Efficient Numerical PDE Methods to Solve Calibration and Pricing Problems in Local Stochastic Volatility Models

Artur Sepp
Bank of America Merrill Lynch, London
artur.sepp@baml.com

Global Derivatives Trading & Risk Management 2011
Paris
April 12-15, 2011

Plan of the presentation

1) Volatility modelling

2) Local stochastic volatility models: stochastic volatility, jumps, local volatility

3) Calibration of parametric local volatility models using partial differential equation (PDE) methods

4) Calibration of non-parametric local volatility volatility models with jumps and stochastic volatility using PDE methods

5) Numerical methods for PDEs

6) Illustrations using SPX and VIX data

References

Some theoretical and practical details for my presentation can be found in:

1) Sepp, A. (2012) An Approximate Distribution of Delta-Hedging Errors in a Jump-Diffusion Model with Discrete Trading and Transaction Costs, Quantitative Finance, 12(7), 1119-1141
`http://ssrn.com/abstract=1360472`

2) Sepp, A. (2008) VIX Option Pricing in a Jump-Diffusion Model, Risk Magazine April, 84-89
`http://ssrn.com/abstract=1412339`

3) Sepp, A. (2008) Pricing Options on Realized Variance in the Heston Model with Jumps in Returns and Volatility, Journal of Computational Finance, Vol. 11, No. 4, pp. 33-70
`http://ssrn.com/abstract=1408005`

Literature

Local volatility: Dupire (1994)

Local volatility with jumps: Andersen-Andreasen (2000)

Local stochastic volatility with jumps: Lipton (2002)

Volatility modelling I

What is important for a competitive pricing and hedging volatility model?

1) Consistency with the observed market dynamics implies stable model parameters and hedges

2) Consistency with vanilla option prices ensures that the model fits the risk-neutral distribution implied by these prices

3) Consistency with market prices of liquid exotic options

Volatility modelling II. Volatility models

1) **Non-parametric local volatility models** (Dupire (1994), Derman-Kani (1994), Rubinstein (1994))
"+" LV models are consistent with vanilla prices by construction
"-" but they tend to be poor in replicating the market dynamics of spot and volatility (implied volatility tends to move too much given a change in the spot, no mean-reversion effect)
"-" especially, it is impossible to tune-up the volatility of the implied volatility as there is simply no parameter for that!

2) **Parametric stochastic volatility models** (Heston (1993))
"+" SV models tend to be more aligned with the market dynamics
"+" introduce the term-structure (by mean-reversion parameters) and the volatility of the variance (by vol-of-vol parameters)
"-" need a least-squares calibration to today's option prices
"-" unfortunately, any change in any of the parameters of the mean-reversion or the vol-of-vol requires re-calibration of other parameters

3) **Stochastic local volatility models** (JP Morgan (1999), Blacher (2001), Lipton (2002), Ren-Madan-Qian (2007))
LSV models aim to include "+"'s and cross "-"'s of first two models

Volatility modelling III. LSV model specification

1) Global factors - specify the dynamics for the instantaneous volatility, include jumps or default risk if this is relevant for product risk

(Guess) Estimate or calibrate model parameters for the dynamics of the instantaneous volatility and jumps using either historical or market data

Parameters are updated infrequently

2) Local factors - specify local factors for either parametric or non-parametric local volatility

Parameters of local factors are updated frequently (on the run) to fit the risk-neutral distributions implied by market prices of vanilla options

3) Mixing weight - specify mixing weight between stochastic and local volatilities, adjust vol-of-vol and correlation accordingly

LSV model dynamics. We consider a specific version of the LSV dynamics with jumps under pricing measure:

$$
\begin{aligned}
& dS(t) = \mu(t)S(t_-)dt + \sigma_{(\mathrm{loc,svj})}(t_-, S(t_-))\vartheta(t_-, Y(t_-))S(t_-)dW^{(1)}(t) \\
& + \Big((e^{-\nu} - 1)dN(t) + \lambda\nu dt\Big)\, S(t_-),\ S(0) = S \\
& dY(t) = -\kappa Y(t)dt + \epsilon dW^{(2)}(t) + \eta dN(t),\ Y(0) = 0
\end{aligned}
\tag{1}
$$

$S(t)$ is the underlying and $Y(t)$ is the factor for instantaneous stochastic volatility with $dW^{(1)}(t)dW^{(2)}(t) = \rho dt$

$\sigma_{(\mathrm{loc,svj})}(t, S(t))$ is the **local volatility** (to be considered later)

$\vartheta(t, Y(t))$ is **volatility mapping** with $\mathbb{V}[Y(t)]$ being variance of $Y(t)$:

$$\vartheta(t, Y(t)) = e^{Y(t) - \mathbb{V}[Y(t)]}$$

For $\rho = 0$, $\mathbb{E}\left[\vartheta^2(t, Y(t)) \,|\, Y(0) = 0\right] = 1$, so that $\vartheta(t, Y(t))$ introduces "**volatility-of-volatility**" effect without affecting the local volatility close to the spot

Stochastic volatility factor I

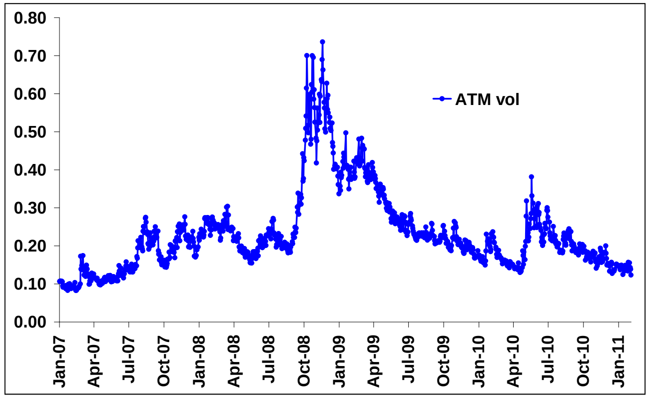
Q: Can we observe the dynamics of stochastic volatility factor $Y(t)$?
A: Consider the model implied ATM volatility squared:

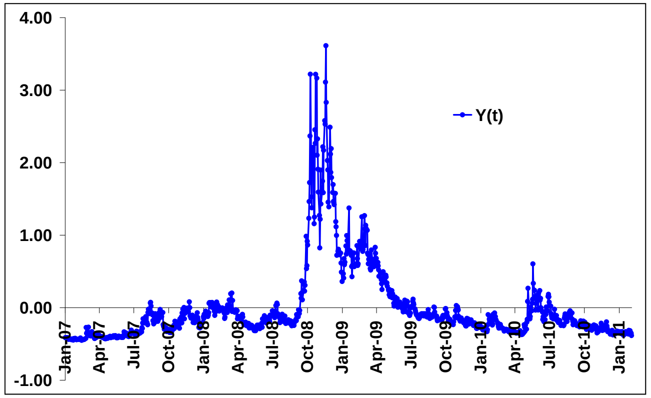
$$V(t) = \sigma_0^2 e^{2Y(t)-2\mathbb{V}[Y(t)]} \approx \sigma_0^2(1+2Y(t))$$

Thus, given daily time series of implied ATM volatility $\sigma_{\text{ATM}}(t_n)$:

$$Y(t_n) = \frac{1}{2}\left(\frac{\sigma^2_{\text{ATM}}(t_n)}{\overline{\sigma^2_{\text{ATM}}}} - 1\right)$$

where $\overline{\sigma^2_{\text{ATM}}}$ is the average ATM volatility during the period





Left: One-month ATM implied volatility for SPX; **Right:** Y(t)

Stochastic volatility factor II. Estimation

To obtain historical estimates for the mean-reversion and vol-of-vol, apply regression model:

$$Y(t_n) = \beta Y(t_{n-1}) + \theta \varsigma_n$$

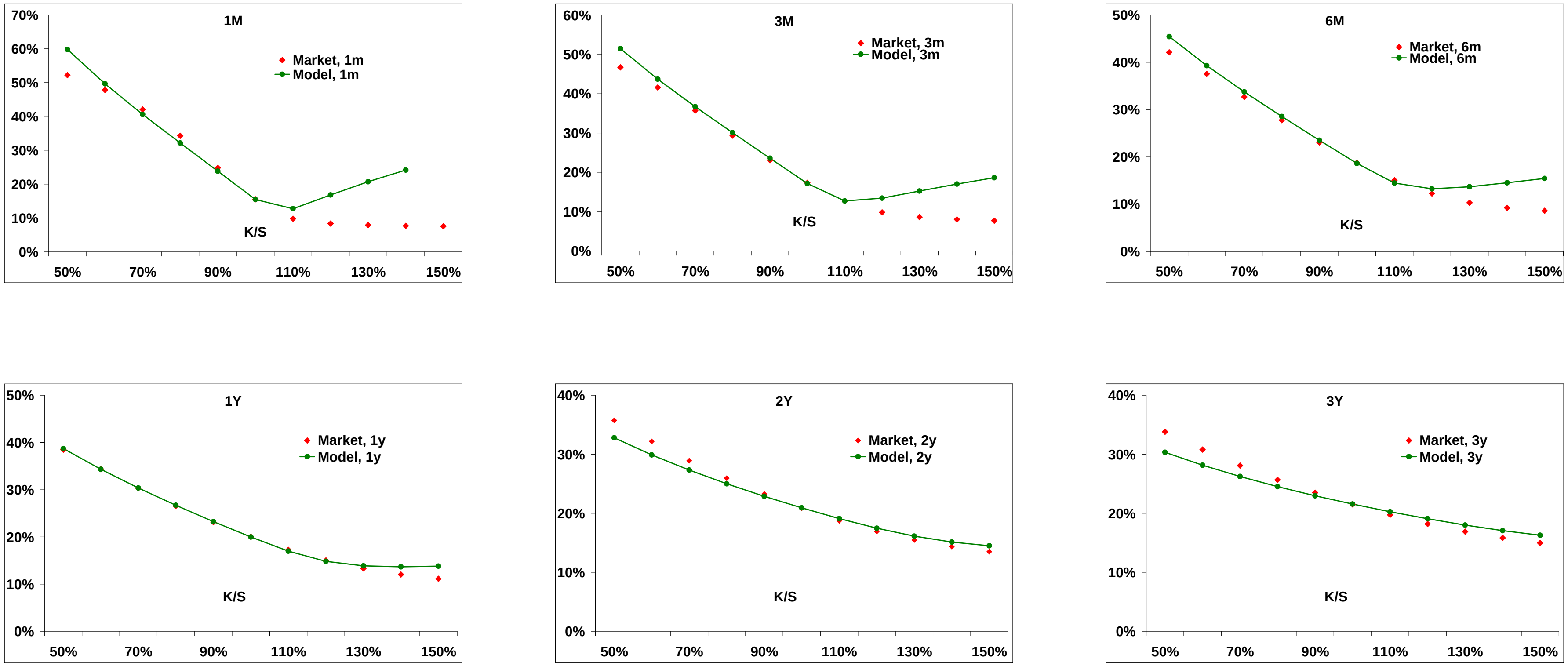
where ς_n is iid normals

Thus:
$\kappa = -252 \ln(\beta)/2$
$\epsilon = \sqrt{252\theta(2\kappa)/(1 - e^{-2\kappa/252})}$

For SPX (using data for last 4 years):
$R^2 = 74\%$
$\kappa = 4.49$
$\epsilon = 2.30$

Stochastic volatility factor III. Fit to implied SPX volatilities (23-March-1011) using flat local volatility (log-normal SV model) and historical parameters with $\rho = -0.8$

Conclusion: pure stochastic volatility model is adequate to describe longer terms skew but need local factors to "correct" mis-pricings



Jumps I

We assume that jumps in $S(t)$ and $Y(t)$ are **simultaneous and discrete** with magnitudes $-\nu < 0$ and $\eta > 0$

Q: Why do we need jumps in addition to local stochastic volatility?
A: Market prices of exotics (eg cliques) imply very steep forward skews so the value of **vol-of-vol parameter must be very large** (I will show an example in a bit)

However, local stochastic volatility with a large vol-of-vol parameter cannot be calibrated consistently to given local volatility

Calibration to both exotics and vanilla can be achieved by adding infrequent but large negative jumps

Q: Why simultaneous jumps?
A: Realistic, do not decrease the implied spot-volatility correlation

Jumps II
Q: Why discrete jumps?
A: Check instantaneous correlation between $X(t) = \ln S(t)$ and $Y(t)$:

$$\varsigma \equiv \frac{< dX(t), dY(t) >}{\sqrt{< dX(t), dX(t) >}\sqrt{< dY(t), dY(t) >}}$$
$$= \frac{\rho\varepsilon\sigma_{(\text{loc,svj})}(t, S(t))\vartheta(t, Y(t)) + \nu\eta\lambda}{\sqrt{(\sigma_{(\text{loc,svj})}(t, S(t))\vartheta(t, Y(t)))^2 + \nu^2\lambda}\sqrt{\epsilon^2 + \eta^2\lambda}}$$

In equity markets, $\varsigma^* \in [-0.5, -0.9]$
If $Y(t) \to \infty$ or $\lambda \to 0$ then $\varsigma \to \rho$
If $Y(t) \to 0$ or $\lambda \to \infty$ then $\varsigma \to \text{sign}(\nu) = -1$ if $\nu < 0$

For discrete jumps, the correlation is the smallest possible: $\varsigma \approx -1$
For exponential jumps: $\varsigma \to 0.5\text{sign}(\nu) = \text{sign}(\nu) = -0.5$ if $\nu < 0$

Jump variance leads to de-correlation (not good for stability of ρ)

Experience: even for 1-d jump-diffusion Merton model, adding jump variance to model calibration leads to less stable calibration

Jumps III. Q: How to estimate jump parameters?
Consider a jump-diffusion model under the historic measure:

$$dS(t)/S(t) = \mu dt + \sigma dW(t) + \left(e^{J} - 1\right) dN(t),\ S(t_0) = S,$$

$N(t)$ is Poisson process with intensity λ

Jumps PDF is normal mixture: $w(J) = \sum_{l=1}^{L} p_l n(J; \nu_l, \upsilon_l),\ \sum_{l=1}^{L} p_l = 1$
L, $L = 1, 2, ..$, is the number of mixtures
p_l, $0 \leq p_l \leq 1$, is the probability of the l-th mixture
ν_l and υ_l are the mean and variance of the l-th mixture, $l = 1, .., L$

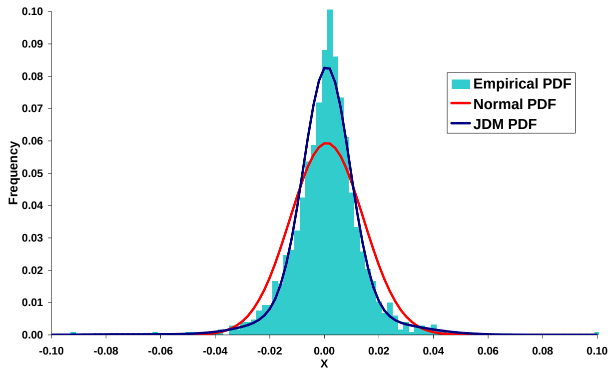
Cumulative probability function can be represented as weighted sum over normal probabilities (see Sepp (2011))

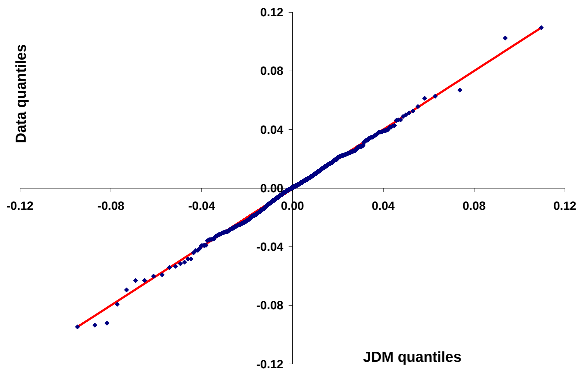
Can be applied for estimation of jumps

Jumps IV. Empirical estimation for the S&P500 using last 10 years
Fit empirical quantiles of sampled daily log-returns to model quantiles by minimizing the sum of absolute differences

The best fit, keeping L as small as possible, is obtained with $L = 4$
Model estimates: $\sigma = 0.1348$, $\lambda = 46.4444$, υ_l is uniform

l	p_l	ν_l	$\sqrt{\upsilon_l}$	$p_l\lambda$	expected frequency
1	0.0208	-0.0733	0.0127	0.9641	every year
2	0.5800	-0.0122	0.0127	26.9399	every two weeks
3	0.3954	0.0203	0.0127	18.3661	every six weeks
4	0.0038	0.1001	0.0127	0.1743	every five years





Jumps V. Implied Calibration of Jumps

Remove intermediate jumps ($l = 2, 3$) and large positive jumps ($l = 4$) - the stochastic local volatility takes care of "moderate" skew

Jump magnitude from historical estimation $\nu = -0.0733$ **is too small to fit one-month SPX skew**

An estimate implied from one-month SPX options is (approximately) $\nu = -0.25$

Assume **only large negative jumps** with $\lambda = 0.2$

The magnitude of jump in volatility η is calibrated to match VIX skews

Typically $\eta = 1.0$ implying that ATM volatility will double following a jump

Local Volatility
Using local volatility we fit vanilla options while having flexibility to specify the dynamics for stochastic volatility and jumps

Local volatility parametrization:

1) Parametric local volatility

Specify a functional form for local volatility (CEV, shifted-lognormal, quadratic, etc)

Calibrate by boot-strap and least-squares

Good for single stocks where implication of Dupire non-parametric local volatility is problematic and unstable due to lack of quotes

2) Non-parametric local volatility

Fit the model local volatility to Dupire local volatility

Good for indices with liquid option market

Parametric local volatility I

Consider the following parametric form:

$$\sigma_{\text{loc}}(t,S) = \sigma_{\text{atm}}(t)\sigma_{\text{skew}}(t,S)\sigma_{\text{smile}}(t,S) \tag{2}$$

where $\sigma_{\text{atm}}(t)$ is **at-the-money forward volatility**

Skew function is specified as ratio of two CEVs:

$$\begin{aligned}\sigma_{\text{skew}}(t,S) &= \frac{(S/S_0)^{\beta(t)-1} + a}{(S/S_0)^{\beta(t)-1} + b} \\ &= \frac{1}{2}\left((1+q) + (1-q)\tanh\left(\frac{1+q}{1-q}(\beta(t)-1)\ln(S/S_0)\right)\right)\end{aligned}$$

where $\beta(t)$ is the skew parameter, q is weight parameter, $0 < q < 1$

Smile function is quadratic:

$$\sigma_{\text{smile}}(t,S) = \sqrt{1 + (\alpha(t)\ln(S/S_0))^2}$$

where $\alpha(t)$ is smile parameter

Parametric local volatility II

Q: What is intuition about the model parameters and initial guesses for model calibration?

A: From the given market implied volatilities, compute ATM volatility, skew and smile as follows:

$\vartheta_{\text{atm}}(T) = \sigma(T, K = S(T))$
$\vartheta_{\text{skew}}(T) \propto \sigma(T, K = S(T)+) - \sigma(T, K = S(T)-)$
$\vartheta_{\text{smile}}(T) \propto \sigma(T, K = S(T)+) - 2\sigma(T, K = S(T)) + \sigma(T, K = S(T)-)$

Consider local volatility at $S = S_0$ to obtain:

$$\begin{aligned} \sigma_{\text{atm}}(t) &= \frac{2\vartheta_{\text{atm}}(t)}{1+q} \\ \beta(t) &= \frac{2\vartheta_{\text{skew}}(t)}{\vartheta_{\text{atm}}(t)} + 1 \\ \alpha^2(t) &= \frac{2\vartheta_{\text{smile}}(t)}{\vartheta_{\text{atm}}(t)} \end{aligned} \tag{3}$$

$\sigma_{\text{atm}}(t)$ is proportional to the ATM volatility
$\beta(t)$ is the skew relative to ATM volatility
$\alpha(t)$ is the smile relative to ATM volatility

Calibration of parametric models I. Backward equation
Consider general pricing equation for options on $S(t)$ under stochastic local volatility model
PV function $U(t,S,Y;T,K)$, $0 \le t \le T$, solves the **backward equation** as function of (t,S,Y):

$$\begin{aligned} &U_t + \mathcal{L}U(t,S,Y) = 0 \\ &U(T,S,Y) = u(S) \end{aligned} \tag{4}$$

where $u(S)$ is pay-off function
$\mathcal{L}$ is **infinitesimal backward operator**:

$$\begin{aligned} &\mathcal{L}U(t,S,Y) = \frac{1}{2}(\sigma_{(\text{loc,svj})}(t,S)\vartheta(t,Y)S)^2 U_{SS} + \mu(t)SU_S \\ &+ \frac{1}{2}\epsilon^2(Y)U_{YY} + \theta(Y)U_Y \\ &+ \rho\sigma_{(\text{loc,svj})}(t,S)\vartheta(t,Y)\epsilon(Y)SU_{SY} + \lambda I(S) \\ &I(S) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(Se^J, Y+\Upsilon)\varpi(J)\varsigma(\Upsilon)dJd\Upsilon - \nu SU_S - U \end{aligned}$$

For generality, jump sizes in S and Y have PDFs $\varpi(J)$ and $\varsigma(\Upsilon)$, respectively; ν is jump compensator

Calibration of parametric models II. Forward equation

Transition probability density function, $G(0, S_0, Y_0; t, S', Y')$ solves the **forward equation** as function of (t, S', Y'):

$$\begin{aligned} &G_t(t, S', Y') - \mathcal{L}^\dagger G(t, S', Y') = 0, \\ &G(0, S', Y') = \delta(S' - S_0)\delta(Y' - Y_0) \end{aligned} \tag{5}$$

where $\delta()$ is Dirac delta function

$\mathcal{L}^\dagger$ is forward **operator adjoint to** $\mathcal{L}$:

$$\begin{aligned} \mathcal{L}^\dagger G(t, S', Y') = &-\frac{1}{2}\Big((\sigma_{(\mathrm{loc,svj})}(T, S')\vartheta(T, Y')S')^2 G\Big)_{S'S'} + \Big(\mu(T)S'G\Big)_{S'} \\ &-\frac{1}{2}\Big(\epsilon^2(Y')G\Big)_{Y'Y'} + \Big(\theta(Y')G\Big)_{Y'} \\ &-\Big(\rho\sigma_{(\mathrm{loc,svj})}(T, S')\vartheta(T, Y')\epsilon(Y')S'G\Big)_{S'Y'} - \lambda I(S') \\ I(S') = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(S'e^{-J}, Y' - \Upsilon)e^{-J}\varpi(J)\varsigma(\Upsilon)dJd\Upsilon + (\nu S'G)_{S'} - G \end{aligned}$$

Calibration of parametric models III Option PV can be computed by convolution (Duhamel's or Feynman-Kac formula):

$$U(0, S_0, Y_0; T, K) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(0, S_0, Y_0; T, S', Y')u(S')dS'dY' \qquad (6)$$

Calibration by bootstrap: For parametric local volatility, model parameters are assumed to be piece-wise constant in time with jumps at times $\{T_m\}$, $\{T_m\}$ is set of maturity times of listed options

1) Given calibrated set of piece-wise constant model parameters at time T_{m-1} and known values of $G(T_{m-1}, S, Y)$, make a guess for parameters at time T_m and compute $G(T_m, S, Y)$

2) Compute PV-s of vanilla options using discrete version of (6)

3) By changing parameters for time T_m, minimize the sum of squared differences between model and market prices

4)After convergence, store $G(T_m, S, Y)$ and go to the next time slice

Calibration of parametric models IV. Numerical schemes
Consistency condition for adjoint operators $\mathcal{L}$ and $\mathcal{L}^{\dagger}$ (can be checked by integration by parts) is based on theoretical arguments:

$$\begin{aligned}\int_0^T \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Big[\Big\{\mathcal{L}^{\dagger} G(0, S_0, Y_0; t', S', Y')\Big\} U(t', S', Y') \\ -G(0, S_0, Y_0; t', S', Y') \Big\{\mathcal{L} U(t', S', Y')\Big\}\Big] dS' dY' dt' = 0\end{aligned} \tag{7}$$

It is not necessarily true for discrete numerical schemes (Lipton (2007))

Implications for numerical schemes:
1) For adjoint operators, if M is spacial discretisation of $\mathcal{L}$ then M^T should be spacial discretisation of $\mathcal{L}^{\dagger}$
2) If M^T is diagonally dominant with positive diagonal and non-positive off-diagonal elements (for 1-d problems with zero correlation), then $(M^T)^{-1} > 0$ and computed probabilities are non-negative and sum up to one (Andreasen (2010), Nabben (1999))
3) Both forward and backward equations should be solved using the same schemes
4) $\mathcal{L}$ and $\mathcal{L}^{\dagger}$ have the same values of operator norm, so the convergence of the forward scheme implies convergence of backward scheme

Calibration of parametric models V. Non-linear least square fit

Model parameters: $\Theta = (\theta_1, ..., \theta_q)$, Model prices: $U(\Theta) = (u_1, ..., u_n)$
Jacobian: $\mathbf{J}_{n\times q} = (J_{n'q'})$: $J_{n'q'} = \frac{\partial u_{n'}}{\partial \theta_{q'}}$ (computed numerically)
Market prices: $M = (m_1, ..., m_n)$, Weight (diagonal) matrix: $\mathbf{W}_{n\times n}$

Gauss-Newton normalized equations:

$(\mathbf{J}^T\mathbf{W}\mathbf{J})(\Delta\Theta) = \mathbf{J}^T\mathbf{W}(\Delta U)$, $\Delta U = M - U(\Theta^{(k)})$, $\Theta^{(k+1)} = \Theta^{(k)} + \Delta\Theta$
Use SVD for stability, dimensionality is small so calibration is fast

Levenberg-Marquardt method:

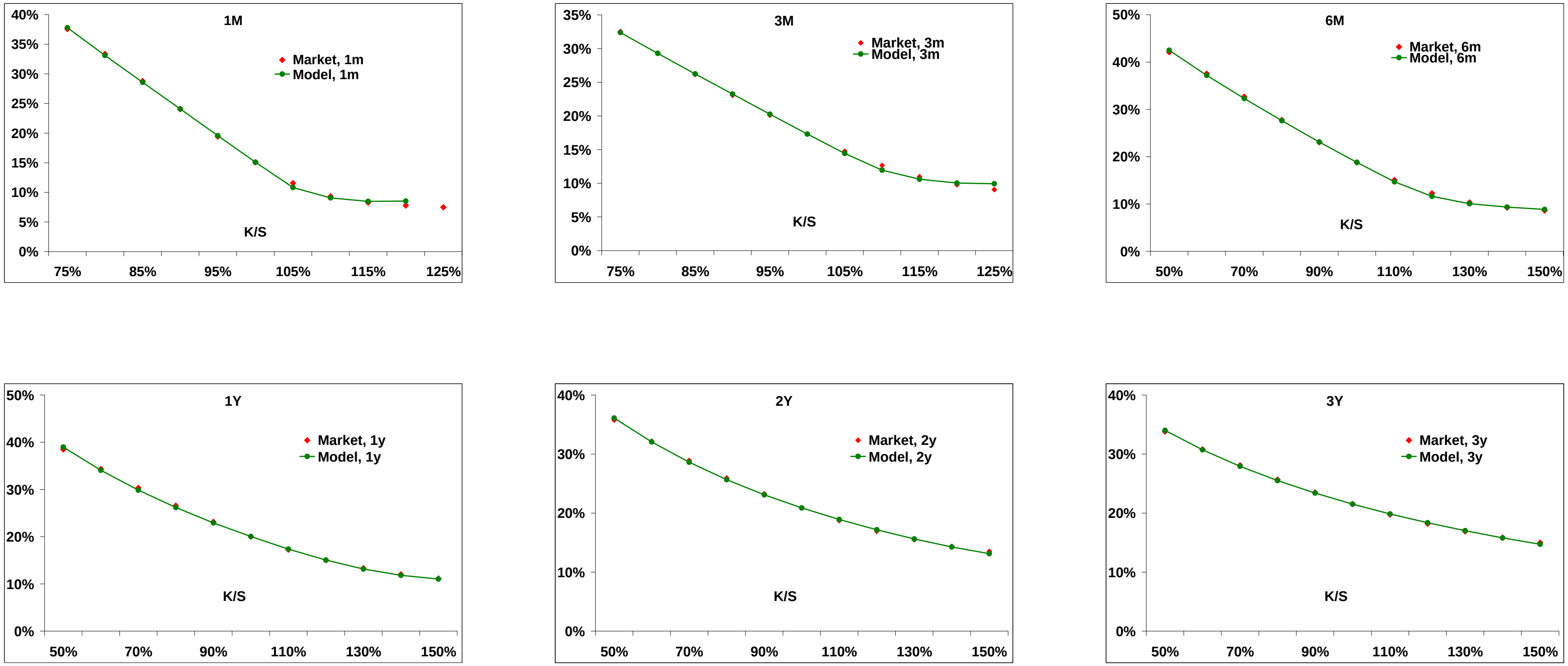
$$(\mathbf{J}^T\mathbf{W}\mathbf{J} + \lambda\mathbf{I})(\Delta\Theta) = \mathbf{J}^T\mathbf{W}(\Delta U)$$

λ is Marquardt parameter to make the matrix diagonally-dominant
Large value of λ - the steepest descent (tendency to be slow)
How to set λ? (NR in C (Press (1988)) use deterministic value)
One possibility: $\lambda = 1/\text{trace}(\mathbf{J}^T\mathbf{W}\mathbf{J})$

Experience: for carefully chosen functional form of the local volatility with robust initial estimates, Gauss-Newton works fast

Calibration of parametric models VI. Illustration

Implied volatilities for parametric volatility with vol-of-vol reduced by 50%: the fit is acceptable



Local volatility calibration I. Diffusion model

Consider a one-dimensional diffusion:

$$dS(t) = \mu(t)S(t)dt + \sigma_{(\text{loc,dif})}(t, S(t))S(t)dW(t),\ \ S(0) = S \tag{8}$$

where $\sigma_{(\text{loc,dif})}$ is the local volatility of the diffusion

Calibration objective: how we should specify $\sigma_{(\text{loc,dif})}(T, S')$?

In terms of call option prices, we have **Dupire equation** (1994):

$$\sigma^2_{(\text{loc,dif})}(T, K) = \frac{C_T(T, K) + \mu(T)KC_K(T, K) - \mu(T)C(T, K)}{\frac{1}{2}K^2C_{KK}(T, K)} \tag{9}$$

Local volatility calibration II. Jump-diffusion model. Andersen-Andreasen (2000) consider Merton jump-diffusion with local volatility:

$$
\begin{aligned}
dS(t) =& \mu(t)S(t_-)dt + \sigma_{(\text{loc,jd})}(t, S(t_-))S(t_-)dW(t) \\
&+ \Big((e^J - 1)dN(t) - \lambda\nu dt\Big) S(t_-), \ \ S(0) = S
\end{aligned} \tag{10}
$$

Here $\sigma^2_{(\text{loc,jd})}$ is specified by **Andersen-Andreasen equation** (2000):

$$
\begin{aligned}
\sigma^2_{(\text{loc,jd})}(T, K) &= \frac{C_T(T,K) + \mu(T)KC_K(T,K) - \mu(T)C(T,K) - \lambda I^{\dagger}(K)}{\frac{1}{2}K^2 C_{KK}(T,K)} \\
&= \sigma^2_{(\text{loc})}(T,K) - \frac{\lambda I^{\dagger}(K)}{\frac{1}{2}K^2 C_{KK}(T,K)}
\end{aligned} \tag{11}
$$

$$
I^{\dagger}(K) = \int_{-\infty}^{\infty} C(Ke^{-J'})e^{J'}\varpi(J')dJ' + \nu K C_K - (\nu+1)C \tag{12}
$$

Calibration: i) Given Dupire local volatility $\sigma_{(\text{loc,dif})}(T, K)$ solve forward equation for call prices $C(T,K)$ using FD for 1-d problem
ii) Compute $I^{\dagger}(K)$ and imply $\sigma_{(\text{loc,jd})}(T, K)$

Local volatility calibration III. Stochastic local volatility:

$$dS(t) = \mu(t)S(t)dt + \sigma_{(\text{loc,sv})}(t, S(t))\vartheta(t, Y(t))S(t)dW^{(1)}(t),\ S(0) = S,$$
$$dY(t) = \theta(Y(t))dt + \epsilon(Y(t))dW^{(2)}(t),\ Y(0) = Y$$

The model is consistent with the **Dupire local volatility** if

$$\sigma^2_{(\text{loc,sv})}(T, K) = \frac{\sigma^2_{(\text{loc,dif})}(T, K)}{V(T, K)} \tag{13}$$

where $V(T, S')$ is **the conditional variance**:

$$V(T, S') = \frac{\int_{-\infty}^{\infty} \vartheta^2(T, Y')G(t, S, Y; T, S', Y')dY'}{\int_{-\infty}^{\infty} G(t, S, Y; T, S', Y')dY'} \tag{14}$$

Calibration:(More details to follow)
i) Stepping forward in time, solve for density $G(t, S, Y; T, S', Y')$ using FD for 2-d problem
ii) Compute $V(T, K)$ and imply $\sigma^2_{(\text{loc,sv})}(T, K)$ using (13)

Local volatility calibration IV. Stochastic local volatility with jumps specified by dynamics (1)
Local stochastic volatility for jump-diffusion, $\sigma^2_{(\text{loc,svj})}(T,K)$, is specified by **Lipton equation** (2002):

$$\begin{aligned}\sigma^2_{(\text{loc,svj})}(T,K) &= \frac{C_T(T,K)+\mu(T)KC_K(T,K)-\mu(T)C(T,K)-\lambda I^{\dagger}(K)}{\frac{1}{2}V(T,K)K^2C_{KK}(T,K)}\\ &= \frac{1}{V(T,K)}\left(\sigma^2_{(\text{loc,dif})}(T,K)-\frac{\lambda I^{\dagger}(K)}{\frac{1}{2}K^2C_{KK}(T,K)}\right)\\ &= \frac{\sigma^2_{(\text{loc,jd})}(T,K)}{V(T,K)}\end{aligned} \tag{15}$$

Calibration: i) Given Dupire local volatility $\sigma_{(\text{loc,dif})}(T,K)$ solve forward equation for call prices $C(T,K)$ using FD for 1-d problem and imply $\sigma_{(\text{loc,jd})}(T,K)$ using (11)
ii) Stepping forward in time, solve for density $G(t,S,Y;T,S',Y')$ using FD for 2-d problem with jumps
iii) compute $V(T,K)$ using (14) and imply $\sigma^2_{(\text{loc,sv})}(T,K)$ using (15)

PDE based methods

Non-parametric local stochastic volatility can only be implemented using numerical methods for partial integro-differential equations

These methods should be flexible to handle jumps in one and two dimensions and the correlation term for two dimensions

I will review some of the available methods

1-d Problem. For the backward problem:

$$\mathcal{L} = \mathcal{D} + \lambda \mathcal{J}$$

Here $\mathcal{D}$ is the diffusion-convection operator:

$$\mathcal{D}U(t,S) \equiv \frac{1}{2}\sigma^2(t,S)S^2 U_{SS} + \mu(t)SU_S + \lambda U$$

$\mathcal{J}$ is the integral operator:

$$\mathcal{J}U(t,S) \equiv \int_{-\infty}^{\infty} U(t, Se^{J'})\varpi(J')dJ'$$

For the forward problem, we consider the operator $\mathcal{L}^\dagger$ adjoint to $\mathcal{L}$:

$$\mathcal{L}^\dagger = \mathcal{D}^\dagger + \lambda \mathcal{J}^\dagger$$

For both problems, denote the discretized diffusion and integral operators by $\mathbf{D}^n$ and $\mathbf{J}$, respectively

i) The diffusion operator is space and time dependent
ii) Care must be taken for discretization of the operator for the mean-reverting process

1-d Problem for diffusion problem. N - number of spacial steps
$\mathbf{D}^n$ - discrete diffusion matrix at time t_n with dimension $N \times N$
U^n - the solution vector at time t_n with dimension $N \times 1$
Time-stepping with θ-method:

$$\left(\mathbf{I} - \theta \mathbf{D}^{n+1}\right) U^{n+1} = \left(\mathbf{I} + \theta \mathbf{D}^{n+1}\right) U^n$$

i) $\theta = 0$ - explicit method requires very fined grids and is highly unstable - it should be avoided
ii) $\theta = 1/2$ - Crank-Nicolson scheme is unconditionally stable and is of order $(\delta S)^2$ and $(\delta t)^2$, but might lead to negative densities
iii) $\theta = 1$ - implicit method is unconditionally stable and does not lead to negative densities, but it is of order (δt) and becomes less accurate for longer maturities

My favourite is the predictor-corrector scheme:

$$\left(\mathbf{I} - \mathbf{D}^{n+1}\right) \widetilde{U} = U^n$$

$$\left(\mathbf{I} - \mathbf{D}^{n+1}\right) U^{n+1} = U^n + \frac{1}{2} \mathbf{D}^{n+1} (U^n - \widetilde{U})$$

with order of $(\delta S)^2$ and $(\delta t)^2$; does not lead to negative densities

Numerics for jump-diffusions I

Consider the forward equation with additive jumps (multiplicative jumps can be handled in the log-space):

$$\mathcal{J}G(x) = \int_{-\infty}^{\infty} G(x-j)\varpi(J')dJ'$$

Consider discrete negative jumps with size $-\eta$, $\eta > 0$:

$$\mathcal{J}G(x) = G(x+\eta)$$

Discretization is a simple linear interpolation:

$$\mathcal{J}G_i = wG_k + (1-w)G_{k+1}$$

where $k = \min\{k : x_{k+1} \geq x_i + \nu\}$ and $w = \frac{x_{k+1}-(x_i+\nu)}{x_{k+1}-x_k}$

Numerics for jump-diffusions II

In the matrix form, for uniform spacial grid:

$$\mathbf{J} = \begin{pmatrix} 0 & 0 & ... & w & 1-w & 0 & ... & 0 & 0 \\ 0 & 0 & ... & 0 & w & 1-w & ... & 0 & 0 \\ ... & & & & & & & & \\ 0 & 0 & ... & 0 & 0 & 0 & ... & w & 1-w \\ 0 & 0 & ... & 0 & 0 & 0 & ... & 0 & 1 \\ ... & & & & & & & & \\ 0 & 0 & ... & 0 & 0 & 0 & ... & 0 & 0 \end{pmatrix}$$

$\mathbf{J}_{(p,N)} = 1$, where $p = \min\{p : x_p + \nu \geq x_N\}$

In general, for two-sided jumps, $\mathbf{J}$ is a full matrix:
The implicit method for the integral term leads to $O(N^3)$ complexity and is not practical
The explicit method leads to $O(N^2)$ complexity but needs extra care for stability

For exponential jumps, Lipton (2003) develops recursive scheme with $O(N)$ complexity

Numerics for jump-diffusions III

In case of discrete jumps, we have two-banded matrix with p rows that have non-zero elements

Consider solving a linear system:

$$(\mathbf{I} - \beta\mathbf{J})X = R$$

where $\beta = \lambda\delta t$, $0 < \beta < 1$

We can solve this system by back-substitution with cost of $O(N)$ operations

Numerics for jump-diffusions IV

Consider θ and θ_J schemes for the diffusion and the jump parts:

$$\left(\mathbf{I} - \theta \mathbf{D}^{n+1} - \theta_J \lambda_{n+1} \mathbf{J}\right) U^{n+1} = \left(\mathbf{I} + \theta \mathbf{D}^{n+1} + \theta_J \lambda_{n+1} \mathbf{J}\right) U^{n}$$

where $\lambda_{n+1} = (t_{n+1} - t_n)\lambda(t_{n+1})$

The accuracy with $\theta = \theta_J = \frac{1}{2}$ is supposed to be $O(dx^2) + O(dt^2)$

However, the matrix in the lhs will be full: the cost to invert is $O(N^3)$

Implicit-explicit method:

$$\left(\mathbf{I} - \mathbf{D}^{n+1}\right) U^{n+1} = \left(\mathbf{I} + \lambda_{n+1} \mathbf{J}\right) U^{n}$$

with accuracy of $O(dx^2) + O(dt)$

θ-**explicit method**:

$$\left(\mathbf{I} - \theta \mathbf{D}^{n+1}\right) U^{n+1} = \left(\mathbf{I} + \theta \mathbf{D}^{n+1} + \lambda_{n+1} \mathbf{J}\right) U^{n}$$

with accuracy of $O(dx^2) + O(dt)$

Numerics for jump-diffusions V. Andersen-Andreasen (2000) scheme

Make the first half step with $\theta = 1$ and $\theta_J = 0$ and the second half step with $\theta = 1$ and $\theta_J = 1$:

$$\left(\mathbf{I} - \frac{1}{2}\mathbf{D}^{n+1}\right)\widetilde{U} = \left(\mathbf{I} + \frac{1}{2}\lambda_{n+1}\mathbf{J}\right)U^n$$
$$\left(\mathbf{I} - \frac{1}{2}\lambda_{n+1}\mathbf{J}\right)U^{n+1} = \left(\mathbf{I} + \frac{1}{2}\mathbf{D}^{n+1}\right)\widetilde{U}$$

with accuracy of $O(dx^2) + O(dt^2)$

When $\mathbf{J}$ is full, the second equation can be solved by the application of the discrete Fourier transform (DFT) with the cost of $O(N \log N)$

Beware of deficiencies of the DFT

For discrete jumps, the second equation solved with cost of $O(N)$ operations

Numerics for jump-diffusions VI. d'Halluin et al (2005) scheme

Consider fixed point iterations:

Set $V^1 = U^n$

Iterate for $p = 1, .., \overline{p}$ ($\overline{p} = 2$ is good enough):

$$\left(\mathbf{I} - \theta \mathbf{D}^{n+1}\right) V^{p+1} = \left(\mathbf{I} + \theta \mathbf{D}^{n+1}\right) U^n + \lambda_{n+1} \mathbf{J} V^p,$$

Set $U^{n+1} = V^{\overline{p}}$

The accuracy is $O(dx^2) + O(dt)$

Numerics for jump-diffusions VII

My favourite is the implicit scheme with predictor-corrector applied twice:

$$
\begin{aligned}
\widetilde{U}^{(0)} &= \left(\mathbf{I} + \mathbf{D}^{n+1} + \lambda_{n+1}\mathbf{J}\right) U^n \\
\left(\mathbf{I} - \mathbf{D}^{n+1}\right) \widetilde{U} &= \widetilde{U}^{(0)} \\
\left(\mathbf{I} - \mathbf{D}^{n+1}\right) U^{n+1} &= \widetilde{U}^{(0)} + \frac{1}{2}(\mathbf{D}^{n+1} + \lambda_{n+1}\mathbf{J})(\widetilde{U} - U^n)
\end{aligned}
$$

The accuracy is $O(dx^2) + O(dt^2)$

Numerical Methods for Two Dimensional Problem I

We consider the forward problem for $U(t, x_1, x_2)$:

$$\begin{aligned} &U_T - \mathcal{M}U = 0 \\ &U(0, x_1, x_2) = \delta(x_1 - x_1(0))\delta(x_2 - x_2(0)) \end{aligned} \qquad (16)$$

where

$$\mathcal{M} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{C} + \lambda\mathcal{J}$$

$\mathcal{D}_1$ and $\mathcal{D}_2$ are 1-d diffusion-convection operators in x_1 and x_2 directions, respectively
$\mathcal{C}$ is the correlation operator
$\mathcal{J}$ is the integral operator for joint jumps in x_1 and x_2

Let $\mathbf{D}_1$ and $\mathbf{D}_2$ denote the discretized 1-d diffusion-convection operators in x_1 and x_2 directions, respectively

$\mathbf{C}$ and $\mathbf{J}$ are the discretized correlation and integral operator, respectively

Douglas-Rachford (1956) scheme

Make a predictor and apply two orthogonal corrector steps:

$$\begin{aligned}
\widetilde{U}^{(0)} &= (\mathbf{I} + \mathbf{C} + \lambda\mathbf{J} + \mathbf{D}_1 + \mathbf{D}_2)U^n \\
(\mathbf{I} - \theta\mathbf{D}_1)\widetilde{U} &= \widetilde{U}^{(0)} - \theta\mathbf{D}_1 U^n \\
(\mathbf{I} - \theta\mathbf{D}_2)U^{n+1} &= \widetilde{U} - \theta\mathbf{D}_2 U^n
\end{aligned}$$

In the second line, for each fixed index j we apply the diffusion operator in x_1 direction; and solve the tridiagonal system of equations to get the auxiliary solution $\widetilde{U}(\cdot, x_2(j))$

In the third line, keeping i fixed, we apply the diffusion step in x_2 direction and solve the system of tridiagonal equations to get the solution $U^{n+1}(x_1(i), \cdot)$ at time t_{n+1}

Complexity is $O(N_1 N_2)$ per time step

Craig-Sneyd (1988) scheme

Start as with Douglas-Rachford scheme make a second predictor and again apply two orthogonal corrector steps:

$$\begin{aligned}
&\widetilde{U}^{(0)} = (\mathbf{I} + \mathbf{C} + \lambda\mathbf{J} + \mathbf{D}_1 + \mathbf{D}_2)U^n \\
&(\mathbf{I} - \theta\mathbf{D}_1)\widetilde{U}^{(1)} = \widetilde{U}^{(0)} - \theta\mathbf{D}_1U^n \\
&(\mathbf{I} - \theta\mathbf{D}_2)\widetilde{U}^{(2)} = \widetilde{U}^{(1)} - \theta\mathbf{D}_2U^n \\
&\widetilde{U}^{(3)} = \widetilde{U}^{(0)} + \frac{1}{2}(\mathbf{C} + \lambda\mathbf{J})(\widetilde{U}^{(2)} - U^n) \\
&(\mathbf{I} - \theta\mathbf{D}_1)\widetilde{U}^{(4)} = \widetilde{U}^{(3)} - \theta\mathbf{D}_1U^n \\
&(\mathbf{I} - \theta\mathbf{D}_2)U^{n+1} = \widetilde{U}^{(4)} - \theta\mathbf{D}_2U^n
\end{aligned}$$

Hundsdorfer-Verwer (2003) scheme (in't Hout-Foulon (2008))

Similar to Craig-Sneyd scheme with predictor including $\mathbf{D}_1$ and $\mathbf{D}_2$ and the second corrector applied on $\widetilde{U}^{(2)}$:

$$\widetilde{U}^{(0)} = (\mathbf{I} + \mathbf{C} + \lambda\mathbf{J} + \mathbf{D}_1 + \mathbf{D}_2)U^n$$
$$(\mathbf{I} - \theta\mathbf{D}_1)\widetilde{U}^{(1)} = \widetilde{U}^{(0)} - \theta\mathbf{D}_1 U^n$$
$$(\mathbf{I} - \theta\mathbf{D}_2)\widetilde{U}^{(2)} = \widetilde{U}^{(1)} - \theta\mathbf{D}_2 U^n$$
$$\widetilde{U}^{(3)} = \widetilde{U}^{(0)} + \frac{1}{2}(\mathbf{C} + \lambda\mathbf{J} + \mathbf{D}_1 + \mathbf{D}_2)\left(\widetilde{U}^{(2)} - U^n\right)$$
$$(\mathbf{I} - \theta\mathbf{D}_1)\widetilde{U}^{(4)} = \widetilde{U}^{(3)} - \theta\mathbf{D}_1\widetilde{U}^{(2)}$$
$$(\mathbf{I} - \theta\mathbf{D}_2)U^{n+1} = \widetilde{U}^{(4)} - \theta\mathbf{D}_2\widetilde{U}^{(2)}$$

Discretisation of integral term

Direct methods are infeasible because of $O(N_1^2 N_2^2)$ complexity

DFT method (Clift-Forsyth (2008)) has $O(N_1 N_2 \log N_1 N_2)$ complexity but suffers from problems associated with the DFT

Explicit methods with $O(N_1 N_2)$ complexity are available for discrete and exponential jumps (Lipton-Sepp (2011))

The simplest case is if jumps are discrete with sizes η_1 and η_2:

$$\mathcal{J}G = G(x_1 - \eta_1, x_2 - \eta_2)$$

This term is approximated by bi-linear interpolation with the second order accuracy leading to the $O(N_1 N_2)$ complexity

Summary I. LSV model calibration

1) Specify time and space grids

2) Compute the local volatility using Dupire equation (9) on the specified grid (interpolating implied volatilities in strikes and maturities)

3) Calibrate local stochastic volatility on the specified grid

4) Use backward PDE methods or Monte-Carlo simulations of the local stochastic volatility model to compute values of exotic options (see Andreasen-Huge (2010) for consistent schemes)

Summary II. Local volatility calibration at time t_n
given $G^{n-1}(x_1(i), x_2(j))$ and $V(t_{n-1}, x_1(i))$
i) Compute the local variance by:

$$\sigma^2_{(\text{loc,sv})}(t_n, x_1(i)) = \frac{\sigma^2_{(\text{loc,dif})}(t_n, x_1(i))}{V(t_{n-1}, x_1(i))}$$

ii) Compute $G^n(x_1(i), x_2(j))$ by solving the forward PDE

iii) Compute $V(t_n, x_1(i))$ by

$$V(t_n, x_1(i)) = \frac{\sum_i \sum_j \vartheta^2(t_n, x_2(j)) G^n(x_1(i), x_2(j))}{\sum_j G^n(x_1(i), x_2(j))}$$

and set

$$\sigma^2_{(\text{loc,sv})}(t_n, x_1(i)) = \frac{\sigma^2_{(\text{loc,dif})}(t_n, x_1(i))}{V(t_n, x_1(i))}$$

D) Repeat **B)** and **C)** and go to the next time step

For stochastic volatility with jumps we use $\sigma^2_{(\text{loc,jd})}$

Summary III. Predictor-corrector schemes

When we use predictor-corrector schemes:

i) Compute the predictor step using local stochastic volatility from previous time step

ii) Update local stochastic volatility and compute the corrector step with new volatility

iii) After the corrector step, compute new local stochastic volatility and go to the next step

Summary IV. Discrete dividends

At ex-dividend time t_n:

$$S(t_n+) = S(t_n-) - D_n$$

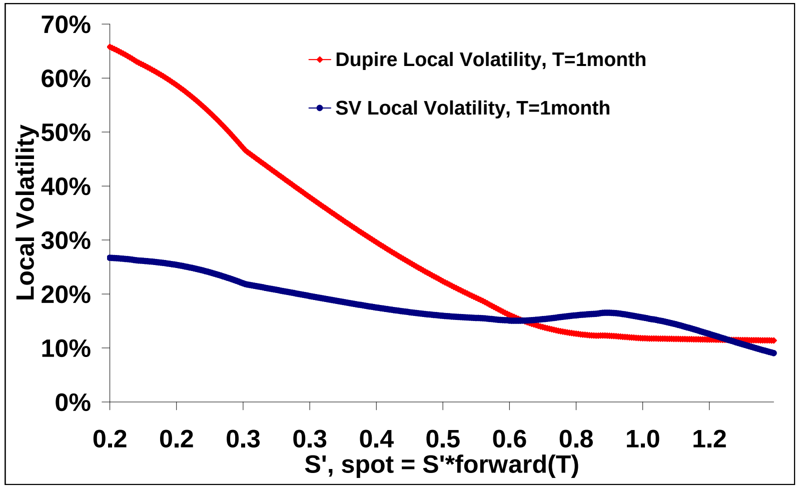
where D_n is the cash dividend at time t_n

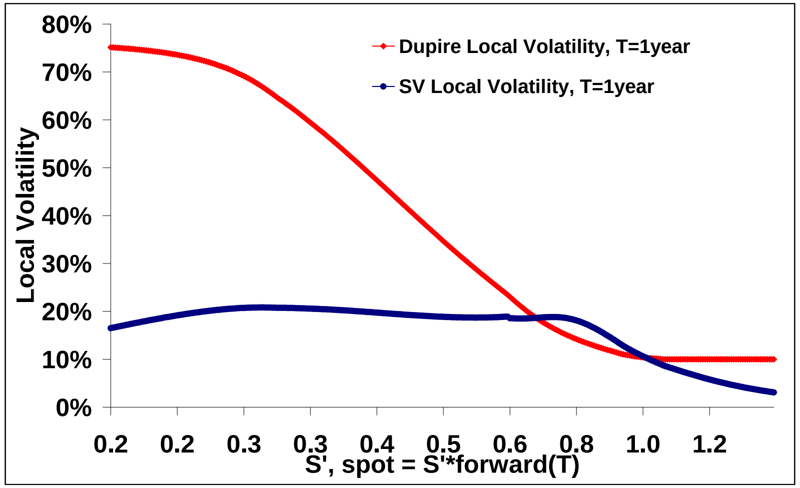
Note that this corresponds to the negative jump in $S(t)$ with a constant magnitude D_n at deterministic time t_n

Therefore the developed method for discrete negative jumps are readily applied for discrete dividends

It is important that $\sigma^2_{\text{(loc,dif)}}$ is consistent with the discrete dividends

Illustration I. Local volatilities for options on the S&P 500

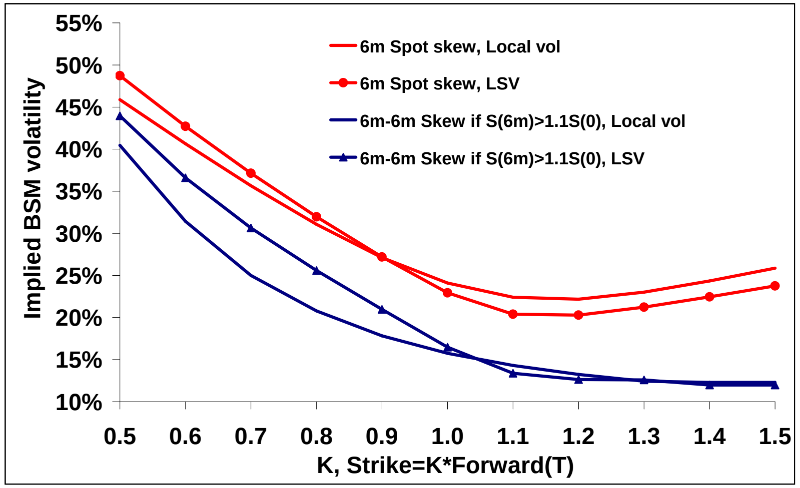


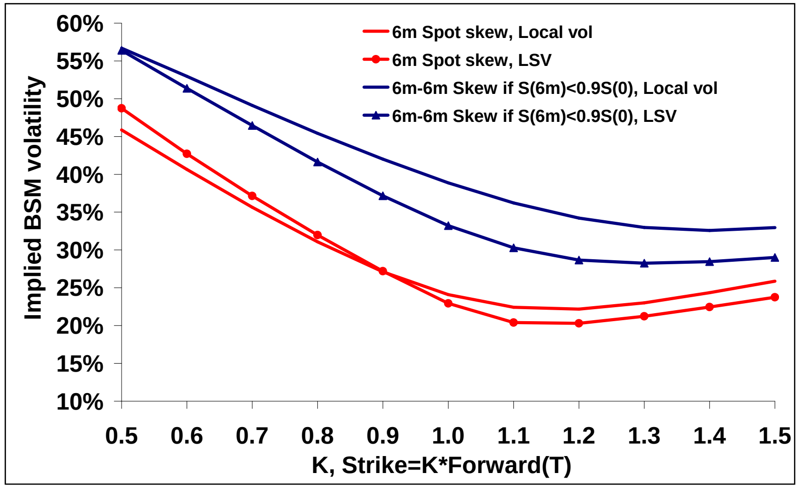


Left: Dupire local volatility and LSV local volatilities at T=1 month
Right: ... at T=1 year

LSV local volatility is an adjustment to the skew implied by the stochastic volatility part and is mostly flat

Illustration II. Forward volatilities for options on the S&P 500





6m-6m skew: the implied volatility for a vanilla call option starting in 6 month and maturing in one year from now with strike fixed in 6 months: $K = S(T = 6m)$

Left: 6m-6m skew if $S(T = 6m) > 1.1S(0)$: the implied volatility for the forward-start option conditional that $S(T = 6m) > 1.1S(0)$

Right: 6m-6m skew if $S(T = 6m) < 0.9S(0)$: the implied volatility for the forward-start option conditional that $S(T = 6m) < 0.9S(0)$

Illustration III. Implications

Conditional that spot moves up, the LSV model preserves the shape of the volatility skew while the pure local volatility does not

Conditional that spot moves down, the pure local volatility implies that the ATM volatility goes too high and the skew flattens unlike the LSV model

The LSV is more closer to the actual dynamics!

Illustration IV. Quarterly 95−105% **spread clique on SPX**, $T = 2y$:

$$u = \sum_n \max\left\{\min\left\{\frac{S(t_n)}{S(t_{n-1})} - 1,\ c\right\},\ -f\right\} \quad c = f = 5\%$$

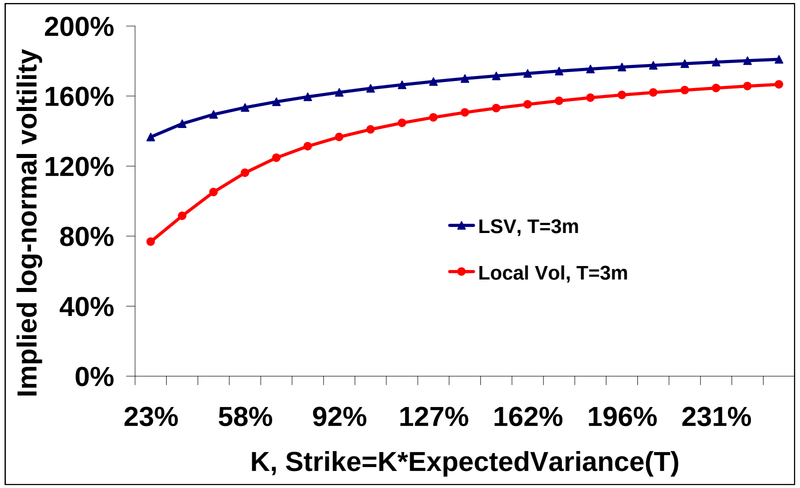
Marking model[1]	**5.42%**	**considered as "market price"**
1-d LocalVol	0.90%	tiny, no fit to steep forward skew!
1-d LocalVol+jump[3]	2.40%	still small
Heston	<u>4.33%</u>	<u>under-price</u>
Heston+jump [2]	*6.08%*	*over-price*
LSV, 70% [3]	3.78%	too small
LSV, 100% [3]	<u>4.48%</u>	<u>under-price, in line with Heston</u>
LSV, 130% [3]	4.77%	still under-price !
LSV+jump[2] 70% [3]	**5.37%**	**closest to the market**
LSV+jump[2] 100% [3]	*6.37%*	*over-price*

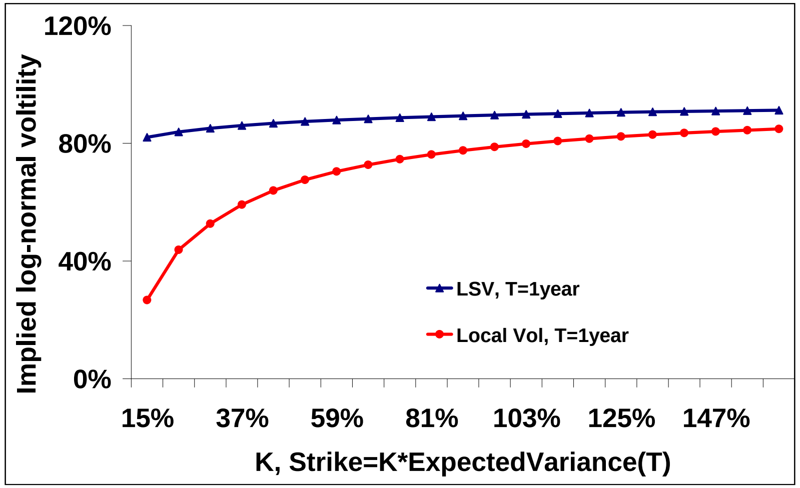
[1]internal "local beta" model for cliques developed by Alex Langnau

[2]1 jump per 5 years: $\lambda = 0.2$, $\nu = -25\%$

[3]vol-of-vol is set by $x\% * \varepsilon$, $\varepsilon = 2.3$

Illustration V. Implied volatilities of options on the daily realized variance of the S&P 500





Left: Implied log-normal volatilities on options on the daily realized variance of the S&P 500 with maturity T=3 month
Right: ... at T=1 year

LSV local volatility implies higher volatility of the realized variance ad the vol-of-vol parameter can be "tuned-up" to match the market

Options on the VIX I

Augment the model dynamics (1) with the third variable for the realized quadratic variance of $S(t)$:

$$dI(t) = \Big(\sigma_{(\mathrm{loc,svj})}(t_-, S(t_-))\vartheta(t_-, Y(t_-))\Big)^2 dt + \nu^2 dN(t), \;\; I(0) = 0$$

Consider the expected variance realized over period $[T, T + T_{\mathrm{vix}}]$:

$$\widetilde{I}(T, T + T_{\mathrm{vix}}) = \frac{1}{T_{\mathrm{vix}}}\mathbb{E}\left[\int_T^{T+T_{\mathrm{vix}}} dI(t')\right]$$

where $T_{\mathrm{vix}} = 30/365.25$ is the annualized tenor of the VIX

A call option on the VIX maturing at T is computed by:

$$C(T, K) = \mathbb{E}\left[\Big(\widetilde{I}(T, T + T_{\mathrm{vix}}) - K\Big)^+ \mid \widetilde{I}(T, T) = 0\right]$$

Options on the VIX II. Valuation

Consider

$$\widetilde{I}(T,T+T_{\mathsf{vix}}) = \frac{1}{T_{\mathsf{vix}}}\mathbb{E}\left[\int_T^{T+T_{\mathsf{vix}}} dI(t')\right] \approx \frac{1}{T_{\mathsf{vix}}}\sum_{t_n\in[T,T+T_{\mathsf{vix}}]}\left(\mathbb{E}\left[V(t_n)\right]+\nu^2\lambda\right)dt_n$$

where $V(t)$ is instantaneous variance:

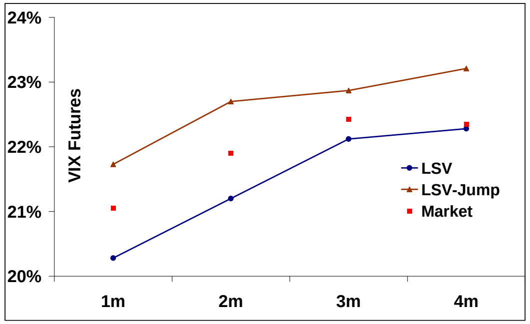
$$V(t) = \left(\sigma_{(\mathsf{loc,svj})}(t,S(t))\vartheta(t,Y(t))\right)^2$$

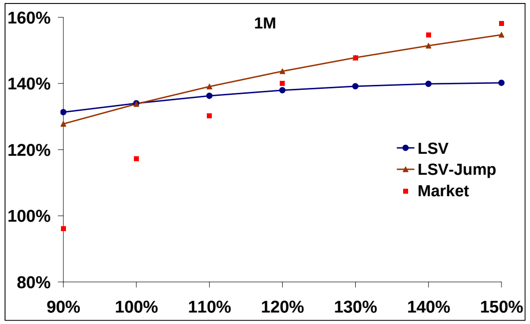
$U(t,S,Y) \equiv \widetilde{I}(T,T+T_{\mathsf{vix}})$ solves backward problem (4), $t \in [T,T+T_{\mathsf{vix}}]$:

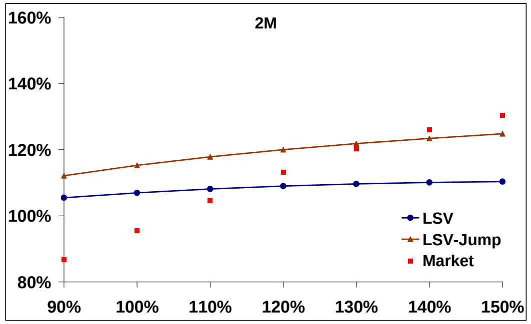
$$U_t + \mathcal{L}U(t,S,Y) = -\frac{1}{T_{\mathsf{vix}}}\left(V(t)+\nu^2\lambda\right)$$
$$U(T+T_{\mathsf{vix}},S,Y) = 0$$

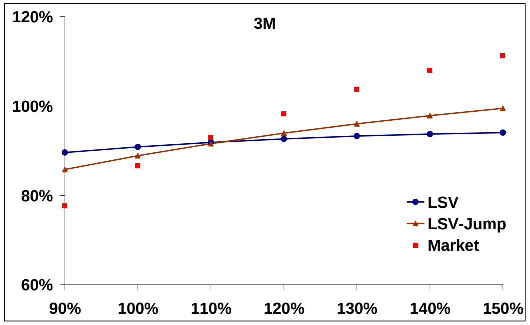
Valuation: **i)** Solve 2-d backward problem for $\widetilde{I}(T,T+T_{\mathsf{vix}})$ as function of (S,Y)
ii) Compute $G(t,S,Y;T,S',Y')$ and evaluate $C(T,K)$ by convolution
iii) Can price call with different strikes at a time

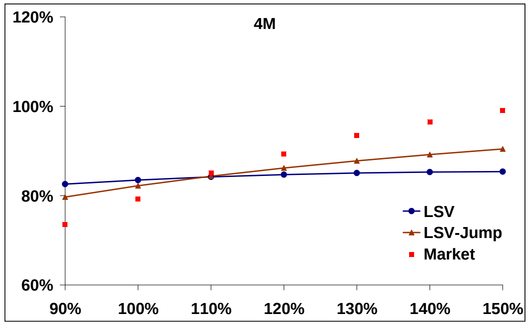
Options on the VIX III. VIX Implied volatilities











Conclusion: consistency with options on the SPX does not lead to consistency with options on the VIX; Need extra flexibility for jumps in volatility (time and/or space dependency)

Conclusions

I have presented theoretical and practical grounds for stochastic local volatility models and highlighted details of model implementation

The presentation is based on my experience in implementation of the LSV model with jumps for BAML equity derivatives library

During this project I benefited from interactions with Alex Lipton, Hassan El Hady and other members of BAML quantitative analytics

Thank you for your attention!

References

Andersen, L. and Andreasen J. (2000), "Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing", *Review of Derivatives Research* **4**, 231-262

Andreasen, J. (2010), "Volatility Interpolation and Finite Difference Based Simulation", *Presentation at ICBI Global Derivatives*

Andreasen, J. and Huge N. (2010), "Finite Difference Based Calibration and Simulation ", *Working paper*, ssrn.com/abstract= 1697545

Blacher, G. (2001), "A new approach for designing and calibrating stochastic volatility models", *Presentation at ICBI Global Derivatives*

Craig I. and Sneyd A. (1988), "An alternating-direction implicit scheme for parabolic equations with mixed derivatives ", *Comput. Math. Appl.* **16(4)** 341-350

Clift S. and Forsyth P. (2008), "Numerical solution of two asset jump diffusion models for option valuation", *Applied Numerical Mathematics* **58**, 743-782

Derman, E., and Kani, I. (1994), "The volatility smile and its implied tree", *Goldman Sachs Quantitative Strategies Research Notes*.

d'Halluin Y., Forsyth P. and Vetzal K. (2005), "Robust numerical methods for contingent claims under jump diffusion processes", *IMA Journal of Numerical Analysis* **25**, 87-112

Dupire, B. (1994), "Pricing with a smile", *Risk*, **7**, 18-20.

Douglas J. and Rachford H.H. (1956), "On the numerical solution of heat conduction problems in two and three space variables", *Trans. Amer. Math. Soc.* **82** 421-439

Heston, S. (1993), "A closed-form solution for options with stochastic volatility with applications to bond and currency options", *Review of Financial Studies* **6**, 327-343

in 't Hout, K.J., Foulon S. (2008), "ADI finite difference schemes for option pricing in the Heston model with correlation," *Working paper*

Hundsdorfer W and Verwer J. G. (2003), "Numerical Solution of Time-Dependent Advection-Diffusion- Reaction Equations", *Springer*

JP Morgan (1999), "Pricing exotics under smile", *Risk*, **11** 72-75

Lipton, A. (2002). "The vol smile problem", *Risk, February*, 81-85.

Lipton, A. (2003), "Evaluating the latest structural and hybrid models for credit risk", *Presentation at ICBI Global Derivatives*

Lipton, A. (2007), "Pricing of credit-linked notes and related products", *Merrill Lynch research papers*

Lipton A. and Sepp A. (2011), "Credit Value Adjustment in the Extended Structural Default Model," *in The Oxford Handbook of Credit Derivatives, ed. Lipton A. and Rennie A.*, 406-463

Merton, R. (1973), "Theory of Rational Option Pricing", *The Bell Journal of Economics and Management Science* **4(1)**, 141-183.

Nabben, R. (1999), "On Decay Rates of Tridiagonal and Band Matrices", *SIAM J. Matrix Anal. Appl* **20**, 820-837

Press, W., Vetterling W., Teukolsky S., and Flannery B. (1988), "Numerical Recipes in C: The Art of Scientific Computing", *Cambridge University Press*

Ren Y., Madan D., Qian Q. (2007), "Calibrating and pricing with embedded local volatility models", *Risk*, **9**, 138-143

Rubinstein, M. (1994), "Implied binomial trees", *Journal of Finance*, **49**, 771-818

Sepp, A. (2011) "An Approximate Distribution of Delta-Hedging Errors in a Jump-Diffusion Model with Discrete Trading and Transaction Costs," *Quantitative Finance*, forthcoming, (ssrn.com/abstract=1360472)