Riding on the Smiles

Martino Grasselli, University of Padova and ESILV

6th World Congress of the Bachelier Finance Society
Toronto, June 22-26, 2010

Joint work with J. da Fonseca
Outline of the presentation:

2. Calibration of single asset multi-dimensional stochastic volatility models
3. Calibration of multi-asset multi-dimensional stochastic volatility models
4. Price approximations
On the calibration of the Heston (1993) model

\[
\frac{dS_t}{S_t} = \sqrt{v_t}dW_t^1
\]
\[
dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^2
\]
\[
dW_t^1dW_t^2 = \rho dt
\]

\(\rho\) controls the link between vol and asset returns

\[\downarrow\]

The Skew or Leverage
Analytic and Financial properties

- Characteristic function of the asset returns

\[ \mathbb{E}_t \left[ e^{i\omega \log(S_{t+\tau})} \right] = e^{A(\tau)v_t + B(\tau) \log(S_t) + C(\tau)} \]

- \( A(\tau) \) solves a Riccati ODE: explicit solution!

- Quasi closed form option prices via Fast Fourier Transform (Carr and Madan 1999)

- Sensitivity analysis, vol of vol asymptotic expansion..

- Each parameter has a clear financial interpretation
Quoting vanilla options

The implied volatility $\sigma_{imp}$ is the quantity such that

$$C_{mkt}(t, T, S_t, K) = c_{bs}(t, T, S_t, K, \sigma_{imp}^2(T - t))$$

(1)
The Smiles
Important facts

The skew is controlled by $\rho$

$\Downarrow$

Term structure of skews

$\Downarrow$

We should have different values for $\rho$

and

above $T - t > 0.1$ the smiles are similar
The choice of the Criterium: pitfall of the price
LSE

Calibration of vanilla options (OTM), maturities available

$$\min \frac{1}{N} \sum_{i=1}^{N} (C_{\text{model}}(t, T_i, K_i) - C_{\text{mkt}}(t, T_i, K_i))^2$$

<table>
<thead>
<tr>
<th>error</th>
<th>$\rho$</th>
<th>$t_{min}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.25E-07</td>
<td>-0.7095</td>
<td>0.05</td>
</tr>
<tr>
<td>2.06E-07</td>
<td>-0.7001</td>
<td>0.1</td>
</tr>
</tbody>
</table>

I don’t take the first maturity
• short term options have small (if no) impact on the solution

• the calibration seems to be good

• poor fit of short term options

What is the problem?

short term options have small time value w.r.t long term options  
\[
\downarrow
\]
small/no impact on the objective
The volatility LSE

\[ \min \frac{1}{N} \sum_{i=1}^{N} (\sigma_{imp}^{model}(t, T_i, K_i) - \sigma_{imp}^{mkt}(t, T_i, K_i))^2 \]

- more weight on short term options
- adding jumps does not help because jumps impact the very short part of the smile
**Calibration tests (Vol norm)**

<table>
<thead>
<tr>
<th>error</th>
<th>$\rho$</th>
<th>$(T - t)_{min}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00010773</td>
<td>-0.5562</td>
<td>0.05</td>
</tr>
<tr>
<td>4.31E-05</td>
<td>-0.6324</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Calibration date: 28/08/08

Maturities 0.06= 19/09, 0.13= 17/10 .. 4.31

- to fit the **short term skew** a low correlation is needed.
Why extending the Heston model?

• The **dynamics** of the implied volatility surface (vanilla options) and the Variance Swap curve are driven by **several factors**

• On the FX market the skew is **stochastic**

• We have a term structure of skew: short term skew \(\neq\) long term skew
Double-Heston model

(Christoffensen, Heston, Jacobs 2007)

\[
\frac{dS_t}{S_t} = \sqrt{v^1_t} dZ^1_t + \sqrt{v^2_t} dZ^2_t
\]

\[
dv^1_t = \kappa^1 (\theta^1 - v^1_t) dt + \sigma^1 \sqrt{v^1_t} dW^1_t
\]

\[
dv^2_t = \kappa^2 (\theta^2 - v^2_t) dt + \sigma^2 \sqrt{v^2_t} dW^2_t
\]

\[
dZ^1_t dW^2_t = \rho^1 dt
\]

\[
dZ^2_t dW^1_t = \rho^2 dt
\]

but

\[
\begin{align*}
&dZ^1_t dZ^2_t = dW^1_t dW^2_t = dZ^1_t dW^2_t = dZ^2_t dW^1_t = 0
\end{align*}
\]
Recall the Duffie-Filipovic-Schachermayer (2003)’s condition

If $X_t = (X_t^1, X_t^2)^\top$ is a vector affine square root process (thus positive):

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \ldots dt + \begin{pmatrix} \times & 0 \\ 0 & \times \end{pmatrix} d \begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix}$$

$$\Downarrow$$

We have strong constraints on the diffusion

$$\Downarrow$$

Strong constraints on the correlation!!

$$\Downarrow$$

We can not correlate $\nu_t^1$ and $\nu_t^2$ in the Double-Heston
Main question

Is it possible to find an AFFINE model allowing for nontrivial correlation among factors?

↓

Yes, choose a suitable State Space Domain!
Wishart multi-dim Stochastic Vol

- Extended by Da Fonseca, Grasselli and Tebaldi (2008)

\[
\frac{dS_t}{S_t} = r dt + \text{Tr} \left[ \sqrt{\Sigma_t} dZ_t \right]
\]

- \( d\Sigma_t = (\beta Q^\top Q + M \Sigma_t + \Sigma_t M^\top) dt + \sqrt{\Sigma_t} dW_t Q + Q^\top dW_t^\top \sqrt{\Sigma_t} \)
- \( Z_t = \) Matrix Brownian Motion correlated with \( W_t \) (Matrix Brownian Motion)
- \( Vol(S_t) = \text{Tr} [\Sigma_t] \) linear combination of the Wishart elements
• $d\Sigma_t = (\beta Q^\top Q + M\Sigma_t + \Sigma_t M^\top)dt + \sqrt{\Sigma_t}dW_t Q + Q^\top dW_t^\top \sqrt{\Sigma_t}$

• $\Omega \Omega^\top = \beta Q^\top Q$ with $\beta$ large enough (Gindikin’s condition)

• $M$ negative definite $\iff$ mean reverting behavior

• $\Sigma_t$ SYMMETRIC MATRIX SQUARE ROOT PROCESS ($n \times n$)

• $Q$ vol-of-vol.

• $(W_t; t \geq 0)$ is a matrix Brownian motion ($n \times n$)
Correlation in the Wishart model

- $R \in M_n$ (identified up to a rotation) completely describes the correlation structure:

$$Z_t = W_t R^\top + B_t \sqrt{I - RR^\top}$$

- This choice is compatible with affinity of the model!!

- Other (few) choices are possible but harder to manage.
• The Wishart Affine model is solvable. That is, the conditional characteristic function can be written as:

\[ \mathbb{E}_t e^{\imath \omega \log(S_{t+\tau})} = e^{\text{Tr}[A(\tau)\Sigma_t] + B(\tau) \log(S_t) + C(\tau)} \]

• \( A(\tau) \) solves a Riccati ODE that can be linearized! (Grasselli and Tebaldi 2008)
Stochastic correlation between stock returns and vol

\[ \text{Corr}_t (d\ln(S), d\text{Vol} (\ln(S))) = \rho_t = \frac{2 \text{Tr} [\Sigma_t RQ]}{\sqrt{\text{Tr} [\Sigma_t]} \sqrt{\text{Tr} [\Sigma_t Q^\top Q]}} \]

- Stochastic correlation between the stock and its volatility
- Multi-dimensional correlation/volatility SHOULD allow for more complex skew effects
Calibration single-asset stochastic volatility models:

<table>
<thead>
<tr>
<th>Model</th>
<th>error</th>
<th>( \rho_1(\rho_{11}) )</th>
<th>( \rho_2(\rho_{12}) )</th>
<th>( \rho_{21} )</th>
<th>( \rho_{22} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heston</td>
<td>0.00010773</td>
<td>-0.556</td>
<td>xxx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BiHeston</td>
<td>7.61E-05</td>
<td>-0.393</td>
<td>-0.866</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wishart</td>
<td>7.19E-05</td>
<td>-0.258</td>
<td>0.017</td>
<td>-0.343</td>
<td>-0.766</td>
</tr>
</tbody>
</table>

- the Wishart/BiHeston perform better than Heston model (not surprising!)
- the Wishart model performs slightly better than BiHeston model but numerical the cost is higher
What about adding jumps?

<table>
<thead>
<tr>
<th>Model</th>
<th>error</th>
<th>( \rho_1(\rho_{11}) )</th>
<th>( \rho_2(\rho_{12}) )</th>
<th>( \rho_{21} )</th>
<th>( \rho_{22} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heston</td>
<td>0.00010773</td>
<td>-0.556</td>
<td>xxx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BiHeston</td>
<td>7.61E-05</td>
<td>-0.393</td>
<td>-0.866</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wishart</td>
<td>7.19E-05</td>
<td>-0.258</td>
<td>0.017</td>
<td>-0.343</td>
<td>-0.766</td>
</tr>
<tr>
<td>BiBates</td>
<td>2.82E-05</td>
<td>-0.527</td>
<td>0.814</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Jumps do not change significantly the parameters of the BiHeston
- Improve the very short term fit (less than 3 weeks)
- No conflict with diffuse part
DAX calibration date: 28/08/08

![Graph showing implied volatility vs. forward moneyness for DAX with two curves labeled 'mkt 19/09/2008' and 'model 19/09/2008'.]
DAX calibration date: 28/08/08

Implied volatility

Forward moneyness

mkt 21/10/2008, model 21/10/2008
A Closer look at the $\sigma_{imp}$ for short time

Using perturbation method (vol of vol) as in Benabid, Bensusan, El Karoui (2009) we can prove that for $(T - t \sim 0)$ as a function of the forward moneyness $m_f$

$$\sigma_{imp}^2 \sim \text{Tr} [\Sigma_t] + \frac{\text{Tr} [RQ \Sigma_t]}{\text{Tr} [\Sigma_t]} m_f$$

A Double-Heston model would lead to

$$\sigma_{imp}^2 \sim v_1 + v_2 + \left( \frac{v_1 \rho_1 \sigma_1 + v_2 \rho_2 \sigma_2}{v_1 + v_2} \right) \frac{m_f}{2}$$

- $\Sigma_{12}$ controls the slope of the skew and $\Sigma_{11} + \Sigma_{22}$ controls the level of the smile (as far as $RQ$ is non diagonal).
- in the Double-Heston the factors impact both level and skew!
Conclusions

- as far as we are interest with vanilla options the BiHeston and Wishart performs equally
- but the Wishart allows a better management of the implied volatility risks
- the numerical cost of the Wishart model is much more important. How to speed up the pricing process?
- if the calibrated model will be used to price a derivative which is sensitive to the slope of the skew then the Wishart model is of interest
- the selected model depends on
  - the complexity of the smile to be calibrated
  - the sensitivity of the derivative to be priced with the calibrated model
- it raises the problem of how to aggregate the ratios from different models
Numerical results for price approximations
Dax 28/08/08 mat 0.06

forward moneyness

Impl vol

model
approx 1order
approx 2order
The Multi-asset model

How to build a multi asset framework:

- Consistent with the smile in vanilla options
- With a general correlation structure
- Analytic as much as possible
Using Heston’s model

\[ dS_t^1 = S_t^1 r\, dt + S_t^1 \sqrt{V_t^1} \, dZ_t^1 \]
\[ dV_t^1 = \kappa_1 (\theta_1 - V_t^1)\, dt + \sigma_1 \sqrt{V_t^1} \, dW_t^1 \]
\[ dS_t^2 = S_t^2 r\, dt + S_t^2 \sqrt{V_t^2} \, dZ_t^2 \]
\[ dV_t^2 = \kappa_2 (\theta_2 - V_t^2)\, dt + \sigma_2 \sqrt{V_t^2} \, dW_t^2 \]

\[ dZ_1^1 \, dZ_2^1 = 0 \iff \text{Affinity of the model} \]

\[ \downarrow \]

\[ \frac{dS_t^1}{S_t^1} \frac{dS_t^2}{S_t^2} = 0 \]
The Wishart Affine Stochastic Correlation model

Da Fonseca, Grasselli and Tebaldi (RDR-2007):

The model: \( S_t = (S_t^1, \ldots, S_t^n)^\top \) and \( \Sigma_t \in M_{(n,n)} \)

\[
dS_t = \text{diag}[S_t] \left( \mu dt + \sqrt{\Sigma_t} dZ_t \right)
\]

\[
d\Sigma_t = \left( \Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top \right) dt + \sqrt{\Sigma_t} dW_t Q + Q^\top (dW_t)^\top \sqrt{\Sigma_t}
\]

\( dZ_t \) is a vector BM (n,1) and \( dW_t \) is a matrix BM (n,n):

\[
\frac{dS^i}{S^i} \frac{dS^j}{S^j} = \Sigma^{ij} dt
\]
How to correlate $dZ$ and $dW$?

In Da Fonseca, Grasselli and Tebaldi (RDR-2007):

\[
\text{Affinity of the infinitesimal generator} \\
\dZ_t = dW_t \rho + \sqrt{1 - \rho^\top \rho} dB_t
\]

where $\rho$ is a vector $(n,1)$ and $dB$ is a vector BM$(n,1)$.

- only $n$ parameters to specify the skew
- parsimonious model
- Characteristic function has an exponential affine form, it involves the computation of the exponential of a matrix.
Pricing plain vanilla options on single assets

- In the WASC model, the single assets evolve according to a Heston-like dynamics.

- Assets’ returns and volatilities are partially correlated:

  \[ \text{Corr}_t \left( \text{Noise}(Y^1), \text{Noise}(Vol(S^1)) \right) = \frac{Q_{11}\rho_1 + Q_{21}\rho_2}{\sqrt{Q_{11}^2 + Q_{21}^2}} \]

- Vol-Of-Vol(\(S_1\)) = \(2\sqrt{Q_{11}^2 + Q_{21}^2}\)

- Skew in the implied volatility is related with the correlation, cross-asset effects appear(systematic vs specific dependence)
Calibration results in the multi-asset model

<table>
<thead>
<tr>
<th></th>
<th>error (WASC)</th>
<th>error (Heston)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dax</td>
<td>2.52E-05</td>
<td>1.105E-04</td>
</tr>
<tr>
<td>SP</td>
<td>1.39E-04</td>
<td>1.59E-04</td>
</tr>
</tbody>
</table>

• we calibrate a stochastic correlation model using only vanilla options!

• vanilla options are basket products
A closer look at $\sigma_{imp}$ for short time

We can prove

$$
\sigma_{imp}^{Dax} = \sum_{t}^{11} + (\rho_1 Q_{11} + \rho_2 Q_{21}) m_f + m_f^2 \left[ \frac{4(Q_{11}^2 + Q_{21}^2) - 7(\rho_1 Q_{11} + \rho_2 Q_{21})^2}{6\sum_{t}^{11}} \right]
$$

Recall for Heston we have

$$
\sigma_{imp}^2 = v + \sigma^2 \frac{\rho^2}{2} m_f + \frac{\sigma^2 m_f^2 (4 - 7 \rho^2)}{24v}
$$

- the expansions for the smile are similar
- the same problem as for Heston: we have a concave relation! those asymptotics can not be used to build a starting point for the calibration!
- at first order $\rho$ and $\sigma$ appear as a product $\rightarrow$ identification problem (same for Wasc)
- this aggregation of parameters allows to understand the parameter values
A competitor

\[ ds_1(t) = s_1(t)(\sqrt{v_1(t)}dw_1(t) + \sqrt{v_0(t)}dw_0(t)) \]
\[ ds_2(t) = s_2(t)(\sqrt{v_2(t)}dw_2(t) + \sqrt{v_0(t)}dw_0(t)) \]
\[ dv_1(t) = \kappa_1(\theta_1 - v_1(t))dt + \sigma_1\sqrt{v_1(t)}(\rho_1dw_1(t) + \sqrt{1 - \rho_1^2}d\tilde{w}_1(t)) \]
\[ dv_2(t) = \kappa_2(\theta_2 - v_2(t))dt + \sigma_2\sqrt{v_2(t)}(\rho_2dw_2(t) + \sqrt{1 - \rho_2^2}d\tilde{w}_2(t)) \]
\[ dv_0(t) = \kappa_0(\theta_0 - v_0(t))dt + \sigma_0\sqrt{v_0(t)}(\rho_0dw_0(t) + \sqrt{1 - \rho_0^2}d\tilde{w}_0(t)) \]

- this model allows stochastic correlation and is more tractable (the CF is computationally less complicated than the Wasc).
- in this model we have a factor model for the covariance matrix whereas for the Wasc model the covariance matrix is the factor, might be of interest when dealing with estimation
Conclusions

• we build a model which is tractable

• this model allows for stochastic volatilities and stochastic correlation

• we provide some results on calibration using single underlying options with the consequence that vanilla options are basket products.

some open problems

• building estimation strategy, for the Wasc see Da Fonseca, Grasselli, Ielpo (2009).

• how to increase the dimension of the model and still be able to estimate it

• how to aggregate the risks of different models: Heston, BiHeston, Wishart, Wasc and others...
Thanks for your attention!