On the relationship between implied volatilities and volatility swaps: a Malliavin calculus approach

Elisa Alòs

Universitat Pompeu Fabra
Barcelona

BACHELIER FINANCE SOCIETY ONE WORLD SEMINARS

Jointly with Kenichiro Shiraya (The University of Tokyo)

Contents

Contents

Volatility swaps, ATMI?

Volatility swaps, (the square root of) variance swaps, and the ATMI are 'volatility measures' that are considerered to be similar at short maturites.

In this talk we will discuss 'how close' are volatility swaps and the ATMI, from an analytical point of view.

Our approach does not requiere the volatility to be Markovian. That is, it can be applied to *rough* volatilities.

Some related research :

For a study of the connection between variance swaps and the ATMI we refer to El Euch, Fukasawa,Gatheral, and Rosenbaum (2018).

The difference between vol swaps and variance swaps is the object of a working paper with A. Muguruza.

Volatility swaps and implied volatilities

The main objective of this talk is the study of the difference between the fair strike of a volatility swap

$$E\sqrt{\frac{1}{T-t}\int_t^T \sigma_u^2 du}.$$

and the at-the-money implied volatility (ATMI) of a European call option.

If the volatility is not a constant but a stochastic process, the ATMI is not constant, but depends on T.

It is well-known that the difference between these two quantities converges to zero as the time to maturity decreases (see for example Carr and Lee (2008, 2009)).

In this paper, we make use of a Malliavin calculus approach to derive an exact expression for this difference.

Volatility swaps and implied volatilities

This representation allows us to establish that the order of the convergence is different in the correlated (the volatility and the asset price are correlated) and in the uncorrelated (the asset price and the volatility are uncorrelated) case.

Moreover, we will see that it depends on the behavior of the Malliavin derivative of the volatility process.

In particular, we will prove that for volatilities driven by a fractional Brownian motion, this order depends on the corresponding Hurst parameter H.

Moreover, in the case $H \geq 1/2$, we develop a model-free approximation formula for the of the volatility swap, in terms of the ATMI and its skew.

Contents

The Malliavin derivative

Let us consider a standard Browian motion $W = \{W_t, t \in [0, T]\}$ defined in a complete probability space $(\Omega, \mathcal{F}, P)$ .Set $H = L^2([0, T])$, and denote $W(h) = \int_0^T h(s)\, dW_s$ the Wiener integral of a deterministic function $h \in H$.

Let $\mathcal{S}$ be the set of random variables of the form

$$F = f(W(h_1), ..., W(h_n)),$$

where $n \geq 1, f \in \mathcal{C}_b^\infty(\mathbb{R}^n)$ (f and all its derivatives are bounded), and $h_1, .., h_n \in H$. Given a random variable F of this form, we define its derivative as the stochastic process $\left\{D_t^W F,\ t \in [0, T]\right\}$ given by

$$D_t^W F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), ..., W(h_n))\, h_i(t),\ \ t \in [0, T].$$

The Malliavin derivative

Example : If $F = W_T,\ D_t F = \mathbf{1}_{[0,T]}(t)$.

The operator D^W and the iterated operators $D^{W,n}$ are closable and unbounded from $L^2(\Omega)$ into $L^2([0,T]^n \times \Omega)$, for all $n \geq 1$. We denote by $\mathbb{D}_W^{n,2}$ the clousure of $\mathcal{S}$ with respect to the norm defined by

$$\|F\|_{n,2}^2 = \|F\|_{L^2(\Omega)}^2 + \sum_{k=1}^{n} \left\| D^{W,k} F \right\|_{L^2\left([0,T]^k \times \Omega\right)}^2 .$$

Malliavin calculus : examples

$\mathbb{D}_W^{1,2}$ denotes the domain of the Malliavin derivative operator D^W with respect to the Brownian motion W.

Example : If $F = W_T$, $D_t^W F = \mathbf{1}_{[0,T]}(t)$.

Example : If $F = W_T^H = \int_0^T K(T,s) dW_s$, $D_t^W F = K(T,s)\mathbf{1}_{[0,T]}(t)$.

We see that the Malliavin derivatives of both processes are very different

We recall that, for any u adapted and any $r > s$ $D_r^W u_s = 0$.

Malliavin derivatives are easy to compute

Let us see the example of a SABR volatility

$$\sigma_t = \sigma_0 \exp\left(-\frac{\alpha^2}{2} t + \alpha W_t\right).$$

Then, the chain rule gives us that the Malliavin derivative with respect to W is given by

$$D_r^W \sigma_t = \alpha \sigma_t \mathbf{1}_{[0,t]}(r).$$

Some useful properties of the Malliavin derivative

Notice that, for an adapted process u and for $s > t$

$$D_s^W u_t = 0$$

The Malliavin derivative operator satisfies the chain rule : for any $F \in \mathbb{D}_W^{1,2}$ and any function f satisfying some regularity conditions

$$D_t f(F) = f'(F) D_t F$$

In particular, the two above properties give us that, for $s > t$

$$D_s\Big(BS\big(t, X_t, K, v_t\big)\Big) = \frac{\partial BS}{\partial \sigma}\big(t, X_t, K, v_t\big) D_s^W v_t$$
$$= \frac{\partial BS}{\partial \sigma}\big(t, X_t, K, v_t\big) \frac{1}{2(T-t)v_t}\Big(\int_s^T D_s^W \sigma_r^2 dr\Big)$$

The Clark-Ocone-Haussman representation

Proposition (Clark-Ocone-Haussman formula)

Consider a random variable $F \in \mathbb{D}^{1,2}$. *Then*

$$F = E(F) + \int_0^T E_r(D_r F) dW_r.$$

Contents

The main problem and notations

Let us consider the following model for the log-price :

$$X_t = x - \frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t \sigma_s \left(\rho dW_s + \sqrt{1-\rho^2} dB_s\right), \quad t \in [0, T]. \tag{1}$$

Here, x is the current log-price, W and B are standard Brownian motions defined on a complete probability space $(\Omega, \mathcal{G}, P)$, and σ is a square-integrable and **right-continuous** stochastic process adapted to the filtration generated by W. In the following, we denote by $\mathcal{F}^W$ and $\mathcal{F}^B$ the filtrations generated by W and B. Moreover we define $\mathcal{F} := \mathcal{F}^W \vee \mathcal{F}^B$.

Remark

This model includes the fractional volatility models with $H < 1/2$ *introduced in Alòs, León and Vives (2007) (named recently as 'rough'volatilities).*

The main property of these models is their ability to reproduce a short-end blow-up of the skew slope of the implied volatility, as observed in real market data.

The main problem and notations

In the sequel, we make use of the following notation :

- $v_t = \sqrt{\frac{Y_t}{T-t}}$, where $Y_t = \int_t^T \sigma_u^2 du$. That is, v represents the future average volatility. Notice that $E_t[v_t]$ is the fair strike of a volatility swap with maturity time T.
- $BS(t, T, x, k, \sigma)$ denotes the price of a European call option under the classical Black-Scholes model with constant volatility σ, current log stock price x, time to maturity $T-t$, strike price $K = \exp(k)$.

The main problem and notations

- Fixed t, T, X_t, k we define the implied volatility $I(t, T, X_t, k)$ as the quantity such that

$$.BS(t, T, X_t, k, I(t, T, X_t, k)) = V_t := E_t(S_T - \exp(k))_+.$$

- $H(t, T, x, k, \sigma) := \left(\frac{\partial^3}{\partial x^3} - \frac{\partial^2}{\partial x^2}\right) BS(t, T, x, k, \sigma)$.
- $BS^{-1}(k, x) := BS^{-1}(t, T, X_t, k, x)$.

Contents

The uncorrelated case

Proposition

Consider the model (1) with $\rho = 0$*. Then, under some general conditions, the at-the-money implied volatility admits the representation*

$$
\begin{aligned}
& I\left(t, T, X_t, k^*\right) \\
= \;& E_t\left[v_t\right] + \frac{1}{32(T-t)} \\
& \times E_t\left[\int_t^T A_r\left(E_r\left[N'(d_+\left(k_t^*, v_t\right))\frac{\int_r^T D_r^W \sigma_s^2 ds}{v_t}\right]\right)^2 dr\right],
\end{aligned}
$$

where $A_r := \frac{BS^{-1}(k_t^*,\Lambda_r)}{(N'(d_+(k_t^*,BS^{-1}(k_t^*,\Lambda_r)\ )))^2}$ *and* $\Lambda_r := E_r\left[BS\left(t, T, X_t, K, v_t\right)\right]$

The uncorrelated case : proof

Notice that, in the uncorrelated case, the Hull and White formula gives us that the option price can be writen as

$$V_t = E_t \left[BS \left(t, T, X_t, K, v_t \right) \right].$$

Then the implied volatility satisfies that

The uncorrelated case : proof

$$
\begin{aligned}
& I(t, T, X_t, k) \\
&= BS^{-1}(k, V_t) \\
&= E_t\left[BS^{-1}(k, E_t\left[BS(t, T, X_t, K, v_t)\right])\right]
\end{aligned}
$$

$$
E_t[v_t] = E_t\left[BS^{-1}(k, BS(t, T, X_t, K, v_t))\right]
$$

That is,

$$I(t, T, X_t, k)\, E_t\left[BS^{-1}(k, \Lambda_t)\right]$$

and

$$E_t[v_t] = E_t\left[BS^{-1}(k, \Lambda_t)\right],$$

where

$$\Lambda_r := E_r\left[BS(t, T, X_t, K, v_t)\right].$$

The uncorrelated case : proof

Now, Λ admits a representation

$$\Lambda_r = \Lambda_0 + \int_0^r U_s dW_s$$

and then, as in Alòs and León (2017), we can write

$$BS\left(t, T, X_t, k, v_t\right) = E_t\left[BS\left(t, T, X_t, k, v_t\right)\right] + \int_t^T U_s dW_s,$$

where U_s can be computed (under some hypotheses) by Clark-Ocone-Haussman formula and W is the Brownian motion that drives the volatility process. Then the result follows from Itô's formula.

$$
\begin{aligned}
& E_t\left[BS^{-1}(k,\Lambda_t) - BS^{-1}(k,\Lambda_t)\right] \\
= & -\int_t^T \left(BS^{-1}\right)'\left(\Lambda_t + \int_t^r U_s dW_s\right) U_r dW_r \\
& -\frac{1}{2}\int_t^T \left(BS^{-1}\right)''\left(k,\Lambda_t + \int_t^r U_s dW_s\right) U_r^2 dr\Bigg] \\
= & -\frac{1}{2}E_t\int_t^T \left(BS^{-1}\right)''(k,\Lambda_r)\, U_r^2 dr, \qquad (2)
\end{aligned}
$$

where $(BS^{-1})''$ is the second order derivative with respect to Λ.

Then the results from the facts that

$$\left(BS^{-1}\right)''(k^*, \Lambda_r) = -\frac{BS^{-1}(k^*, \Lambda_r)}{4(T-t)\left(\exp(X_t)N'(d_+\left(k, BS^{-1}(k^*, \Lambda_r)\right))\right)^2},$$

and

$$\begin{aligned} U_r &= E_r\left(D_r^W\left(BS\left(t, T, X_t, k^*, v_t\right)\right)\right) \\ &= E_r\left(\exp(X_t)N'(k^*, d_+\left(v_t\right))\frac{\int_r^T D_r^W \sigma_s^2 ds}{2\sqrt{T-t}v_t}\right). \end{aligned}$$

The uncorrelated case : limit results

The main hypothesis we assume is the following one : There exists $\delta \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ such that the term

$$\frac{1}{(T-t)^{3+2\delta}} E_t \int_t^T \left(E_r \left[\int_r^T D_r^W \sigma_s^2 ds \right] \right)^2 dr,$$

has a finite limit as $T \to t$.

Remark

In classical diffusions, $\delta = 0$. For models based on the fBm, $\delta = H - \frac{1}{2}$.

The uncorrelated case : limit results

Theorem

Consider the model (1) with $\rho = 0$*. Then, under some general hypotheses*

$$\lim_{T \to t} \frac{I(t, T, X_t, k^*) - E_t[v_t]}{(T-t)^{2+2\delta}}$$

$$= \frac{1}{32\sigma_t} \lim_{T \to t} \frac{1}{(T-t)^{3+2\delta}} E_t \left[\int_t^T \left(E_r \left[\int_r^T D_r^W \sigma_s^2 ds \right] \right)^2 dr \right].$$

The uncorrelated case : limit results

Remark

In fractional volatility models, $\delta = H - 1/2$*. Then*
$2 + 2\delta = 2 + 2\left(H - 1/2\right)$ *and*
$I\left(t, T, X_t, k^*\right) - E_t\left[v_t\right] = O\left(T - t\right)^{1+2H}$.

The uncorrelated case : limit results

Remark

Previous results for the curvature (see Alós and León (2017)) give us that, in the uncorrelated case

$$\lim_{T\to t}\frac{I\left(t,T,X_t,k^*\right)-E_t\left[v_t\right]}{(T-t)^{2+2\delta}}=\frac{\sigma_t^4}{8}\lim_{T\to t}\frac{\frac{\partial^2 I}{\partial k^2}\left(t,T,X_t,k^*\right)}{(T-t)^{2\delta}}.$$

That is,

$$\lim_{T\to t}\frac{I\left(t,T,X_t,k^*\right)-E_t\left[v_t\right]}{(T-t)^{2+2\delta}}=\frac{1}{8}\lim_{T\to t}\frac{I\left(t,T,X_t,k^*\right)^4\frac{\partial^2 I}{\partial k^2}\left(t,T,X_t,k^*\right)}{(T-t)^{2\delta}}.$$

Contents

The correlated case : a decomposition formula

The key tool in this case is this decomposition of the implied volatility

Theorem

Consider the model (1) and assume that hypotheses (H1), (H2), (H3), (H4) and (H5) hold. Then

$$I(t,T,X_t,k^*) = I^0(t,T,X_t,k^*) + \frac{\rho}{2}\int_t^T \left(BS^{-1}\right)'(k^*,\Gamma_s)\, H(s,X_s,k^*,v_s)\Phi_s ds \quad (3)$$

where $I^0(t,T,X_t,k^*)$ *denotes the implied volatility in the uncorrelated case* $\rho = 0$,

$$\Gamma_s := E_t\left[BS(t,T,X_t,k,v_t)\right] + \frac{\rho}{2}E_t\int_t^s H(r,X_r,k,v_r)\Phi_r dr,$$

and $\Phi_t := \sigma_t \int_t^T D_t^W \sigma_r^2 dr$.

Theorem

Under some regularity conditions,

- *If* $\delta < 0$

$$
\begin{aligned}
&\lim_{T\to t} \frac{I(t,T,X_t,k_t^*) - E_t[v_t]}{(T-t)^{1+2\delta}} \\
&= \lim_{T\to t} \frac{3\rho^2}{8\sigma_t^3 (T-t)^{4+2\delta}} E_t\left(\int_t^T \int_s^T D_s^W \sigma_r^2 drds\right)^2 \\
&\quad - \lim_{T\to t} \frac{\rho^2}{2\sigma_t (T-t)^{3+2\delta}} E_t \int_t^T \left(\int_s^T D_s^W \sigma_r dr\right)^2 ds \\
&\quad - \lim_{T\to t} \frac{\rho^2}{2\sigma_t (T-t)^{3+2\delta}} E_t \int_t^T \int_s^T \int_r^T D_s^W D_r^W \sigma_u^2 dudrds.
\end{aligned}
$$

Theorem (cont)

- *If* $\delta > 0$

$$\lim_{T\to t} \frac{I(t,T,X_t,k_t^*) - E_t[v_t]}{(T-t)^{1+\delta}}$$
$$= \lim_{T\to t} \frac{\rho}{4(T-t)^{2+\delta}} E_t \int_t^T \int_s^T D_s^W \sigma_r^2 dr ds.$$

Theorem (cont)

- *If* $\delta = 0$

$$
\begin{aligned}
&\lim_{T\to t} \frac{I(t,T,X_t,k_t^*) - E_t[v_t]}{(T-t)} \\
&= \lim_{T\to t} \frac{3\rho^2}{8\sigma_t^3 (T-t)^4} E_t \left(\int_t^T \int_s^T D_s^W \sigma_r^2 dr ds \right)^2 \\
&\quad - \lim_{T\to t} \frac{\rho^2}{2\sigma_t (T-t)^3} E_t \int_t^T \left(\int_s^T D_s^W \sigma_r dr \right)^2 ds \\
&\quad - \lim_{T\to t} \frac{\rho^2}{2\sigma_t (T-t)^3} E_t \int_t^T \int_s^T \int_r^T D_s^W D_r^W \sigma_u^2 du dr ds \\
&\quad + \lim_{T\to t} \frac{\rho}{4(T-t)^2} E_t \int_t^T \int_s^T D_s^W \sigma_r^2 dr ds.
\end{aligned}
$$

The correlated case : limit results

Remark

In the fractional volatility case $\delta = H - 1/2$. *Then, in the correlated case*

- *If* $H \leq 1/2$, *then* $I(t, T, X_t, k^*) - E_t[v_t] = O(T-t)^{2H}$.
- *If* $H \geq 1/2$, *then* $I(t, T, X_t, k^*) - E_t[v_t] = O(T-t)^{H+1/2}$.

Notice that the order of the convergence is not affected only by the Hurst parameter but also by the correlation.

.

The correlated case : limit results

Remark

Moreover, taking into account the representation of the short-time limit skew in term of the Malliavin derivative of the volatility process (see Alòs, León and Vives (2007)) we get the following model-free approximation formula in the case $H > 1/2$

$$E_t\left[v_t\right] \approx I\left(t, T, X_t, k^*\right) - \frac{I\left(t, T, X_t, k^*\right)^2}{2} \frac{\partial I}{\partial k}\left(t, T, X_t, k^*\right)\left(T - t\right). \tag{4}$$

Estimating the Hurst parameter

Let us consider a linear regression analysis with dependent variable

$$\ln |I(0, T, X_0, k_0^*) - E[v_0]|$$

and independent variable $\ln T$. According to our previous results, the corresponding slope will be approximately $2H$ for $H \leq 1/2$ and $1/2 + H$ for $H > 1/2$.

This gives us a tool to estimate the Hurst parameter of the fractional volatility model.

In fact, if the obtained slope a is less than 1, then we will estimate H as $a/2$, while if $a \geq 1$, the Hurst parameter will be estimated by $a - 1/2$.

Some numerical results

TABLE – Hurst parameters obtained from linear regressions.

H index		Maturities	$T \leq 0.5$	$T \leq 0.4$	$T \leq 0.3$	$T \leq 0.2$	$T \leq 0.1$	$T \leq 0.01$
0.1	(A)	Slopes	0.244	0.241	0.237	0.232	0.227	0.215
		estimated H	0.122	0.120	0.118	0.116	0.113	0.107
	(B)	Slopes	-0.401	-0.401	-0.401	-0.402	-0.401	-0.401
		estimated H	0.099	0.099	0.099	0.098	0.099	0.099
0.3	(A)	Slopes	0.655	0.654	0.652	0.649	0.646	0.638
		estimated H	0.328	0.327	0.326	0.325	0.323	0.319
	(B)	Slopes	-0.200	-0.200	-0.200	-0.200	-0.200	-0.200
		estimated H	0.300	0.300	0.300	0.300	0.300	0.300
0.5	(A)	Slopes	1.002	1.002	1.002	1.001	1.002	1.002
		estimated H	0.501	0.501	0.501	0.501	0.501	0.501
	(B)	Slopes	0.000	0.000	0.000	-0.000	0.000	0.000
		estimated H	0.500	0.500	0.500	0.500	0.500	0.500
0.7	(A)	Slopes	1.242	1.241	1.240	1.238	1.235	1.229
		estimated H	0.742	0.741	0.740	0.738	0.735	0.729
	(B)	Slopes	0.200	0.200	0.200	0.200	0.200	0.200

Many thanks!