Pricing Kernels and Option Valuation

Peter Christoffersen
Rotman School of Management, University of Toronto,
Copenhagen Business School, and
CREATES, University of Aarhus

1st Lecture on Thursday
Background

• Financial econometricians often estimate time-series models with dynamic volatility using underlying returns only. This yields the physical (P) distribution.

• Financial practitioners often rely on frequent recalibration of option valuation models using option prices only. This yields the risk-neutral (Q) distribution.

• Bates (HS, 1996) “The central empirical issue in option research is whether the distributions implicit in option prices are consistent with the time series properties of the underlying asset prices.”
**P, Q and M (SDF)**

- The $P$ and $Q$ distributions for risky asset with price $S$ are linked by the pricing kernel $M$. If we assume a representative agent with time sep utility then

\[
S_0 = E^P_0 [M_T S_T] = E^P_0 \left[ \beta \frac{U'(S_T)}{U'(S_0)} S_T \right] = \exp(-rT) E^Q_0 [S_T]
\]

- So that

\[
f^Q(S_T) \propto \exp(rT) \beta \frac{U'(S_T)}{U'(S_0)} f^P(S_T)
\]

- Note: $S_T$ will be $C(S_T)$ in option valuation.

- See Hansen and Renault (EQF, 2010).
The Pricing Kernel

• Investigating “...whether the distributions implicit in option prices are consistent with the time series properties of the underlying asset prices...” seems to require a specification search on the pricing kernel.

• Compared with the enormous literature on specifying $P$, the derivatives literature on specifying $M$ is small.

• The literature seems to be driven by wanting $Q$ to be similar to $P$ and (relatedly) wanting $Q$ to be such that a (quasi) closed-form expression can be had for the (European) option price.
Overview

• 1) What are the stylized facts in the relative distributions of index return and option prices that the pricing kernel needs to capture?
• 2) What is the pricing kernel implied in the stochastic volatility model of Heston (1993)?
• 3) Does the Heston (1993) pricing kernel capture (qualitatively) the stylized facts in 1)?
• 4) Can the Heston and Nandi (2000) GARCH model be augmented with a pricing kernel similar to that in Heston (1993)?
• 5) GARCH versus SV. Affine versus Non-affine.
Table 1: Returns and Options Data

Panel A: Return Characteristics (Annualized)

<table>
<thead>
<tr>
<th></th>
<th>1990-2010</th>
<th>1996-2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>7.90%</td>
<td>5.46%</td>
</tr>
<tr>
<td>St. Deviation</td>
<td>18.28%</td>
<td>20.66%</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.201</td>
<td>-0.180</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>12.258</td>
<td>10.981</td>
</tr>
</tbody>
</table>

Panel B. Option Data by Moneyness

<table>
<thead>
<tr>
<th></th>
<th>F/X ≤ .96</th>
<th>.96 &lt; F/X ≤ .98</th>
<th>.98 &lt; F/X ≤ 1.02</th>
<th>1.02 &lt; F/X ≤ 1.04</th>
<th>1.04 &lt; F/X ≤ 1.06</th>
<th>F/X &gt; 1.06</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Contracts</td>
<td>3,032</td>
<td>1,778</td>
<td>6,287</td>
<td>2,315</td>
<td>1,744</td>
<td>6,553</td>
<td>21,709</td>
</tr>
<tr>
<td>Average IV</td>
<td>20%</td>
<td>19%</td>
<td>19%</td>
<td>21%</td>
<td>22%</td>
<td>26%</td>
<td>22%</td>
</tr>
<tr>
<td>Average Price</td>
<td>27.90</td>
<td>32.63</td>
<td>40.87</td>
<td>41.25</td>
<td>36.97</td>
<td>25.86</td>
<td>33.58</td>
</tr>
<tr>
<td>Average Spread</td>
<td>1.66</td>
<td>1.74</td>
<td>1.92</td>
<td>1.76</td>
<td>1.65</td>
<td>1.49</td>
<td>1.70</td>
</tr>
</tbody>
</table>

Panel C. Option Data by Maturity

<table>
<thead>
<tr>
<th></th>
<th>DTM ≤ 30</th>
<th>30 &lt; DTM ≤ 60</th>
<th>60 &lt; DTM ≤ 90</th>
<th>90 &lt; DTM ≤ 120</th>
<th>120 &lt; DTM ≤ 180</th>
<th>DTM &gt; 180</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Contracts</td>
<td>968</td>
<td>4,824</td>
<td>3,710</td>
<td>1,769</td>
<td>3,074</td>
<td>7,364</td>
<td>21,709</td>
</tr>
<tr>
<td>Average IV</td>
<td>20.91%</td>
<td>20.93%</td>
<td>21.38%</td>
<td>23.35%</td>
<td>22.21%</td>
<td>21.84%</td>
<td>21.69%</td>
</tr>
<tr>
<td>Average Price</td>
<td>12.92</td>
<td>19.54</td>
<td>26.97</td>
<td>32.54</td>
<td>33.76</td>
<td>49.00</td>
<td>33.58</td>
</tr>
<tr>
<td>Average Spread</td>
<td>0.89</td>
<td>1.29</td>
<td>1.64</td>
<td>1.84</td>
<td>1.76</td>
<td>2.05</td>
<td>1.70</td>
</tr>
</tbody>
</table>
1) Three Stylized Facts

- Relative Fat Tails in Options versus Returns

- Implied Volatility Exceeds Physical Volatility

- Long-term Implied Volatility Overreaction
  - Stein (1989), Posheshman (2001)
Model-free Physical and Risk Neutral Conditional Densities (I)

• Following Ait-Sahalia and Lo (1998) we first fit options each Wednesday using an ad-hoc Black-Scholes approach relying on a second-order polynomial in strike and maturity for implied volatility:

\[ \hat{C}(S(t), X, \tau, r) = C_{BS}(S(t), X, \tau, r; \hat{\sigma}(S(t), X, \tau)) . \]

• Then we compute the risk neutral density by taking the second derivative w.r.t. the strike price.

\[ f^*_t(S(T)) = \exp(r) \left[ \frac{\partial^2 \hat{C}(S(t), X, \tau, r)}{\partial X^2} \right]_{X=S(T)} \]
Model-free Physical and Risk Neutral Conditional Densities (II)

• Switching from prices to returns we get

\[ \hat{f}_t^*(R(t,T)) = \frac{\partial}{\partial u} \Pr \left( \ln \left( \frac{S(T)}{S(t)} \right) \leq u \right) = S(t) \exp(u) \hat{f}_t^*(S(t)\exp(u)) \]

• Challenge: How to get a model-free view of the physical conditional density of returns? We compute it using a Gaussian kernel on monthly returns normalized by GARCH.

• The plot below shows the physical histogram using current GARCH (dots) and the risk-neutral (solid) density on the first Wednesday of each year.
Physical histogram using GARCH (dots) and the risk-neutral (solid) density on the first Wednesday of each year.

One-month horizon.
Log Ratio of Risk-Neutral and Physical One-Month Densities using all Wednesdays each year.

The U-shape patterns is remarkably robust across the 52 weeks in the year.

An adequate pricing kernel needs to have this feature.
“Realized” one-month volatility is simply the standard deviation of a forward-looking one-month window of daily returns.
Straddle Strategy

• On the third Wednesday of every month we sell a position of SPX calls and SPX puts using the closest to at-the-money, one-month contracts.
• We scale the position so that the premiums collected correspond to 10% of current capital.
• The contracts are held until maturity.
• The index below is computed by starting with $100 in capital on 1/1/1996 and then cumulating all the cash flows from the option sales. Cash earns the risk-free rate.
Mean Return = 1.32% p.m.
Volatility = 6.77%
T-stat = 2.51
Skewness = -1.39
X-Kurtosis = 3.23
Stein’s (1989) Overreaction Test

• Stein’s test of overreaction:
  \[ E_t \left[ (IV_{t+1(M)}^{ST} - IV_t^{ST}) - 2(IV_t^{LT} - IV_t^{ST}) \right] = 0, \]

• Test this moment using weekly regression
  \[ (IV_{t+4}^{1M} - IV_t^{1M}) - 2(IV_t^{2M} - IV_t^{1M}) = a_0 + a_1 IV_t^{1M} + e_{t+4}, \]

• When IV\(^{1M}\) is high then IV\(^{2M}\) is “too” high yielding a negative \(a_1\).

• If \(a_1\) is negative then IV\(^{2M}\) “overreacts” to changes in IV\(^{1M}\).

• Stein’s sample is small. Are findings robust?
<table>
<thead>
<tr>
<th>Sample Period</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>-0.237</td>
<td>0.0093</td>
<td>-25.62</td>
</tr>
<tr>
<td>1996</td>
<td>-0.326</td>
<td>0.0255</td>
<td>-12.76</td>
</tr>
<tr>
<td>1997</td>
<td>-0.179</td>
<td>0.0266</td>
<td>-6.75</td>
</tr>
<tr>
<td>1998</td>
<td>-0.272</td>
<td>0.0387</td>
<td>-7.04</td>
</tr>
<tr>
<td>1999</td>
<td>-0.276</td>
<td>0.0157</td>
<td>-17.58</td>
</tr>
<tr>
<td>2000</td>
<td>-0.21</td>
<td>0.0138</td>
<td>-15.25</td>
</tr>
<tr>
<td>2001</td>
<td>-0.225</td>
<td>0.0233</td>
<td>-9.66</td>
</tr>
<tr>
<td>2002</td>
<td>-0.184</td>
<td>0.0273</td>
<td>-6.74</td>
</tr>
<tr>
<td>2003</td>
<td>-0.234</td>
<td>0.02</td>
<td>-11.68</td>
</tr>
<tr>
<td>2004</td>
<td>-0.361</td>
<td>0.0194</td>
<td>-18.59</td>
</tr>
<tr>
<td>2005</td>
<td>-0.439</td>
<td>0.0246</td>
<td>-17.89</td>
</tr>
<tr>
<td>2006</td>
<td>-0.538</td>
<td>0.0234</td>
<td>-22.99</td>
</tr>
<tr>
<td>2007</td>
<td>-0.388</td>
<td>0.0386</td>
<td>-10.05</td>
</tr>
<tr>
<td>2008</td>
<td>-0.16</td>
<td>0.0479</td>
<td>-3.34</td>
</tr>
<tr>
<td>2009</td>
<td>-0.228</td>
<td>0.0156</td>
<td>-14.59</td>
</tr>
</tbody>
</table>
2) The Heston Model (1993)

• The physical measure

\[ dS(t) = (r + \mu v(t))S(t)\,dt + \sqrt{v(t)}S(t)\,dz_1(t), \]
\[ dv(t) = \kappa(\theta - v(t))\,dt + \sigma\sqrt{v(t)}\left(\rho dz_1(t) + \sqrt{1 - \rho^2}\,dz_2(t)\right), \]

• The conventionally assumed risk-neutral measure

\[ dS(t) = rS(t)\,dt + \sqrt{v(t)}S(t)\,dz_1^*(t), \]
\[ dv(t) = (\kappa(\theta - v(t)) - \lambda v(t))\,dt + \sigma\sqrt{v(t)}(\rho dz_1^*(t) + \sqrt{1 - \rho^2}\,dz_2^*(t)), \]

• Empirically \( \mu > 0 \) and \( \lambda < 0. \)
Pricing Kernel (SDF) Proposition

• There is a unique stochastic discount factor reconciling the physical and risk neutral

\[ M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^\phi \exp \left( \delta t + \eta \int_0^t \nu(s) ds + \xi (\nu(t) - \nu(0)) \right) \]

• Where

\[
\begin{align*}
\delta &= - (1 + \phi) r - \xi \kappa \theta, \\
\eta &= - \phi \mu + \frac{1}{2} \phi + \xi \kappa - \frac{1}{2} \left( \phi^2 + 2 \phi \xi \sigma \rho + \xi^2 \sigma^2 \right), \\
\phi &= - \mu - \xi \sigma \rho, \\
\xi &= \frac{\mu \sigma \rho - \lambda}{\sigma^2 (1 - \rho^2)}.
\end{align*}
\]
Sketch of Proof

• Let $B(t)$ be a risk-free bond. Requiring that $B(t)M(t)$ is a martingale yields the restrictions on $\delta$ and $\eta$.

• Requiring that $S(t)M(t)$ is a martingale yields the restriction on $\phi$.

• Let $U(t)$ be an asset that depends on $S(t)$ and $v(t)$. Requiring that $U(t)M(t)$ is a martingale yields the restriction on $\xi$. 
Implications

• The proposition can be rewritten to get the equity and volatility risk prices as functions of preference parameters:

\[ \begin{align*}
\mu &= -\phi - \xi \sigma, \rho \\
\lambda &= \rho \sigma - \mu (1-\rho^2) \sigma^2 \xi = \rho \sigma - \phi \sigma^2 \xi.
\end{align*} \] (10)

– If marginal utility is decreasing in returns: \( \phi < 0 \).
– If hedging needs increase when volatility is high then SDF should be increasing in volatility: \( \xi > 0 \).
– Empirically, \( \rho < 0 \) (and \( \sigma > 0 \)) so then \( \mu > 0 \) and \( \lambda < 0 \).

• Note: Variance preference \( \xi > 0 \) not needed for \( \mu > 0 \) and \( \lambda < 0 \).
The Pricing Kernel

• When variance is constant the pricing kernel corresponds to the power utility setup in Rubinstein (1976) and Brennan (1979) with risk aversion $\phi$.

• Allowing for stochastic variance and variance preference, $\xi$, could be justified by $v(t)$ governing the variance of aggregate production in a CIR (1985) model with non-log utility.

• Benzoni, Collin-Dufresne, and Goldstein (2009): Uncertainty directly affects preferences

• Bakshi, Madan and Panayotov (2010): Short-sale constraints.
3) Does the Pricing Kernel Capture the Stylized Facts?

• \( \xi > 0 \) not needed for \( \lambda < 0 \) and \( \mu > 0 \) but if \( \xi > 0 \) then U-shaped pricing kernel.

• Expected future risk-neutral variance:
  \[
  E_t^*(v(t + \Delta)) = \exp^{-\kappa^*\Delta} v(t) + (1 - \exp^{-\kappa^*\Delta})\theta^*.
  \]

• Where \( \theta^* = \kappa \theta / (\kappa + \lambda) \) and \( \kappa^* = \kappa + \lambda \).

• Thus with \( \lambda < 0 \), we get that risk neutral volatility exceeds physical on average.

• Also option market “overreaction” due to greater sensitivity of \( E_t^*(v(t+\Delta)) \) to \( v(t) \).
4) A Discrete Time GARCH Approach

- Estimation of the continuous time SV model jointly on option and returns data can be cumbersome due to the latent volatility factor.
- Is it possible to develop a version of the discrete time Heston and Nandi (2000) affine GARCH model (with quasi closed-form option prices) that captures the stylized facts?
- The existing version of Heston and Nandi (2000) can fit returns well and options well but not with the same parameter values. A nontrivial wedge is needed between its physical and risk-neutral distributions.
Augmented Heston-Nandi (2000)

Heston and Nandi assume the discrete time return process

\[
\ln(S(t)) = \ln(S(t-1)) + r + (\mu - \frac{1}{2})h(t) + \sqrt{h(t)}\epsilon(t),
\]

\[
h(t) = \omega + \beta h(t-1) + \alpha (\epsilon(t-1) - \gamma \sqrt{h(t-1)})^2,
\]

Now augment: If we assume the SDF

\[
M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^{\phi} \exp \left( \delta t + \eta \sum_{s=1}^{t} h(s) + \xi (h(t+1) - h(1)) \right)
\]

Has the flavor of the pricing kernel implicit in Heston (1993). Use this to “risk-neutralize”
GARCH Q Measure

• New risk neutral measure

\[
\ln(S(t)) = \ln(S(t-1)) + r - \frac{1}{2} h^*(t) + \sqrt{h^*(t)} z^*(t),
\]

\[
h^*(t) = \omega^* + \beta h^*(t-1) + \alpha^*(z^*(t-1) - \gamma^* \sqrt{h^*(t-1)})^2,
\]

• Thus the risk-neutral process takes the same GARCH form but with new parameters:

\[
h^*(t) = h(t) / (1 - 2\alpha \xi),
\]

\[
\omega^* = \omega / (1 - 2\alpha \xi),
\]

\[
\alpha^* = \alpha / (1 - 2\alpha \xi)^2,
\]

\[
\gamma^* = \gamma - \phi.
\]

Note: In Heston and Nandi (2000) \(\xi=0\) and only \(\gamma\) changes across measures.
Properties of $Q$ versus $P$

- $h^*(t) > h(t)$,
- Risk-neutral variance is more persistent, 
  \[ E_{t-1}^*(h^*(t+1)) = (\beta + \alpha^*\gamma^2)h^*(t) + (1 - \beta - \alpha^*\gamma^2)E^*(h^*(t)), \]
- Risk-neutral expected variance > expected variance, 
  \[ E^*(h^*(t)) = (\omega^* + \alpha^*)/(1 - \beta - \alpha^*\gamma^2) \]
- Risk-neutral variance of variance is higher. 
  \[ Var_{t-1}^*(h^*(t+1)) = 2\alpha^2 + 4\alpha^2\gamma^2h^*(t). \]
The pricing kernel is U-shaped when $\xi > 0$

- Recall that we assumed

$$M(t) = M(0) \left( \frac{S(t)}{S(0)} \right)^\phi \exp \left( \delta t + \eta \sum_{s=1}^{t} h(s) + \xi (h(t+1) - h(1)) \right)$$

- Substituting in return and variance dynamics gives one-day log pricing kernel:

$$\ln \left( \frac{M(t)}{M(t-1)} \right) = \frac{\xi \alpha}{h(t)} (R(t) - r)^2 - \mu (R(t) - r) + \\
\left( \eta + \xi (\beta - 1) + \xi \alpha \left( \mu - \frac{1}{2} + \gamma \right)^2 \right) h(t) + \delta + \xi \omega + \phi r$$
The Heston-Nandi is Applicable

• The new SDF yields a risk neutral process of the same form (but with different parameters and variance path) as in Heston and Nandi (2000).

• We can use the HN’s characteristic function \( g \), but with the new parameter mapping. Call price is:

\[
C(S(t), h^*(t + 1), X, T) = S(t) P_1(t) - X \exp(-r(T - t)) P_2(t)
\]

\[
P_1(t) = \left( \frac{1}{2} + \frac{\exp(-r(T - t))}{\pi} \int_0^\infty \text{Re} \left[ \frac{X^{-i\varphi} g_{t,T}^*(i\varphi + 1)}{i\varphi S(t)} \right] d\varphi \right)
\]

\[
P_2(t) = \left( \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \text{Re} \left[ \frac{X^{-i\varphi} g_{t,T}^*(i\varphi)}{i\varphi} \right] d\varphi \right).
\]
Estimation of the New Model

• We investigate the relationship between the physical and risk-neutral measures.
• We thus want to fit the model using both option and underlying return data using a joint log likelihood:

\[
\max_{\Theta, \Theta^*} \ln L^R + \ln L^O, \\
\ln L^R \propto -\frac{1}{2} \sum_{t=1}^{T} \{ \ln (h(t)) + (R(t) - r - \mu h(t))^2 / h(t) \}.
\]

\[
\ln L^O \propto -\frac{1}{2} \sum_{i=1}^{N} \{ \ln (s_e^2) + \varepsilon_i^2 / s_e^2 \}. \\
\varepsilon_i = \left( C_i^{Mkt} - C_i^{Mod} \right) / BSV_i^{Mkt},
\]
Table 3: Parameter Estimation and Model Fit. Joint Estimation using Returns and Options

<table>
<thead>
<tr>
<th>Physical Parameters</th>
<th>Joint Estimation</th>
<th>Return-Based Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Premia</td>
<td>Equity Premium Only</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.410E-06</td>
<td>1.410E-06</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.755</td>
<td>0.755</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>411.19</td>
<td>409.63</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0</td>
<td>1.594</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Risk Neutral Parameters</th>
<th>Joint Estimation</th>
<th>Option-Based Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{(1-2\alpha\xi)}{}^{-1}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\omega^*$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\alpha^*$</td>
<td>1.410E-06</td>
<td>1.410E-06</td>
</tr>
<tr>
<td>$\beta^*$</td>
<td>0.755</td>
<td>0.755</td>
</tr>
<tr>
<td>$\gamma^*$</td>
<td>411.19</td>
<td>411.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total Likelihood</th>
<th>From returns</th>
<th>From options</th>
</tr>
</thead>
<tbody>
<tr>
<td>56,403.5</td>
<td>17,673.7</td>
<td>38,729.7</td>
</tr>
<tr>
<td>56,410.7</td>
<td>17,681.0</td>
<td>38,729.8</td>
</tr>
<tr>
<td>56,480.9</td>
<td>17,749.2</td>
<td>38,731.6</td>
</tr>
<tr>
<td>56,578.5</td>
<td>17,846.9</td>
<td>38,731.7</td>
</tr>
</tbody>
</table>
Does the New Model Capture the U-Shape?

• Take the joint estimation parameters from Table 3.
• Consider 1, 30, 90, 180-days to maturity.
• Using the conditional characteristic function from Heston-Nandi, compute the model-based p.d.f. under P and Q.
• Plot the log difference to see if we can match Figure 2 when the variance risk premium is present.
Model-implied log density ratios with a variance risk premium (solid lines) and without (dashes lines).
Take the fitted model prices and run them through the procedure used to compute the log pricing kernel from market prices earlier.

One month horizon
Construct Straddle returns from model prices
Mean = 1.65% p.m.
Volatility = 6.39%
T-stat = 3.31
Skewness = -1.09
X-Kurtosis = 1.28
Daily Spot Vol annualized in percent.

Key model feature:

The Spot Vols vary across the P and Q measures.
Wedge between $h^*(t)$ and $h(t)$

- The wedge between $h^*(t)$ and $h(t)$ is a key feature of the new GARCH model that distinguishes it from the SV model.
- In the Heston SV model we have $V^*(t)=V(t)$
- This has important implications for the variance term structures in the two models.
- Calibrate SV parameters by matching GARCH persistence as well as unconditional variance under $P$ and $Q$. 
Volatility Term Structures

Integrated Variance under P in SV:

\[
VAR_0(T) \equiv E_0 \left[ \int_0^T \nu(t) \, dt \right] = \theta T + (\nu_0 - \theta) \frac{(1 - e^{-\kappa T})}{\kappa}
\]

Volatility Term Structure under Q in SV:

\[
VTSS_{SV}^* = \sqrt{\theta \frac{\kappa}{\kappa + \lambda} + \left( \frac{\nu_0 - \theta \frac{\kappa}{\kappa + \lambda}}{T} \right) \frac{(1 - e^{-(\kappa + \lambda)T})}{(\kappa + \lambda)}}
\]

Volatility Term Structure under Q in New GARCH Model:

\[
VTSS^* = \sqrt{E^*(h^*(t)) + \left( \frac{h^*(1) - E^*(h^*(t))}{T} \right) \frac{1 - (\beta + \alpha^* \gamma^2)^T}{1 - (\beta + \alpha^* \gamma^2)}}
\]
Volatility Term Structures: New GARCH and SV
5) GARCH versus SV. 
Affine versus Non-affine Variance.

• GARCH estimation on returns only is very easy. Standard MLE even if numerical optimization is required. Estimation on returns and options is relatively easy if affine GARCH is used.

• SV estimation on returns is more complicated due to volatility being a latent factor.

• SV estimation on options can be done relatively easily if affine SV is used.

• Affine versus non-affine variance dynamics. Empirical fit versus Monte Carlo pricing.
Possible GARCH Extensions

- Single component affine normal GARCH model clearly has limitations...
- Multiple volatility components
  - Christoffersen, Jacobs, Ornthanalai, and Wang (JFE, 2008)
- Non-normal return innovations
  - Christoffersen, Heston and Jacobs (JE, 2006)
- GARCH-“Jumps” with time-varying intensity
  - Christoffersen, Jacobs and Ornthanalai (JFE, 2013)
- Non-affine (and Non-normal) models
  - Christoffersen, Elkamhi, Feunou, and Jacobs (RFS, 2010)
SV Estimation on Returns

• Necessary Steps
  – Discretize the model
  – Establish filter for latent volatility
  – Decide on estimation criterion for parameters


• The SIR particle filter provides a convenient and accurate way of extracting the latent stochastic volatility for a given set of model parameters.

• Johannes, Polson, Stroud (RFS, 2009), CJM (RFS, 2010),
SIR Particle Filtering

• Discretize the model.

\[
\ln(S_{t+1}) = \ln(S_t) + \left(\mu - \frac{1}{2}V_t\right) + \sqrt{V_t}z_{t+1}
\]

\[
V_{t+1} = V_t + \kappa V_t^\alpha (\theta - V_t) + \sigma V_t^b w_{t+1}
\]

• Generate initial set of volatility particles for date 0 \((V_0^j)\).

• Update variances by simulating model dynamic to get \((V_t^j)\).
SIR Particle Filtering

• Compute particle weights \((W_t^j)\) based on model likelihood.

\[
W_{t+1}^j = \frac{1}{\sqrt{V_{t+1}^j}} \exp \left( -\frac{1}{2} \left( \ln \left( \frac{s_{t+2}}{s_{t+1}} \right) - \left( \mu - \frac{1}{2} V_{t+1}^j \right) \right)^2 \right)
\]

• Now do an importance resampling of particles. This is done using an integer variable that is constructed using the empirical CDF of the V particles. Each particle is resampled either 0, 1 or multiple times. Weights are recalculated based on the integer variable.
Filtered Volatility

- Once the particles and weights have been computed on all the dates we are ready to compute the filtered volatility simply as the weighted average of the particles

\[
\bar{V}_{t+1} = \sum_{j=1}^{N} W_{t+1}^j V_{t+1}^j
\]
MLIS on Returns

As the particle weights \((W_t^j)\) are based on the model likelihood, the parameters of the model can be easily estimated via Maximum Likelihood Importance Sampling (MLIS):

\[
MLIS (\mu, \kappa, \theta, \rho, \sigma) = \sum_{t=1}^{T} \ln \left( \frac{1}{N} \sum_{j=1}^{N} W_t^j \right)
\]
NLSIS Estimation on Options

- Using the auxiliary particle filter again we now minimize the **Nonlinear Least Squares Importance Sampling** criterion

\[
IVMSE(\kappa, \theta, \rho, \sigma, \lambda) = \frac{1}{NT} \sum_{t,i} \left( IV_{i,t} - BS_i^{-1} \{ C_i(\tilde{V}_t) \} \right)^2
\]

- This is done on a large sample of call option contracts observed across quote dates, strike prices and maturity dates.
- IV grid is used. Filtering V on returns as before.
Model Free Realized Volatility versus Affine SV Model. Motivation for Tomorrow’s Topic
Summary

• 1) Stylized facts:
  – Option distribution tails are fatter than return tails.
  – Average IV is higher than average physical vol.
  – Long-term IV “overreacts” to changes in spot vol.

• 2) The pricing kernel implicit in Heston (1993) can capture these features—at least qualitatively.

• 3) The Heston and Nandi (2000) model can be augmented with a pricing kernel similar to Heston (1993).

• 4) Many extensions and alternatives are possible. Non-affine models should be considered.