MTH 9875 The Volatility Surface: Fall 2017

Lecture 8: Asymptotics and dynamics of the volatility surface

Jim Gatheral
Department of Mathematics

Outline of lecture 8

- Volatility surface asymptotics
  - Short expirations
  - Long expirations
  - Small volatility of volatility
  - Extreme strikes

- Volatility surface dynamics
  - Under stochastic volatility
  - Under local volatility
- Stochastic implied volatility models
- Digital options and digital cliquets

Volatility surface asymptotics under stochastic volatility with jumps

- How does the shape of the volatility surface depend on the choice of model?

- In this lecture, we will see that all stochastic volatility with jumps style models generate volatility surfaces with a similar shape.
  - We will see that it's practically impossible to deduce anything about the specific form of the volatility dynamics from a single observation of the volatility surface.

Short time limit of local volatility under stochastic volatility

First we rewrite our generic stochastic volatility SDEs in terms of the log-stock price $x := \log(F/K)$ and under the risk neutral measure, specializing to the case where α and β do not depend on S or t:

(1)

$$\begin{aligned} dx_t &= -\frac{v_t}{2}\,dt + \sqrt{v_t}\,dZ_t \\ dv_t &= \alpha(v_t,t)\,dt + \eta\,\sqrt{v_t}\,\beta(v_t,t)\,dW_t. \end{aligned}$$

We may rewrite

$$dW_t = \rho\,dZ_t + \bar{\rho}\,dZ_t^{\perp}$$

with $\bar{\rho} = \sqrt{1-\rho^2}$ and $\mathbb{E}\left[dZ_t^{\perp}, dZ_t\right] = 0$.

Eliminating $\sqrt{v}\,dZ_t$, we get

$$\begin{aligned} dv_t = \alpha\,(v_t,t)\ dt + \rho\,\eta\,\beta\,(v_t,t)\ \left\{dx_t + \frac{v_t}{2}\,dt\right\} + \\ \bar{\rho}\,\eta\,\beta\,(v_t,t)\ \sqrt{v_t}\,dZ_t^{\perp}. \end{aligned}$$

Then,

$$\mathbb{E}\left[v + dv\,|dx\right] = v + \alpha\,dt + \rho\eta\,\beta\ \left\{dx + \frac{v}{2}dt\right\}$$

and so for small times to expiration (relative to the variation of α and β), we have

(2)

$$\begin{aligned} v_{\ell}(x,t) &= \mathbb{E}\left[v_t\,|x_t = x\right] \\ &\approx v_0 + \left[\alpha(v_0,0) + \rho\eta\,\frac{v_0}{2}\,\beta(v_0,0)\right]t + \rho\,\eta\,\beta(v_0,0)\,x. \end{aligned}$$

- We can just read off the short term variance skew from the form of the variance SDE.

Short time limit of implied volatility under stochastic volatility

To extend the result to implied volatility, we need the following lemma (see also Problem 3 of HW2):

> **Lemma**
>
> The local volatility skew is twice as steep as the implied volatility skew for short times to expiration

We know that Black-Scholes implied total variance is roughly the integral of local variance along the most probable path from the stock price on the valuation date to the strike price at expiration. This path is approximately a straight line (see Figure 1).

Diagrammatic proof

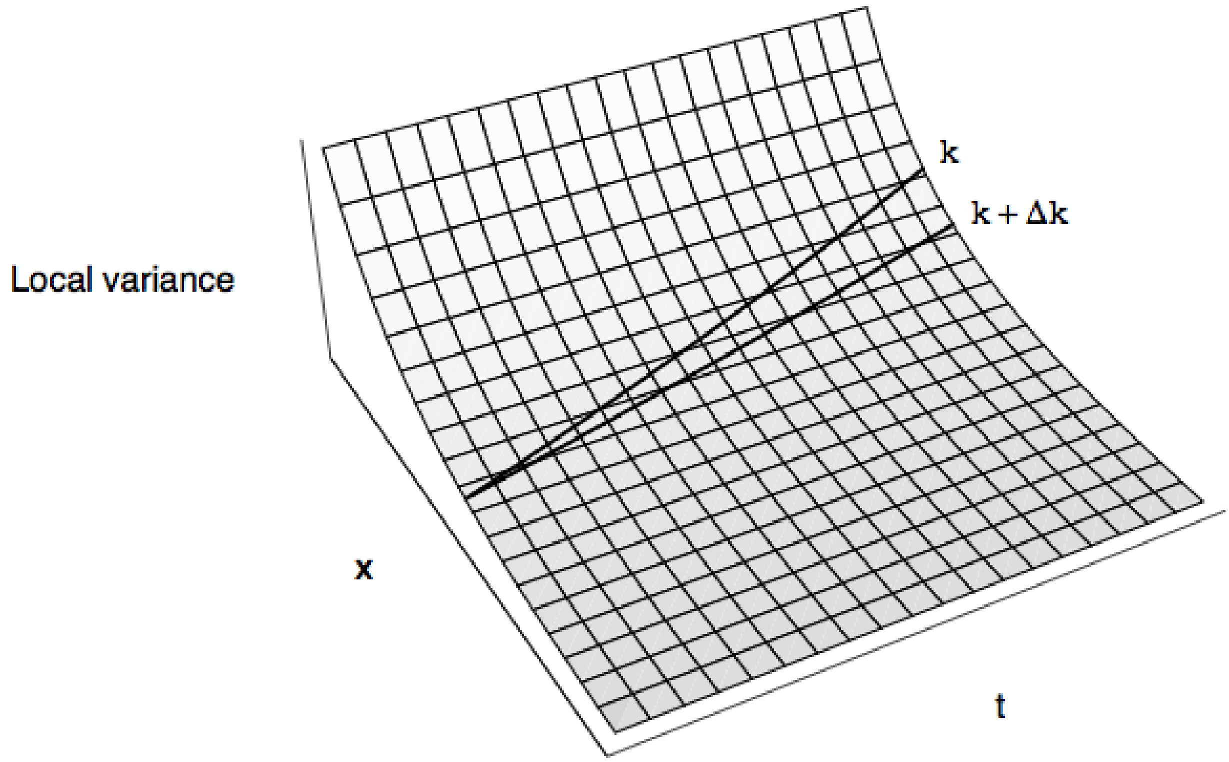

Figure 1: For short expirations, the most probable path is approximately a straight line from spot on the valuation date to the strike at expiration. It follows that $\sigma^2_{BS}(k,T) \approx [v_{loc}(0,0) + v_{loc}(k,T)]/2$ and the implied variance skew is roughly one half of the local variance skew.

Also, from equation (2), we see that the slope of the local variance skew is a roughly constant $\beta(v_0)$ for short times. The BS implied variance skew, being the average of the local variance skews, is one half of the local variance skew.

Formally,

$$
\begin{aligned}
\sigma_{BS}(K,T)^2 &\approx \frac{1}{T}\int_0^T v_{loc}(\tilde{x}_t, t)dt \\
&\approx \text{const.} + \frac{1}{T}\int_0^T \rho\,\eta\,\beta(v_0,0)\,\tilde{x}_t dt \\
&\approx \text{const.} + \frac{1}{T}\int_0^T \rho\,\eta\,\beta(v_0,0)\,x_T\,\frac{t}{T}dt \\
&= \text{const.} + \frac{1}{2}\,\rho\,\eta\,\beta(v_0,0)\,x_T
\end{aligned}
$$

where $\tilde{x}$ represents the *most likely path* from the stock price at time zero to the strike price at expiration.

We conclude that for short times to expiration, the BS implied variance skew is given by

(3)

$$\frac{\partial}{\partial x}\sigma_{BS}(x,t)^2 \;=\; \frac{\rho\,\eta}{2}\beta(v_0)$$

Recall that in the Heston model, $\beta(v) = 1$; we see that equation (3) is consistent with the short-dated volatility skew behavior that we derived earlier for the Heston model.

Medvedev-Scaillet: The no jump case

[Medvedev and Scaillet][12] develop a perturbation expansion for small times to expiration τ and fixed normalized log-strike z defined as

$$z := \frac{k}{\sigma_{BS}(k,\tau)\,\sqrt{\tau}}$$

First we specialize their result to the case where the underlying process is a diffusion of the form

(4)

$$\begin{aligned}\frac{dS_t}{S_t} &= \sigma_t\,dZ_t\\ d\sigma_t &= a\,(\sigma_t)\,dt \,+\, b\,(\sigma_t)\;dW_t\end{aligned}$$

In the no-jump (pure stochastic volatility) case, the implied volatility I has the following short-term asymptotics (with $\sigma = \sigma_0$)

$$I(z,\tau;\sigma) = \sigma + I_1(z;\sigma)\,\sqrt{\tau} + I_2(z;\sigma)\,\tau + O(\tau\,\sqrt{\tau})$$

where I_1 and I_2 are functions of the z and the instantaneous volatility $\sigma = \sqrt{v}$ only:

(5)

$$\begin{aligned}I_1(z;\sigma) &= \frac{\rho\,b(\sigma)\,z}{2}\\ I_2(z;\sigma) &= \frac{1}{6}\left\{\frac{b(\sigma)^2\,(1-\rho^2)}{\sigma} + \rho^2\,b(\sigma)\,\partial_\sigma b(\sigma)\right\}\,z^2\\ &\quad+\frac{a(\sigma)}{2} + \frac{\rho\,\sigma\,b(\sigma)}{4} + \frac{1}{24}\,\frac{\rho^2\,b(\sigma)^2}{\sigma}\\ &\quad+\frac{1}{12}\,\frac{b(\sigma)^2}{\sigma} - \frac{1}{6}\,\rho^2\,b(\sigma)\,\partial_\sigma b(\sigma).\end{aligned}$$

Remarks on Medvedev-Scaillet

- First note that the limit of implied volatility as the log-strike $k \to 0$ and the time to expiration $\tau \to 0$ is just the instantaneous volatility σ.

  - Critics of stochastic volatility models love to point out that instantaneous volatility is not an observable, supposedly a deficiency of SV models.
  - In practice, in liquid option markets, the implied volatility surface is typically very smooth and we can extrapolate to the zero expiration, at-the-money strike limit with little uncertainty.

- The Medvedev-Scaillet expansion as originally presented is purely formal. Recently, [Friz, Gerhold and Pinter][7] showed how to make this expansion rigorous.
  - One application (naturally) is to rough volatility models.

The short-dated volatility skew

To compute the short-dated volatility skew, we substitute

$$z = \frac{k}{I(z,\tau;\sigma)\,\sqrt{\tau}}$$

into (5) and take the limit $\tau \to 0$, to obtain

(6)

$$\left.\frac{\partial I}{\partial k}\right|_{k=0} \to \frac{\rho\, b(\sigma)}{2\,\sigma}$$

which proves our earlier result (3) derived using heuristic methods.

Remarks on the short-dated skew in SV models

- Under SV, the short-dated volatility skew is not explicitly time-dependent.
  - It depends only on the form of the SDE for volatility.

- In contrast, as we saw, local volatility models imply that short-dated skews in the future will be much flatter than today's.

- So even if we find a stochastic volatility model and a local volatility model that price all European options identically today, forward-starting options (that is options whose strikes are to be set some time in the future) cannot possibly be priced identically by these two models.
  - Both models fit the options market today but the volatility surface dynamics implied by the two models are quite different.

A volatility skew conjecture

- Equations (3) and (6) suggest a generalization:
  - Perhaps all stochastic volatility models, whether analytically tractable or not, generate the same BS implied volatility skew up to a factor of $\beta(v)$, not just in the limit $\tau \to 0$ but for all $\tau \geq 0$?

- We will now present further evidence that makes this claim more plausible.

The SABR model

Pat Hagan's well-known SABR (or "stochastic alpha beta rho") model [Hagan, Lesniewski et al.][9] has dynamics

$$\begin{aligned} dS_t &= \sigma_t\, S_t^\beta\, dZ_t \\ d\sigma_t &= \xi\, \sigma_t\, dW_t \end{aligned}$$

with $\mathbb{E}\left[dZ_t\, dW_t\right] = \rho\, dt$.

- Volatility does not mean revert in the SABR model so it is only good for short expirations.
- Nevertheless the model has the virtue of having an exact expression for the implied volatility smile in the short-expiration limit $\tau \to 0$
  - The resulting functional form can be used to fit observed short-dated implied volatilities and the model parameters α, β and ρ thereby extracted.

The SABR model with $\beta = 1$

In the special case $\beta = 1$, the SABR implied volatility formula reduces to

(7)

$$\sigma_{BS}(k) = \sigma_0\, \frac{y}{f(y)}\, \left\{1 + \left[\frac{1}{4}\,\rho\,\xi\,\sigma_0 + \frac{2-3\rho^2}{24}\,\xi^2\right]\,\tau + O(\tau^2)\right\}$$

with

$$y := -\xi\,\frac{k}{\sigma_0}$$

and

$$f(y) = \log\left\{\frac{\sqrt{1-2\,\rho\,y+y^2}+y-\rho}{1-\rho}\right\}$$

Taylor expansion of the SABR formula

Note that the lognormal SABR formula (7) factorizes with one factor depending only on y and the other factor depending only on τ.

Taylor-expanding (7) to second order in y and first order in τ (*i.e.* with $y \sim \sqrt{\tau}$) gives

$$\begin{aligned} \sigma_{BS}(k,\tau) = \sigma_0\, &\left\{1 - \frac{1}{2}\rho\,y + \frac{2-3\,\rho^2}{12}\,y^2\right. \\ &\left.+ \left[\frac{1}{4}\,\rho\,\xi\,\sigma_0 + \frac{2-3\rho^2}{24}\,\xi^2\right]\,\tau + O(\tau\,\sqrt{\tau})\right\} \end{aligned}$$

Substituting $a(\sigma) = 0$ and $b(\sigma) = \xi\,\sigma_0$ into equation (5) gives

$$
\begin{aligned}
I_1(z;\sigma_0) &= \frac{1}{2}\,\rho\,\xi\,\sigma_0\,z \\
I_2(z;\sigma_0) &= \frac{1}{6}\,\xi^2\,\sigma_0\,z^2 + \frac{1}{4}\,\rho\,\xi\,\sigma_0^2 \\
&\quad + \frac{1}{24}\,\rho^2\,\xi^2\,\sigma_0 + \frac{1}{12}\,\xi^2\,\sigma_0 - \frac{1}{6}\,\rho^2\,\xi^2\,\sigma_0
\end{aligned}
$$

Consistency with Medvedev-Scaillet

Then, noting that

$$
\xi\,z\,\sqrt{\tau} = \xi\,\frac{k}{\sigma_{BS}} = -y\,\frac{\sigma_0}{\sigma_{BS}} = -y\,\left(1 + \frac{1}{2}\,\rho\,y\right) + O(y^3)
$$

we obtain

$$
\begin{aligned}
\sigma_{BS}(k,\tau) &= \sigma_0 + I_1(z;\sigma_0)\,\sqrt{\tau} + I_2(z;\sigma_0)\,\tau + O(\tau\,\sqrt{\tau}) \\
&= \sigma_0\,\left\{1 - \frac{1}{2}\rho\,y + \frac{2-3\,\rho^2}{12}\,y^2\right. \\
&\qquad \left. + \left[\frac{1}{4}\,\rho\,\xi\,\sigma_0 + \frac{2-3\rho^2}{24}\,\xi^2\right]\,\tau + O(\tau\,\sqrt{\tau})\right\}
\end{aligned}
$$

We see that the Medvedev-Scaillet formula (5) gives precisely the same result as the SABR implied volatility formula (7) for small τ.

Consistency with our generic short-dated skew formula

Finally, we note that the SABR formula implies that

$$
\left.\frac{\partial \sigma_{BS}}{\partial k}\right|_{k=0} = \frac{\rho\,\xi}{2}
$$

which is a special case of the general result (6) with $\beta(v) = \sqrt{v}$ and $\eta = 2\,\xi$.

To see this, apply Itô's Lemma to the SABR volatility process to obtain

$$
dv = \xi^2\,v\,dt + 2\,\xi\,v\,dZ
$$

with $v = \sigma^2$.

Long expirations

[Fouque, Papanicolaou and Sircar][6] show using a perturbation expansion approach that in any stochastic volatility model where volatility is mean-reverting, BS implied volatility can be well approximated by a simple function of k and τ for long-dated options.

In particular, they study a model where the log-volatility is a Ornstein-Uhlenbeck process (log-OU for short). That is:

$$
\begin{aligned}
dx &= -\frac{\sigma^2}{2}\,dt + \sigma\,dZ_1 \\
d\log(\sigma) &= -\lambda[\log(\sigma) - \overline{\log(\sigma)}]dt + \xi\,dZ_2
\end{aligned}
$$

They find that the slope of the BS implied volatility skew is given (for large $\lambda\,T$) by

(10)

$$
\frac{\partial}{\partial x}\,\sigma_{BS}(x,T) \approx \frac{\rho\,\xi}{\lambda\,T}
$$

Consistency with skew conjecture

To recast this in terms of v to be consistent with the form of the generic process we wrote down in equation (1), we note that (considering random terms only), $dv \sim 2\,\sigma\,d\sigma$ and in the log-OU model,

$$
d\sigma \sim \xi\,\sigma dZ_2
$$

So

$$
dv \sim 2\,\xi\,v\,dZ_2
$$

Then $\beta(v)$ as defined in equation (1) is given by

$$
\eta\,\beta(v) = 2\xi\,\sqrt{v}
$$

and, from equation (10), the BS implied variance skew is given by

$$
\frac{\partial}{\partial x}\,\sigma_{BS}(x,T)^2 \approx \frac{2\,\rho\,\xi\,\sqrt{v}}{\lambda\,T} = \frac{\rho\,\eta\,\beta(v)}{\lambda\,T}
$$

Consistency with Heston long-dated skew

Recall from Lecture 5 that the Heston skew (where $\beta(v) = 1$) has this same behavior for large $\lambda\,T$.

- It seems that both for long and short expirations, the skew behavior may be identical for all stochastic volatility models up to a factor of $\beta(v)$.
- The natural way to interpolate the asymptotic skew behaviors between long and short expirations would then be:

  (11)

$$
\frac{\partial}{\partial x}\,\sigma_{BS}(x,T)^2 \approx \frac{\rho\,\eta\,\beta(v)}{\lambda'\,T}\left\{1 - \frac{\left(1-e^{-\lambda' T}\right)}{\lambda'\,T}\right\}
$$

  with $\lambda' = \lambda - \frac{1}{2}\,\rho\,\eta\,\beta(v)$.

Including Jumps

Medvedev and Scaillet's main result is a more complicated expression for models that include jumps in the stock price.

Specifically, consider the stochastic volatility with jump model

(8)

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t \, dZ_1 \, + J(\sigma_t) \, dq_t \\ d\sigma_t &= a\,(\sigma_t)\, dt \, + \, b\,(\sigma_t) \, dZ_2 \end{aligned}$$

The jump term dq is a standard Poisson process with intensity $\lambda_J(\sigma_t)$.

In this model, short-dated implied volatilities are given by

$$I(z, \tau; \sigma) = \sigma + \tilde{I}_1(z; \sigma) \sqrt{\tau} + \tilde{I}_2(z; \sigma)\, \tau + O(\tau \sqrt{\tau})$$

$\tilde{I}_1$ and $\tilde{I}_2$ are given by

(9)

$$\begin{aligned} \tilde{I}_1(z; \sigma) &= I_1(z; \sigma) - \mu_J \, g(z) + \eta_J \, h(z) \\ \tilde{I}_2(z; \sigma) &= I_2(z; \sigma) + \frac{1}{2\,\sigma} \left(\mu_J \, g(z) - \eta_J \, h(z)\right)^2 z^2 \\ &\quad - \left\{ -\frac{\mu_J \, \sigma}{2} - \sigma \, \lambda_J + \frac{\mu_J^2}{\sigma} + \frac{\mu_J \, b(\sigma) \, \rho}{2\,\sigma} \right\} g(z)\, z \\ &\quad - \left\{ \frac{\eta_J \, \sigma}{2} + \sigma \, \chi_J - \frac{\mu_J \, \eta_J}{\sigma} - \frac{\eta_J \, b(\sigma)\, \rho}{2\,\sigma} \right\} h(z)\, z \\ &\quad + \frac{\rho \, b(\sigma) \, \mu_J}{2\,\sigma} - \frac{\rho \, \partial_\sigma b(\sigma) \, \mu_J}{2} + \frac{\mu_J^2}{2\,\sigma} - \frac{\sigma \, \mu_J}{2} - \lambda_J \, \sigma \end{aligned}$$

where $\eta_J = \lambda_J \, \int_0^\infty \, x f(x) \, dx$, $\chi_J = \lambda_J \, \int_0^\infty \, f(x) \, dx$are respectively the positive part of the jump compensator and the probability of an upwards jump and

$$g(z) = \frac{N(-z)}{N'(z)}; \; h(z) = \frac{1}{N'(z)}.$$

Corollary

In a jump diffusion model (with volatility deterministic) the limit of the implied volatility skew as $\tau \to 0$ is given by

$$\left.\frac{\partial I}{\partial k}\right|_{k=0} \to -\frac{\mu_J}{\sigma}.$$

To get this result, note that $g'(0) = 1$ and $h'(0) = 0$. The result is exactly consistent with our earlier heuristic derivation in Lecture 6.

Corollary

In the SVJ model, the limit of the implied volatility skew as $\tau \to 0$ is given by

$$\left.\frac{\partial I}{\partial k}\right|_{k=0} \to \frac{\rho\, b(\sigma)}{2\,\sigma} - \frac{\mu_J}{\sigma}$$

This is consistent with our earlier observation that the jump and stochastic volatility effects on the at-the-money variance skew are approximately additive. In fact we have

$$\left.\frac{\partial v_{BS}}{\partial k}\right|_{k=0} \to \rho\, b(\sigma) - 2\,\mu_J \text{ as } \tau \to 0$$

so they are exactly additive at $\tau = 0$!

Perturbation expansions

- There are many ways of generating expansions assuming some parameter is small.
  - The most useful is probably small volatility-of-volatility η.
- [Lewis][11] originally did a hard volatility of volatility perturbation computation for the Heston model years ago.
- [Bergomi and Guyon][3] found a general way to do volatility-of-volatility perturbation expansions up to second order in η for a very general stochastic volatility model with any number of factors.
  - Their result is almost (but not quite) identical to the Lewis solution in the Heston case.
  - The BG expansion works even under rough volatility (*i.e.* even for non-Markovian models).
- Recently, [Alòs, Gatheral and Radoičić] [1] showed how to extend this expansion to all orders in η.
  - The idea is to express the cumulant generating function as an infinite sum of BG-style covariance functionals.
  - [Jacquier and Lorig] [10] present an algorithm to map the expansion coefficients of the characteristic function to the smile expansion.

The Bergomi and Guyon model setup

Bergomi-Guyon (BG) dynamics are:

$$\begin{aligned} dx_t &= -\frac{1}{2}\,\xi_t(t)\,dt + \sqrt{\xi_t(t)}\,dZ_t \\ d\xi_t(u) &= \lambda(t,u,\xi_t).\,dW_t, \quad \xi_0(u) = \xi(u). \end{aligned}$$

$\xi_t(u) = \mathbb{E}\left[v_u | \mathcal{F}_t\right]$ is the forward variance curve at time t and $Z = \{Z_1, \ldots, Z_d\}$ is a d −dimensional Brownian motion.

- In particular, the Heston model may be written in this form.
- So can more complicated multi-factor models such as the Bergomi model and the rough Bergomi model.

The Bergomi and Guyon expansion

Using a technique from quantum mechanics, [Bergomi and Guyon][3] compute an expansion of the volatility smile up to second order in volatility of volatility for stochastic volatility models written in variance curve form.

The Bergomi-Guyon expansion of implied volatility takes the form

(12)

$$\sigma_{\rm BS}(k,t) = \hat\sigma_T + \mathcal{S}_T\,k + \mathcal{C}_T\,k^2 + O(\epsilon^3)$$

Here

$$\begin{aligned}
\hat\sigma_T &= \sqrt{\frac{w}{T}}\,\left\{1 + \frac{1}{4\,w}\,C^{x\xi}\right.\\
&\quad \left. + \frac{1}{32\,w^3}\,\left(12\,(C^{x\xi})^2 + w\,(w+4)C^{\xi\xi} + 4\,w\,(w-4)\,C^{\mu}\right)\right\}\\
\mathcal{S}_T &= \sqrt{\frac{w}{T}}\,\left\{\frac{1}{2\,w^2}\,C^{x\,\xi} + \frac{1}{8\,w^3}\,\left(4\,w\,C^{\mu} - 3(C^{x\xi})^2\right)\right\}\\
\mathcal{C}_T &= \sqrt{\frac{w}{T}}\,\frac{1}{8\,w^4}\,\left(4\,w\,C^{\mu} + w\,C^{\xi\xi} - 6(C^{x\xi})^2\right)
\end{aligned}$$

where $w = \int_0^T\,\xi_0(s)\,ds$ is total variance to expiration T.

Bergomi and Guyon correlation functionals

The various correlation functionals appearing in the BG expansion are:

$$\begin{aligned}
C^{x\,\xi} &= \int_0^T\,dt\,\int_t^T\,du\,\frac{\mathbb{E}\left[dx_t\,d\xi_t(u)\right]}{dt}\\
C^{\xi\,\xi} &= \int_0^T\,dt\,\int_t^T\,ds\,\int_t^T\,du\,\frac{\mathbb{E}\left[d\xi_t(s)\,d\xi_t(u)\right]}{dt}\\
C^{\mu} &= \int_0^T\,dt\,\int_t^T\,du\,\frac{\mathbb{E}\left[dx_t\,dC_t^{x\,\xi}\right]}{dt}.
\end{aligned}$$

What are these correlation functionals?

- $C^{x\,\xi}$ is an integral of the term structure of covariances between returns and forward variances.
- $C^{\xi\,\xi}$ is an integral of the variance curve autocovariance function.
- C^{μ} is somewhat more complicated. Something like the covariance between the skew and underlying returns.

- The BG expansion gives us a direct correspondence between the implied volatility surface and the joint dynamics of the underlying and the forward variance curve.
- In principle (assuming $\mathbb{Q} = \mathbb{P}$), we could even compute the covariance functionals from the time series of implied volatility surfaces.

Example: The Heston model

We computed before that

$$\xi_t(u) = (v_t - \bar{v})\, e^{-\kappa\,(u-t)} + \bar{v}.$$

It follows that in forward variance form, the Heston model reads

$$d\xi_t(u) = e^{-\kappa\,(u-t)}\, dv_t = e^{-\kappa\,(u-t)}\, \eta\, \sqrt{v_t}\, dW_t.$$

Then in particular

$$\mathbb{E}\,[dx_t\, d\xi_t(u)] = \rho\, \eta\, v_t\, e^{-\kappa\,(u-t)}\, dt.$$

With $v_0 = \bar{v}$ to simplify computations, we obtain

$$\begin{aligned} C^{x\xi} &= \rho\, \eta\, \bar{v} \int_0^T dt \int_t^T e^{-\kappa\,(u-t)}\, du \\ &= \rho\, \eta\, \bar{v} \int_0^T dt \int_t^T e^{-\kappa\,(u-t)}\, du. \end{aligned}$$

Term structure of ATM skew in the Heston model

Define the at-the-money (ATM) volatility skew

$$\psi(T) = \partial_k \sigma_{BS}(k,T)|_{k=0}$$

It follows from (12) (with $v_0 = \bar{v}$ for simplicity) that to first order in η,

$$\begin{aligned} \psi(T) = \mathcal{S}_T &= \sqrt{\frac{w}{T}}\, \frac{1}{2\,w^2}\, C^{x\,\xi} \\ &= \frac{\rho\,\eta}{2\,\sqrt{\bar{v}}}\, \frac{1}{\kappa\, T} \left\{ 1 - \frac{1 - e^{-\kappa\, T}}{\kappa\, T} \right\}. \end{aligned}$$

- In the Heston model,
  - $\psi(T) \to$ a constant as $T \to 0$,
  - $\psi(T) \sim \frac{1}{T}$ as $T \to \infty$.

Example: The SABR model

$\mathbb{E}\,[\sigma_u | \mathcal{F}_t] = \sigma_t$ so to first order in the volatility of volatility α,

$$\xi_t(u) = v_t.$$

It follows that, to first order again,

$$d\xi_t(u) = 2\,\sigma_t\,d\sigma_t = 2\,\alpha\,\sigma_t^2\,dW_t.$$

Then

$$\mathbb{E}\left[dx_t\,d\xi_t(u)\right] = 2\,\rho\,\alpha\,{\sigma_t}^3\,dt.$$

Again to first order in α, we obtain

$$C^{x\xi} = 2\,\rho\,\alpha\,\sigma_0^3\,\int_0^T\,dt\,\int_t^T\,du = \rho\,\alpha\,\sigma_0^3\,T^2$$

and so

$$\psi(T) = \mathcal{S}_T = \frac{\rho\,\alpha}{2}.$$

- $\psi(T)$ is independent of T in the SABR model!
  - There is no mean reversion and so no term structure of ATM skew.

The Bergomi model

The n-factor Bergomi variance curve model[1] reads:

(13)

$$\xi_t(u) = \xi_0(u)\,\exp\left\{\sum_{i=1}^n \eta_i\,\int_0^t\,e^{-\kappa_i\,(t-s)}\,dW_s^{(i)} + \text{ drift }\right\}.$$

The Bergomi model generates a term structure of volatility skew $\psi(T)$ that looks something like

$$\psi(\tau) \sim \sum_i\,\frac{1}{\kappa_i\,T}\left\{1 - \frac{1-e^{-\kappa_i\,T}}{\kappa_i\,T}\right\}.$$

- This functional form is related to the term structure of the functional $C^{x\xi}$.
  - Which is in turn driven by the exponential kernel in the exponent in (13).

Term structure of ATM skew in conventional stochastic volatility models

Thus, in general, conventional Markovian stochastic volatility models generate a term structure of ATM skew of the form

$$\psi(\tau) \sim \sum_i\,\frac{1}{\kappa_i\,\tau}\left\{1 - \frac{1-e^{-\kappa_i\,\tau}}{\kappa_i\,\tau}\right\}.$$

In contrast, recall from Lecture 1 that we observe empirically something like $\psi(\tau) \sim \tau^{-\gamma}$ with typically $\gamma \approx 0.4$.

- Markovian stochastic volatility models are not consistent with the shape of the volatility surface!

The skew-stickiness ratio

Recall from Lecture 1 that the *skew-stickiness ratio* or *SSR* is defined as

$$\mathcal{R}(T) = \frac{\beta(T)}{\psi(T)}$$

where $\beta(T)$ is the regression coefficient in the regression

$$\delta\sigma(T) = \alpha + \beta(T)\,\frac{\delta S}{S} + \text{noise}$$

and $\sigma(T)$ is the T-expiration ATM implied volatility. Then

$$\beta(T) = \mathbb{E}\left[\delta\sigma(T) \left| \frac{\delta S}{S}\right.\right].$$

Computation of the regression coefficient $\beta(T) = \mathbb{E}\left[\delta\sigma(T)|\delta S/S\right]$

We approximate ATM volatility using the variance swap.

- By definition this is

$$\mathcal{V}_t(T) = \int_t^T \xi_t(u)\,du.$$

Then

$$2\,\sigma(T)\,\delta\sigma(T)\,T \approx \int_t^T \delta\xi_t(u)\,du.$$

Thus

$$\begin{aligned}\beta(T) &= \mathbb{E}\left[\delta\sigma(T)|\delta S/S\right] \\ &\approx \frac{1}{2\,\sigma(T)\,T}\int_t^T \mathbb{E}\left[\delta\xi_t(u)\left|\frac{\delta S}{S}\right.\right]du.\end{aligned}$$

Also,

$$\mathbb{E}\left[\delta\xi_t(u)\left|\frac{\delta S}{S}\right.\right] \approx \frac{\mathbb{E}\left[\delta\xi_t(u)\,\delta x_t\right]}{\mathbb{E}\left[\delta{x_t}^2\right]} = \frac{\rho(t,u)}{v_t}$$

with $x_t = \log S_t$ and

$$\rho(t,u) = \frac{\mathbb{E}\left[d\xi_t(u)\,dx_t\right]}{dt}.$$

$\beta(T)$ and $C^{x\xi}$

Then, restoring explicit dependence on t and T,

$$C_t^{x\,\xi}(T) \;=\; \int_t^T ds \int_s^T \rho(s,u)\,du$$

and

$$\begin{aligned}\beta(T) \;&\approx\; \frac{1}{2\,\sigma(T)\,T}\,\frac{1}{v_t}\int_t^T \rho(t,u)\,du\\ &\approx\; -\,\frac{1}{2\,\sigma(T)\,T}\,\frac{1}{v_t}\,\frac{\partial C_t^{x\,\xi}(T)}{\partial t}.\end{aligned}$$

Assuming time-homogeneity, $C^{x\xi}$ would be a function of $\tau = T - t$ only and $v_t \approx \sigma^2$ and so we further approximate

$$\beta(\tau) \approx \frac{1}{2\,\sigma^3\,\tau}\frac{\partial C^{x\,\xi}(\tau)}{\partial \tau}.$$

Also, from the Bergomi-Guyon expansion, the ATM skew is given by

$$\psi(\tau) \approx \sqrt{\frac{w}{T}}\,\frac{1}{2\,w^2}\,C_t^{x\,\xi}(T) \approx \frac{1}{2\,\sigma^3}\,\frac{1}{\tau^2}\,C^{x\,\xi}(\tau).$$

Thus

$$\mathcal{R}(\tau) = \frac{\beta(\tau)}{\psi(\tau)} \approx \tau\,\frac{d}{d\tau}\log C^{x\xi}(\tau).$$

> For emphasis, to first order in the volatility of volatility, the SSR $\mathcal{R}$ is given by the time derivative of the spot-volatility correlation functional $C^{x\xi}$.

SSR examples

The SABR model

In the SABR model, $C^{x\xi}(\tau) \propto \tau^2$ so $\mathcal{R}(\tau) \approx 2$.

The Heston model

$$C^{x\xi}(\tau) = \rho\,\eta\,\bar{v}\int_0^\tau dt\int_t^\tau e^{-\kappa\,(u-t)}\,du$$

- For $\tau \ll 1/\kappa$, $C^{x\xi}(\tau) \sim \tau^2$ and $\mathcal{R}(\tau) \approx 2$.
- For $\tau \gg 1/\kappa$, $C^{x\xi}(\tau) \sim \tau$ and $\mathcal{R}(\tau) \approx 1$.

The n-factor Bergomi model

Let κ_1 be the shortest timescale (largest) mean reversion coefficient and κ_n be the longest timescale (smallest) mean reversion coefficient.

- For $\tau \ll 1/\kappa_1$, $C^{x\xi}(\tau) \sim \tau^2$ and $\mathcal{R}(\tau) \approx 2$.
- For $\tau \gg 1/\kappa_n$, $C^{x\xi}(\tau) \sim \tau$ and $\mathcal{R}(\tau) \approx 1$.

For stochastic volatility models in general, $\mathcal{R}(\tau) \approx 2$ for τ small and $\mathcal{R}(\tau) \approx 1$ for τ large.

Implication for the "true" model

Recall from Lecture 1 that empirically, we see $\mathcal{R}(\tau) \sim \frac{3}{2}$ for all τ. That is

$$\mathcal{R}(\tau) \approx \tau \, \frac{d}{d\tau} \log C^{x\xi}(\tau) \approx 2 - \gamma.$$

with $\gamma \approx 1/2$. Thus $C^{x\xi}(\tau) \sim \tau^{2-\gamma}$, and so $\rho(t,u) \sim (u-t)^{-\gamma}$.

An obvious model that would generate $\rho(t,u) \sim (u-t)^{-\gamma}$ is

$$\frac{d\xi_t(u)}{\xi_t(u)} \propto \frac{dW_t}{(u-t)^\gamma}.$$

- In effect, replace all the exponential kernels in the Bergomi model (13) with a power-law kernel.
- Such a model would be non-Markovian. The price of an option would depend on the entire history $\{W_s,\ s<t\}$ of the Brownian motion.
- To be consistent with $\mathcal{R}(\tau) \sim \frac{3}{2}$, we would need $\gamma \approx \frac{1}{2}$.
  - Such a model would also generate the observed term structure of ATM skew.

When we present Rough Volatility, we will see that scaling properties of the time series of realized variance also suggest such a model.

The short-dated skew in the Rough Heston model

The rough Heston model (with no mean reversion) may be written in forward variance form as

$$\begin{aligned}
\frac{dS_t}{S_t} &= \sqrt{v_t}\, dZ_t \\
d\xi_t(u) &= \frac{\nu}{\Gamma(1-\gamma)} \, \frac{\sqrt{v_t}}{(u-t)^\gamma} \, dW_t
\end{aligned}$$

where $\gamma = \frac{1}{2} - H$.

- The rough Heston is thus a natural generalization of the conventional Heston model with the exponential kernel replaced with a power-law kernel.

Thus

$$\rho(t,u) = \frac{\mathbb{E}\left[d\xi_t(u)\, dx_t\right]}{dt} = \frac{\rho\,\nu}{\Gamma(1-\gamma)}\,\frac{v_t}{(u-t)^\gamma}\,dt$$

and

$$\begin{aligned} C_t^{x\,\xi}(T) &= \mathbb{E}\left[\int_t^T ds\, \int_s^T \rho(s,u)\,du \middle| \mathcal{F}_t\right] \\ &= \frac{\rho\,\nu}{\Gamma(1-\gamma)}\int_t^T ds\, \int_s^T \frac{\mathbb{E}\left[v_s \middle| \mathcal{F}_t\right]}{(u-s)^\gamma}\,dt \\ &= \frac{\rho\,\nu}{\Gamma(2-\gamma)}\,\int_t^T \xi_t(s)\,(T-s)^{1-\gamma}\,ds. \end{aligned}$$

The skew-stickiness ratio again

Rough Heston is thus an explicit example of a model in which

$$\mathcal{R}(\tau) \approx \tau\,\frac{d}{dT}\log C_t^{x\xi}(T) \approx 2-\gamma = H + \frac{3}{2}.$$

Asymptotic wing behavior in the Heston model

[Drăgulescu and Yakovenko][4] compute the tail behavior of the Heston probability distribution function f and find it to be linear in $|k|$. That is,

$$f(k) \sim e^{-k\,q^+} \text{ as } k \to +\infty$$

and

$$f(k) \sim e^{+k\,q^-} \text{ as } k \to -\infty,$$

for some constants $q_\pm$.

Thus, the cumulative distribution function F has the tail behavior

(14)

$$F(k) \sim \frac{e^{+k\,q^-}}{k} \text{ as } k \to -\infty$$

and

(15)

$$\tilde{F}(k) = 1 - F(k) \sim \frac{e^{-k\,q^+}}{k} \text{ as } k \to +\infty.$$

Benaim and Friz applied to Drăgulescu and Yakovenko

- [Benaim and Friz][2] show how to get the tail behavior of the smile from the cumulative distribution function F.

- Specifically, for the right tail, substituting from (14) and (15), we obtain for the right tail

$$\frac{\sigma_{BS}(k,T)^2\,T}{k} \sim g\left(-1-\frac{\log\left[1-F(k)\right]}{k}\right) \sim g\left(-1+q^{+}\right) \text{ as } k \to \infty$$

- and for the left tail

$$\frac{\sigma_{BS}(-k,T)^2\,T}{k} \sim g\left(\frac{-\log F(-k)}{k}\right) \sim g\left(q^{-}\right) \text{ as } k \to \infty.$$

So Heston implied variance is linear in the extreme wings.

- As noted by Drăgulescu and Yakovenko, qualitatively similar results from other authors suggest that linearity in the tails is a generic feature of stochastic volatility models, not just the Heston model.

Asymptotics in Summary

- The general shape of the volatility surface doesn't seem to depend very much on the specific choice of conventional Markovian stochastic volatility model with jumps.

- Any such model should generate a similar shape of volatility surface with appropriate numerical choices of the parameters.

Dynamics of the volatility surface: Model dependence

- So it seems that all stochastic volatility models have essentially the same implications for the shape of the volatility surface.

- At first it might therefore seem that it would be hard to differentiate between models.
  - That would certainly be the case if we were to confine our attention to the shape of the volatility surface today.

- If instead we were to study the dynamics of the volatility skew – in particular, how the observed volatility skew depends on the overall level of volatility, we would be able to differentiate between models.

Dynamics of the volatility skew under stochastic volatility

Empirical studies of the dynamics of the volatility skew show that $\frac{\partial}{\partial k}\sigma(k,t)$ is approximately independent of volatility level over time. Translating this into a statement about| the implied variance skew, we get

$$\frac{\partial}{\partial k}\,\sigma_{BS}(k,t)^2 \;=\; 2\,\sigma_{BS}(k,t)\,\frac{\partial}{\partial k}\,\sigma_{BS}(k,t) \;\sim\; \sqrt{v(k,t)}.$$

Comparing this with equation (11), we see that this in turn implies that $\beta(v) \sim \sqrt{v}$.

- Referring back to the definition of $\beta(v)$ in equation (1), we conclude that v is approximately lognormal in contrast to the square root process assumed by Heston.
- This makes intuitive sense given that we would expect volatility to be more volatile if the volatility level is high than if the volatility level itself is low.

Does it matter what dynamics we choose?

- Does it matter whether we model variance as a square root process or as lognormal?
  - In certain cases it does.
  - We are using our model to hedge and the hedge should approximately generate the correct payoff at the boundary.
  - If the payoff that we are hedging depends (directly or indirectly) on the volatility skew, and our assumption is that the variance skew is independent of the volatility level, we could end up losing a lot of money if that's not how the market actually behaves.

- Is any stochastic volatility model better than none at all?
  - Yes! Because whereas having the wrong stochastic volatility model will cause the hedger to generate a payoff corresponding to a skew that may perhaps be off by a factor of 1.5 if volatility doubles,
  - Having only a local volatility model will cause the hedger to generate a payoff that corresponds to almost no forward skew at all.

Dynamics of the volatility skew under local volatility

- Empirically, the slope of the volatility skew decreases with time to expiration.

- In the case of mean-reverting single factor stochastic volatility, the term structure of the BS implied variance skew will look something like equation (11).

- Under local volatility, the slope of the volatility skew will decay over time according to the decay of today's volatility with respect to time to expiration.
  - As $\tau^{-0.4}$ say.

Dynamics of the volatility skew under local volatility

Recall formula (1.10) from TVS for local volatility in terms of implied volatility:

$$v_\ell = \frac{\frac{\partial w}{\partial T}}{1 - \frac{k}{w}\frac{\partial w}{\partial k} + \frac{1}{4}\left(-\frac{1}{4} - \frac{1}{w} + \frac{k^2}{w^2}\right)\left(\frac{\partial w}{\partial k}\right)^2 + \frac{1}{2}\frac{\partial^2 w}{\partial k^2}}.$$

To first order in $\frac{\partial w}{\partial k}$ (which is small for large T) , we have

$$v_\ell \approx \frac{\partial w}{\partial T}\left(1+\frac{k}{w}\frac{\partial w}{\partial k}\right)$$

Differentiating with respect to k then gives

$$\frac{\partial v_\ell}{\partial k} \approx \frac{\partial}{\partial T}\frac{\partial w}{\partial k} + \frac{1}{w}\frac{\partial w}{\partial T}\frac{\partial w}{\partial k}.$$

That is, the local variance skew $\frac{\partial v_{loc}}{\partial k}$ decays with the BS implied total variance skew $\frac{\partial w}{\partial k}$.

Dynamics of the volatility skew under local volatility

- To get the forward volatility surface in a local volatility model, we integrate over the local volatilities from the (forward) valuation date to the expiration of the option along the most likely path.
  - The forward volatility surface will be substantially flatter than today's because the forward local volatility skews are all flatter.

- Contrast this with a stochastic volatility model where implied volatility skews are approximately time-homogeneous.
  - Local volatility models imply that future BS implied volatility surfaces will be flat (relative to today's).
  - Stochastic volatility models imply that future BS implied volatility surfaces will look like today's.

Stochastic implied volatility models

- Many authors have looked at models that allow the entire implied volatility surface to diffuse.

- If the underlying price process is assumed continuous (with no jumps), the statics and dynamics of the implied volatility surface are highly constrained.
  - In particular, non-discounted option prices are risk-neutral expectations of future cashflows and as such must be martingales.

-
  - Changes in the call price reflect both changes in the underlying and changes in implied volatility.
  - Imposing the martingale constraint

$$\mathbb{E}[dC_t] = 0$$

    gives a tight relationship between the various sensitivities and many results such as equation (3) follow immediately from this.

Relationship between volatility surface shape and dynamics under stochastic volatility

- [Durrleman][5] showed how to extract the dynamics of instantaneous variance from the dynamics of the observed implied volatility surface in the limit of very short expirations and very close to at-the-money.

- Conversely, given a stochastic volatility model, he showed how to deduce the shape of the implied volatility surface in that same neighborhood.
  - However, to get these impressive results, one has to assume that the volatility surface remains finite as time to expiration $\tau \to 0$ (see Assumption 5 of [5]).
  - As we have seen earlier, the observed implied volatility surface (away from ATM) explodes as $\tau \to 0$.
  - This can be explained with jumps or rough volatility[8].

Digital options

- The valuation of a digital option involves the volatility skew directly.

- A digital (call) option $D(K, T)$ pays 1 if the stock price S_T at expiration T is greater than the strike price K and zero otherwise.

- It may be valued as the limit of a call spread as the spread between the strikes is reduced to zero.

(16)

$$D(K, T) = -\frac{\partial C(K, T)}{\partial K}$$

where $C(K, T)$ represents the price of a European call option with strike K expiring at time T.

Sensitivity of digital to the volatility skew

To see that the price of a digital is very sensitive to the volatility skew, we rewrite the European call price in equation (16) in terms of its Black-Scholes implied volatility $\sigma_{BS}(K, T)$:

$$\begin{aligned} D(K, T) &= -\frac{\partial}{\partial K} C_{BS}\left(K, T, \sigma_{BS}(K, T)\right) \\ &= -\frac{\partial C_{BS}}{\partial K} - \frac{\partial C_{BS}}{\partial \sigma_{BS}} \frac{\partial \sigma_{BS}}{\partial K}. \end{aligned}$$

Impact of volatility skew in practice

- Consider a one year digital option struck at-the-money (with zero rates and dividends).

- Suppose further that at-the-money volatility is 25% and the volatility skew (typical of SPX for example) is 3% per 10% change in strike.

- The value of a digital call is given by:

$$
\begin{aligned}
D(1,1) &= -\frac{\partial C_{BS}}{\partial K} - \frac{\partial C_{BS}}{\partial \sigma_{BS}}\frac{\partial \sigma_{BS}}{\partial K} \\
&= N\left(-\frac{\sigma}{2}\right) - \text{vega} \times \text{skew} \\
&= N\left(-\frac{\sigma}{2}\right) + \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}} \times 0.3 \\
&\approx N\left(-\frac{\sigma}{2}\right) + 0.4 \times 0.3
\end{aligned}
$$

Getting the skew right is critical!

> If we had ignored the skew contribution, we would have got the price of the digital option wrong by 12% of notional!

Digital cliquets

Here is part of a definition of the word cliquet from the Dictionary of Financial Risk Management:

> The French like the sound of ‘cliquet’ and seem prepared to apply the term to any remotely appropriate option structure. (1) Originally a periodic reset option with multiple payouts or a ratchet option (from vilbrequin à cliquet – ratchet brace). Also called Ratchet Option ...

And since the word is originally French, here is an elegant definition of the “Effet-cliquet” from a French website:

> Mécanisme qui permet de figer une performance même si l’actif correspondant baisse par la suite.

(Mechanism that permits a profit to be locked in even if the underlying subsequently declines.)

Payoff of a digital cliquet

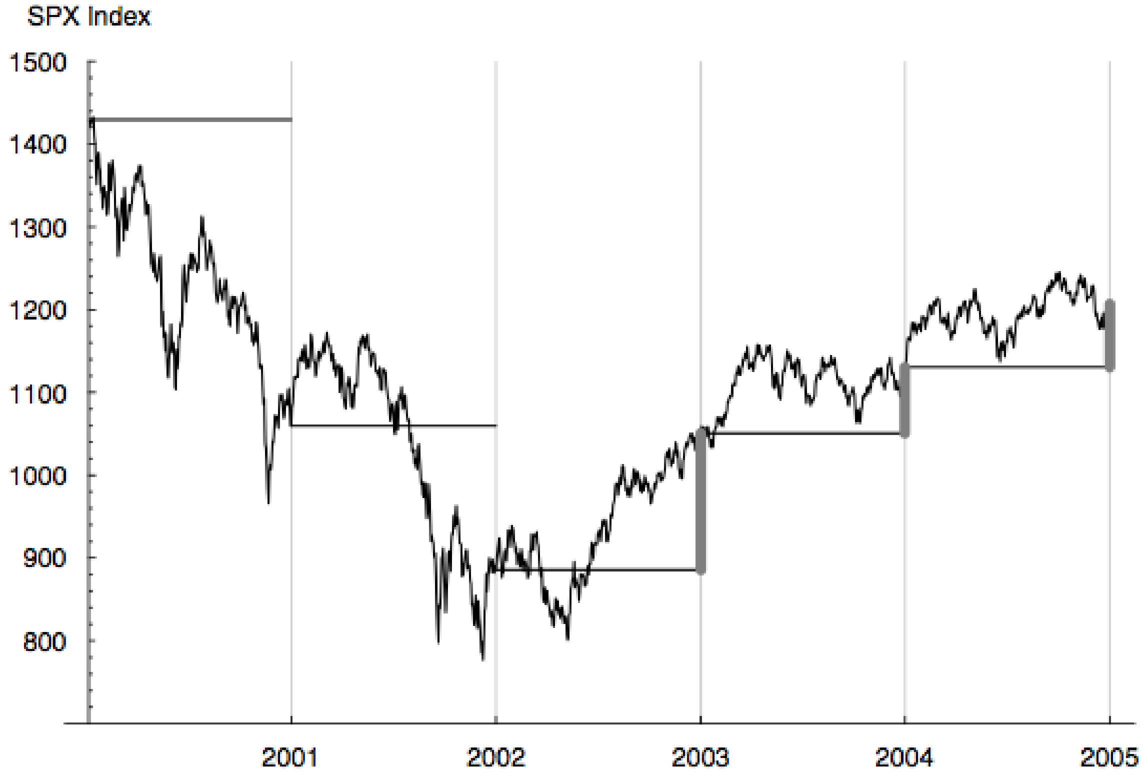

Figure 4: Illustration of a cliquet payoff. This hypothetical SPX cliquet resets at-the-money every year on October 31st. The thick solid lines represent nonzero cliquet payoffs. The payoff of a 5-year European option struck at the 10/31/2000 SPX level of 1429.40 would have been zero.

Generic cliquets

- A cliquet is just a series of options whose strikes are set on a sequence of futures dates.
  - A digital cliquet is a sequence of digital options whose strikes will be set (usually) at the prevailing stock price on the relevant reset date.

- Denoting the set of reset dates by $\{t_1, t_2, \ldots, t_n\}$, the digital cliquet pays $\mathrm{Coupon} \times \theta\left(\mathrm{S_{t_i}} - \mathrm{S_{t_{i-1}}}\right)$ at t_i where $\theta(.)$ represents the Heaviside function.

Economic motivation for cliquets

- The package consisting of a zero coupon bond together with a digital cliquet makes a very natural product for a risk-averse retail investor.
  - The investor typically gets an above market coupon if the underlying stock index is up for the period (usually a year) and a below market coupon (usually zero) if the underlying stock index is down.

- This product was very popular and as a result, many equity derivatives dealers had digital cliquets on their books.

How to lose money

- We deduce that the price of a digital cliquet may vary very substantially depending on the modeling assumptions made by the seller.
- Those sellers using local volatility models will certainly value a digital cliquet at a lower price than sellers using a stochastic volatility.
  - Or more practically, those guessing that the forward skew should look like today's.
- Perversely then, those sellers using an inadequate model will almost certainly win the deal and end up short a portfolio of misvalued forward-starting digital options.
- Or even worse, a dealer could have an appropriate valuation approach but be pushed internally by the salespeople to match (mistaken) competitor's lower prices.

How wrong could the price of the digital cliquet be?

- Consider the example of a (not unrealistic) five year deal that has a 6% coupon annually if the underlying exceeds the prior annual setting and zero otherwise.
- Neglecting the first coupon (because we suppose that all dealers can price a digital which sets today), the error could be up to 12% of the sum of the remaining coupons (48%) or 5.76% of Notional.
- A pricing error of this magnitude is a big multiple of the typical margin on such a trade and would cause the dealer a substantial loss.

References

1. ^ Elisa Alòs, Jim Gatheral, and Radoš Radoičić, Exponentiation of conditional expectations under stochastic volatility, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2983180 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2983180), (2017).
2. ^ Shalom Benaim and Peter Friz, Regular variation and smile asymptotics, *Mathematical Finance* **19**(1), 1–12, (2009).
3. ^ Lorenzo Bergomi and Julien Guyon, The smile in stochastic volatility models. *SSRN* (2011).
4. ^ Adrian A Drăgulescu and Victor M Yakovenko, Probability distribution of returns in the Heston model with stochastic volatility, *Quantitative Finance* **2** 443–453 (2002).
5. ^ Valdo Durrleman, From implied to spot volatilities, *Finance and Stochastics* **14** (2) 157–177 (2010).
6. ^ Jean-Pierre Fouque, George Papanicolaou and Ronnie Sircar, Financial modeling in a fast mean-reverting stochastic volatility environment, *Asia-Pacific Financial Markets* **6** 37-48 (1999).
7. ^ Peter K. Friz, Stefan Gerhold, and Arpad Pinter, Option pricing in the moderate deviations regime, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2759347 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2759347) (2016).
8. ^ Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum, Volatility is rough, *Quantitative Finance* forthcoming (2017).
9. ^ Patrick S Hagan, Deep Kumar, Andrew S Lesniewski and Diana E Woodward, Managing smile risk, *Wilmott Magazine* **1**(1) 84–108 (2002).

10. ^ Antoine Jacquier and Matthew Lorig, From characteristic functions to implied volatility expansions, *Advances in Applied Probability* **47**(3), 837-857, (2015).
11. ^ Alan L. Lewis, *Option Valuation under Stochastic Volatility with Mathematica Code*, Finance Press: Newport Beach, CA (2000).
12. ^ Alexey N Medvedev and Olivier Scaillet, Approximation and calibration of short-term implied volatilities under jump-diffusion stochastic volatility models with jumps by short term asymptotics, *Review of Financial Studies* **20**(2), 427–459 (2007).


```
In [ ]:
```