be time-dependent so as to match the term structure of the market *vanilla* ATMF skew. By pricing the cliquet with the Heston model thus calibrated we are betting on a given level of *forward* skew $\mathcal{S}_{T_2}(T_1)$, derived from calibration to today's term structure of the ATMF skew of vanilla options. Worse, the model will actually generate hedge ratios for our forward-start call spread on vanilla options of maturities $T_1$ and $T_2$.

This is not justified, as it is not possible to hedge forward-skew risk by trading vanilla options – we refer the reader to the illuminating experiment in Section 3.1.7. In case we are unable to take an offsetting position in a different exotic that has comparable forward-smile risk and have to keep the cliquet and its forward skew exposure on our book, it is more reasonable to *select* a conservative level of forward skew we are comfortable with, rather than having it dictated by calibration to market prices of instruments that are incapable of hedging it. The same goes for volatility-of-volatility risk.

Only if we are able to trade cliquets of varying maturities is it licit to make $\sigma$ or $\rho$ time-dependent, so as to match the term structure of the implied forward skew.

Fundamentally, the problem with the Heston model lies not so much with the model itself, its lack of flexibility, or even its peculiar idiosyncrasies, but with its usage (and its users): which practical pricing or hedging issue naturally calls for SDE (6.5)?

[3] Exact calibration is not guaranteed as $V_0(t)$ has to remain positive.

## Chapter's digest

▶ In terms of forward variances, the SDE of the Heston model reads:

$$\begin{cases} dS_t &= \sqrt{\xi_t^t} S_t dW_t \\ d\xi_t^T &= \sigma e^{-k(T-t)} \sqrt{\xi_t^t} dZ_t \end{cases}$$

with the following constraint on the initial values of forward variances:

$$\frac{d\xi_0^T}{dT} = -k\left(\xi_0^T - V^0\right) dT$$

VS volatilities are given by:

$$\widehat{\sigma}_T^2\left(t\right) = V^0 + \frac{1 - e^{-k(T-t)}}{k\left(T-t\right)} \left(V_t - V^0\right)$$

thus are floored.

▶ The drift of the instantaneous variance $V_t$ in stochastic volatility models has nothing to do with the "market price of risk". It is related to the initial slope of the variance curve:

$$dV_t = \left.\frac{d\xi_t^T}{dT}\right|_{T=t} dt + \lambda_t^t dZ_t^t$$

▶ For a flat term-structure of VS volatilities, at order one in volatility of volatility, the ATMF skew of the Heston model is given by:

$$\mathcal{S}_T = \frac{\rho\sigma}{2\sqrt{V^0}} \frac{kT + e^{-kT} - 1}{\left(kT\right)^2}$$

and the (lognormal) volatility of the VS volatility of maturity $T$, $\widehat{\sigma}_T$, has the form:

$$\mathrm{vol}(\widehat{\sigma}_T) \propto \frac{1 - e^{-kT}}{kT}$$

▶ At order one in volatility of volatility, the ATMF skew of the Heston model for short maturities is given by:

$$\mathcal{S}_T = \frac{\rho\sigma}{4\sqrt{V}}$$

and for long maturities:

$$\mathcal{S}_T = \frac{\rho\sigma}{2\sqrt{V^0}} \frac{1}{kT}$$

It decays as $\frac{1}{T}$.