

Markovian projection for equity, fixed income, and credit dynamics.

T. Misirpashaev

NumeriX

April 6, 2007

Financial Mathematics Seminar, Stanford University, 2007

Abstract

We begin with the classic result of Dupire which shows that any diffusion model with stochastic volatility can be reduced to a local volatility model without changing the prices of European options. Specifically, the value of the effective local volatility at state S and time T is equal to the expectation of the stochastic volatility conditional on achieving state S at time T . This leads to a technique of model calibration in which the original model without a low-dimensional Markovian representation is approximated by a low-dimensional Markovian model. We cite the results for the projection on an effective displaced diffusion and Heston models. We then set the goal of extending the technique from diffusions to jump processes used for dynamic modeling of credit basket loss. We identify the one-step Markov chain as the counterpart of the local volatility model and prove the version of the Dupire result applicable to jump processes. We conclude by observing that the local intensity of the effective Markov chain bears a distinctive signature of credit correlation skew, which can be used to predict success or failure of certain models in matching the market of CDO tranches.

Outline

1. Classic “Universal Theory of Volatility” by Dupire
2. Modern applications: European option pricing by Markovian projection
3. Extensions of the classic theory: Gyöngy lemma for multidimensional diffusion processes and Markovian projection on stochastic volatility models
4. Extension to jump processes: local intensity model
5. Applications to credit basket models: signature of correlation skew

Dupire's "Universal Theory of Volatility" (UTV)

Stochastic volatility model in the martingale measure: $dX_t = \gamma_t \cdot dW_t$.

Local volatility model in the martingale measure: $dY_t = g(Y_t, t)dB_t$.

(Note: W_t may have several components, B_t is 1-dimensional.)

Gyöngy (1986) - Dupire (1997) lemma: one-dimensional marginal distributions (and therefore European options) for X_t and Y_t are identical provided $X_0 = Y_0$ and

$$g^2(x, t) = \mathbb{E}[|\gamma_t|^2 | X_t = x].$$

Dupire gave a formula for $g(x, t)$ in terms of European options

$$C(K, T) = \mathbb{E}[(X_T - K)^+],$$

$$g^2(K, T) = \frac{\partial C(K, T) / \partial T}{\frac{1}{2} \partial^2 C(K, T) / \partial K^2}$$

Modern applications of UTV: Markovian Projection

Fast calculation of European options is essential for model calibration. Markovian projection helps because European options can be priced in an equivalent local volatility model.

How to compute the conditional expectation $E[|\gamma_t|^2 | X_t = x]$?

One way is to restrict the space of all local volatility functions $g(x, t)$ to a parametric subspace and do a regression, exploiting the minimizing property of the conditional expectation

$$E[|\gamma_t|^2 | X_t = x] = g^2(x, t) \Rightarrow E[(|\gamma_t|^2 - g^2(x, t))^2] \rightarrow \min$$

(For an alternative, see Avellaneda et al. 2002)

Projection on a displaced diffusion

Choose the subspace

$$g(x, t) = (X_0 + \beta(t)(x - X_0))\sigma(t)$$

Find $\sigma(t)$ and $\beta(t)$ from the minimizing property (Antonov and Misirpashaev, 2006)

$$|\sigma(t)|^2 = E [|\gamma_t|^2]$$

$$\beta(t) = \frac{E [|\gamma_t|^2(x(t) - X_0)]}{2E [|\gamma_t|^2] E [(x(t) - X_0)^2]}$$

Average the shift parameter (Piterbarg, 2005)

$$\bar{\beta}_T = \frac{\int_0^T \beta(t) |\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau dt}{\int_0^T |\sigma(t)|^2 \int_0^t |\sigma(\tau)|^2 d\tau dt}$$

Projection on a displaced diffusion (cont'd)

Price the European option using the Black-Scholes formula

$$\mathbb{E}[(X_T - K)^+] = \frac{X_0}{\bar{\beta}_T} \mathcal{N}(d_+) - \left(K + \frac{X_0(1 - \bar{\beta}_T)}{\bar{\beta}_T} \right) \mathcal{N}(d_-),$$

$$d_{\pm} = \frac{\ln(X_0 / (K\bar{\beta}_T + X_0(1 - \bar{\beta}_T))) \pm V/2}{\sqrt{V}}, \quad V = \bar{\beta}_T^2 \int_0^T |\sigma(t)|^2 dt.$$

Examples of projection on displaced diffusion

All calculations can be completed in the leading order in volatilities for the pricing of European options on the following processes

- basket of equities
- swap rate in a Libor Market Model
- FX rate in a cross-currency Libor Market Model

For details, see Piterbarg (2006), Antonov and Misirpashaev (2006a,b)

Extending the idea of Markovian projection

0. Markovian projection of drift (“Universal Theory of No Volatility”)
1. Markovian projection for a multi-component process with applications to projections onto stochastic volatility models
2. Markovian projection for a jump process with applications to top-down modeling of credit basket loss

Dupire's "Universal Theory of no Volatility" (private communication, unpublished)

A process with stochastic drift $dX_t = \mu_t dt$ and another process with local drift $dY_t = m(Y_t, t)dt$ have the same marginal distributions provided $X_0 = Y_0$ and

$$m(x, t) = \mathbb{E}[\mu_t | X_t = x]$$

The intended application was to model credit basket loss as a continuous variable. We will see later how this changes in a framework with discrete default events.

Markovian projection with multiple components

Take an N -dimensional (non-Markovian) process

$\mathbf{X}(t) = \{X_t^{(1)}, \dots, X_t^{(N)}\}$ with an SDE

$$dX_t^{(n)} = \mu_t^{(n)} dt + \gamma_t^{(n)} \cdot dW_t$$

The process \mathbf{X}_t can be mimicked with a Markovian N -dimensional process \mathbf{Y}_t with the same joint distributions for all components at fixed t .

According to Gyöngy, the process \mathbf{Y}_t satisfies the SDE

$$dY_t^{(n)} = m^{(n)}(\mathbf{Y}_t, t) dt + g^{(n)}(\mathbf{Y}_t, t) \cdot dW_t$$

with

$$\begin{aligned} m^{(n)}(\mathbf{x}, t) &= \mathbb{E}[\mu_t^{(n)} | \mathbf{X}_t = \mathbf{x}] \\ g^{(n)}(\mathbf{x}, t) \cdot g^{(m)}(\mathbf{x}, t) &= \mathbb{E}[\sigma_t^{(n)} \cdot \sigma_t^{(m)} | \mathbf{X}_t = \mathbf{x}] \end{aligned}$$

Choice of process components

The first component is the rate, $dS_t = \Sigma_t \cdot dW_t$. (We set $S_0 = 1$).

The second component should be related to $|\Sigma_t|^2$.

We fix a shift function $\beta(t)$ (for example, from a projection on displaced diffusion) and write the equation for the rate in the form

$$dS_t = (1 + \beta(t)(S_t - 1))\Lambda_t \cdot dW_t$$

where

$$\Lambda_t = \frac{\Sigma_t}{1 + \beta(t)(S_t - 1)}$$

The second equation is for the variance $V_t = |\Lambda_t|^2$,

$$dV_t = \mu_t^V dt + \sigma_t^V \cdot dW_t$$

This completes the SDE's for the non-Markovian pair $\{S_t, V_t\}$.

Projection onto a stochastic volatility model

Target model

$$\begin{aligned} dS_t^* &= (1 + \beta(t)(S_t^* - 1))\sqrt{z_t}\sigma_H(t)\cdot dW_t \\ dz_t &= \theta(t)(1 - z_t)dt + \sqrt{z_t}\sigma_z(t)\cdot dW_t, \quad z_0 = 1 \end{aligned}$$

Answer

$$\begin{aligned} |\sigma_H(t)|^2 &= \mathbf{E}[V_t] \\ \theta(t) &= \frac{d}{dt} (\log \mathbf{E}[V_t]) - \frac{1}{2} \frac{d}{dt} (\log \mathbf{E}[\delta V_t^2]) + \frac{\mathbf{E}[|\sigma_t^V|^2]}{2\mathbf{E}[\delta V_t^2]} \\ |\sigma_z(t)|^2 &= \frac{\mathbf{E}[V_t|\sigma_t^V|^2]}{\mathbf{E}[V_t^2]\mathbf{E}[V_t]} \\ \rho(t) &= \frac{\sigma_t^z \cdot \sigma_t^H}{|\sigma_t^H| |\sigma_t^z|} = \frac{\mathbf{E}[V_t\Lambda_t \cdot \sigma_t^V]}{\sqrt{\mathbf{E}[V_t^2]\mathbf{E}[V(t)|\sigma_t^V|^2]}} \end{aligned}$$

where $\delta V_t = V_t - \mathbf{E}[V_t]$.

From stochastic intensity to local intensity: Gyöngy-Dupire for counting processes

N_t has adapted stochastic intensity λ_t

M_t has local intensity $\Lambda(M, t)$

One-dimensional marginal distributions of N_t and M_t are identical provided $N_0 = M_0$ and

$$\Lambda(M, t) = \mathbf{E}[\lambda_t | N_t = M].$$

(Lopatin and Misirpashaev, 2007). The counterpart of Dupire's formula is

$$\Lambda(M, T) = -\frac{\partial \mathbf{P}[N_T \leq M] / \partial T}{\mathbf{P}[N_T \leq M] - \mathbf{P}[N_T \leq M - 1]}$$

Local intensity model (a.k.a. implied intensity model and 1-step Markov chain)

Forward Kolmogorov equation for the density of loss distribution is easy to solve

$$\frac{\partial p(M, t)}{\partial t} = \Lambda(M - 1, t)p(M - 1, t) - \Lambda(M, t)p(M, t).$$

Local intensity $\Lambda(M, t)$ is directly related to the loss distribution

$$\Lambda(M, t) = -\frac{1}{p(M, t)} \frac{\partial}{\partial t} \sum_{n=0}^M p(n, t)$$

and turns out to bears a clear signature of the correlation skew.

Sketch of intensity averaging formula proof

$$\begin{aligned}
\frac{\partial \mathbb{P}[N_T \leq M]}{\partial T} &= \frac{\partial \mathbb{E}[1_{N_T \leq M}]}{\partial T} = \frac{\partial \mathbb{E}[\mathbb{E}[1_{N_T \leq M} | \{\lambda_\tau\}, 0 \leq \tau \leq T]]}{\partial T} \\
&= \frac{\partial \mathbb{E}\left[\sum_{n=0}^M \frac{1}{n!} e^{-\int_0^T \lambda_\tau d\tau} \left(\int_0^T \lambda_\tau d\tau\right)^n\right]}{\partial T} \\
&= \mathbb{E}\left[\frac{\partial \left(\sum_{n=0}^M \frac{1}{n!} e^{-\int_0^T \lambda_\tau d\tau} \left(\int_0^T \lambda_\tau d\tau\right)^n\right)}{\partial T}\right] \\
&= \mathbb{E}\left[-\lambda_T e^{-\int_0^T \lambda_\tau d\tau} \left(\int_0^T \lambda_\tau d\tau\right)^M\right] \\
&= \mathbb{E}[\mathbb{E}[-\lambda_T 1_{N_T=M} | \{\lambda_\tau\}, 0 \leq \tau \leq T]] \\
&= \mathbb{E}[-\lambda_T 1_{N_T=M}] = -\mathbb{E}[\lambda_T | N_T = M] \mathbb{P}[N_T = M],
\end{aligned}$$

Application of stochastic intensity counting processes to top-down modeling of credit baskets

L_t is a counting process conditional on an adapted stochastic intensity process λ_t . Examples:

Hawkes process

$$d\lambda_t = \kappa(\rho(t) - \lambda_t)dt + \eta dL_t$$

More general affine process (Errais, Giesecke, and Goldberg, 2007)

$$d\lambda_t = \kappa(\rho(t) - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t + \eta dL_t + dJ_t$$

A minimal non-affine model (Lopatin and Misirpashaev, 2007)

$$d\lambda_t = \kappa(\rho(L_t, t) - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$$

Calibration problem: how to recognize whether the model is capable of producing the “correlation skew”?

Correlation skew

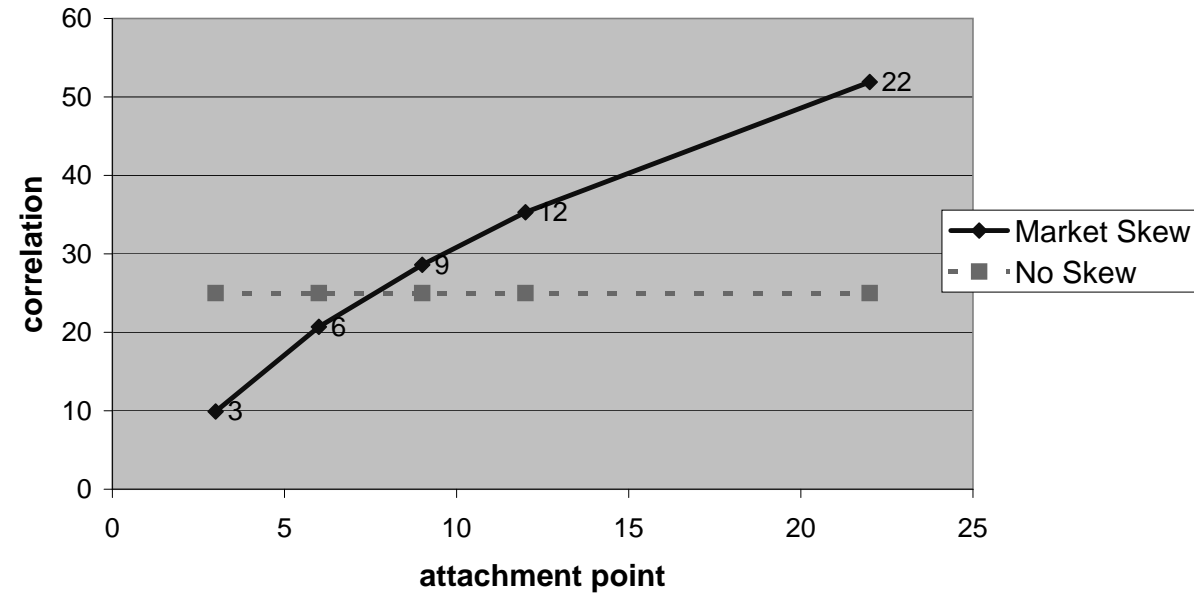


Figure 1: Base correlations skew implied from iTraxx 5y CDO tranches on Oct 12, 2005

Relating intensity to skew is not straightforward

Deterministic intensity $\lambda(t)$ produces no correlations (obvious, no default clustering).

Stochastic intensity λ_t can produce positive default correlations, however the intuition about stronger default clustering does not necessarily result in a stronger skew.

A more reliable indicator is needed to predict the ability of the model to generate skew.

Signature of correlation skew in local intensity

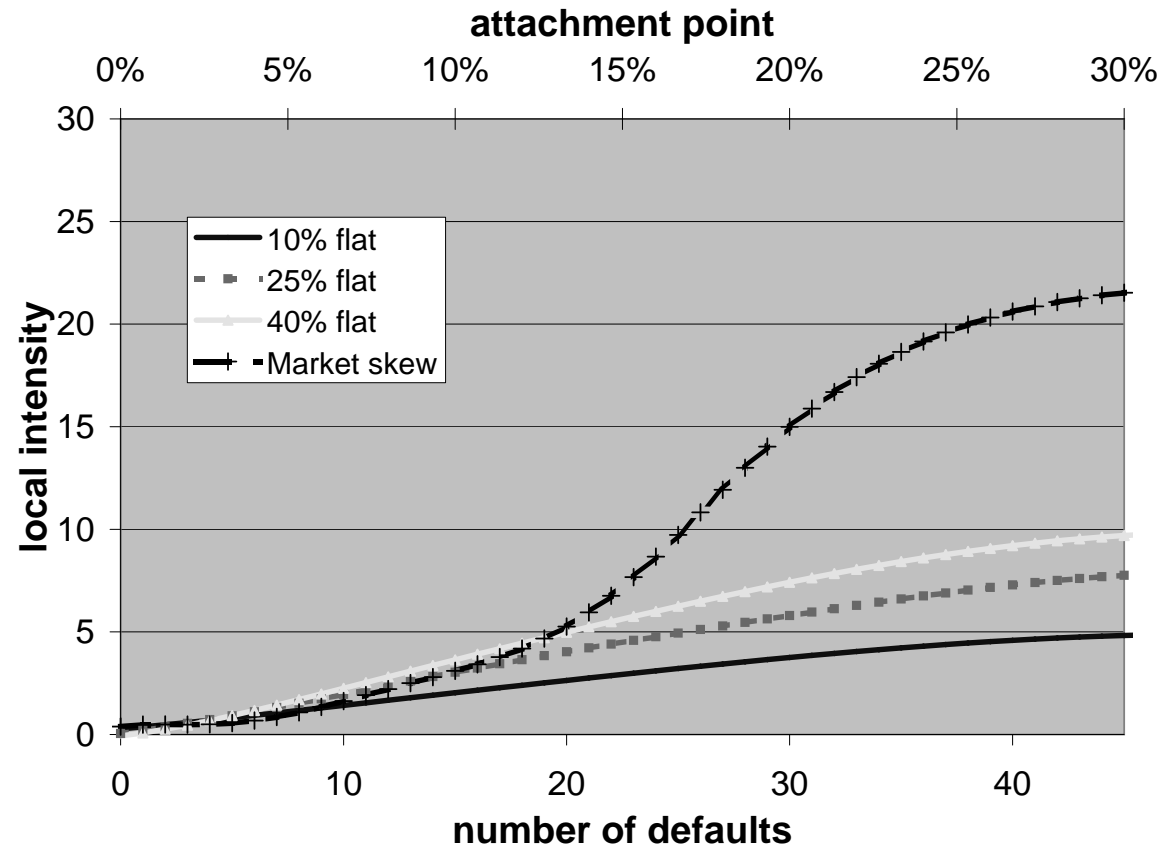


Figure 2: Local intensity consistent with flat Gaussian correlations or market correlations skew for iTraxx 5y (38) on Oct 12, 2005.

Rules of thumb for the local intensity and correlation skew

Flat local intensity \Leftrightarrow no correlations

Sub-linear local intensity \Leftrightarrow flat correlations, no skew

Super-linear local intensity \Leftrightarrow correlation skew

From local intensity to base correlations skew

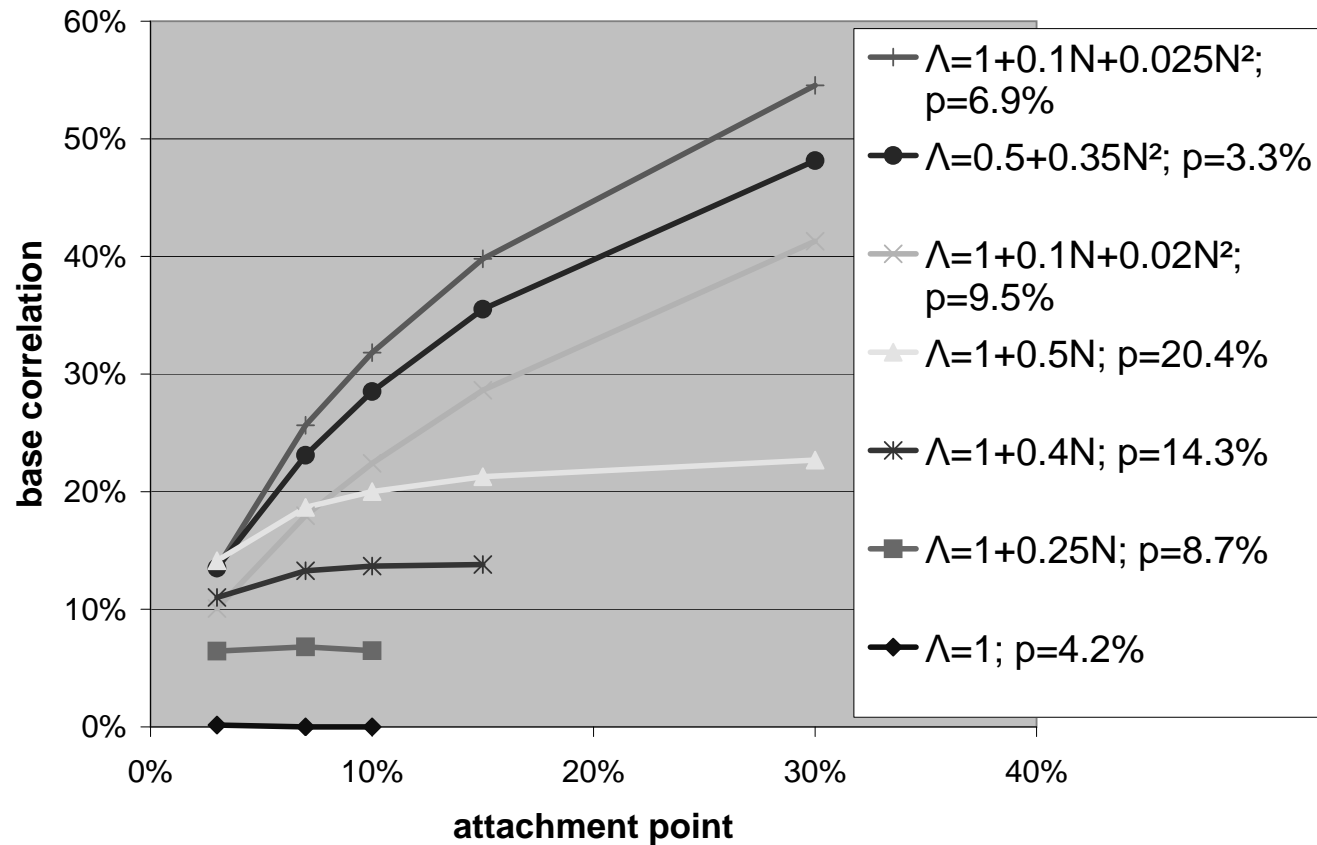


Figure 3: Implied base correlations and expected loss from a given local intensity. The number of assets is assumed to be 125, maturity 5y.

Local intensity in affine models

$$d\lambda_t = \kappa(\rho(t) - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t + \eta dL_t + dJ_t$$

(J_t has intensity $h_0 + h_1\lambda_t$, jump size J)

$$\mathbb{E}[\lambda_T | L_T = L] = \frac{\int p(\lambda, L, T)\lambda d\lambda}{\int p(\lambda, L, T)d\lambda} = -\frac{\partial}{\partial u} \ln \left(\int_0^{2\pi} \frac{dw}{2\pi} e^{-iwL} f(\lambda_0, 0, u, w, 0) \right) \Big|_{u=0}$$

$$f(\lambda, L, u, w, t) = \mathbb{E}[e^{u\lambda_T + iwL_T} | \lambda_t = \lambda, L_t = L]$$

$$\begin{aligned} \frac{\partial f}{\partial t} + \kappa(\rho - \lambda)\frac{\partial f}{\partial \lambda} + \frac{1}{2}\sigma^2\lambda\frac{\partial^2 f}{\partial \lambda^2} + \lambda[f(\lambda + \eta, L + 1, \dots) - f(\lambda, L, \dots)] \\ + (h_0 + h_1\lambda)[f(\lambda + J, L, \dots) - f(\lambda, L, \dots)] = 0. \end{aligned}$$

Local intensity in affine models (cont'd)

(Duffie et al (2000); Giesecke and Goldberg, 2006; Errais et al, 2007)

$$t \mapsto T - t, \quad f = \exp(iwL + a(t) + b(t)\lambda)$$

$$\dot{a}(t) = \kappa\rho b(t) + h_0(e^{Jb(t)} - 1), \quad a(0) = 0,$$

$$\dot{b}(t) = -\kappa b(t) + \frac{1}{2}\sigma^2 b(t)^2 + e^{iw+\eta b(t)} - 1 + h_1(e^{Jb(t)} - 1), \quad b(0) = u.$$

This system is easily solved numerically

Local intensity: Hawkes process

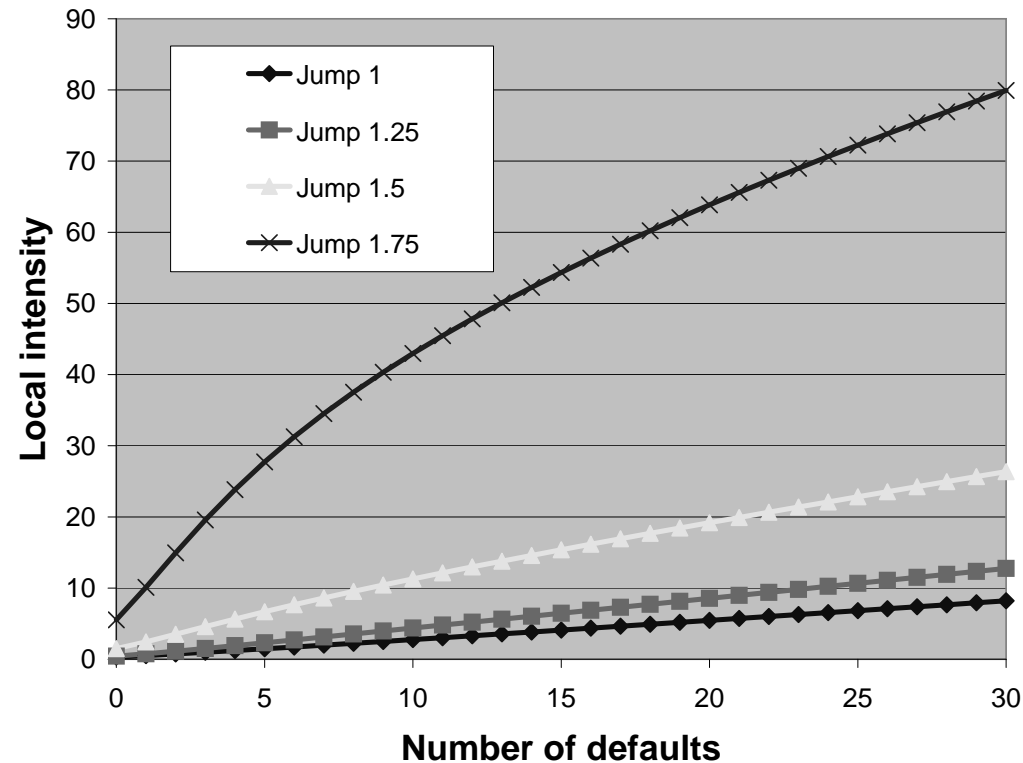


Figure 4: Local intensity of the Hawkes process for different values of the intensity jump upon default. Other parameters are $\kappa = 1$, $\rho = 0.3$, $\lambda_0 = 1$, maturity 5y.

Local intensity from stochastic intensity with jumps

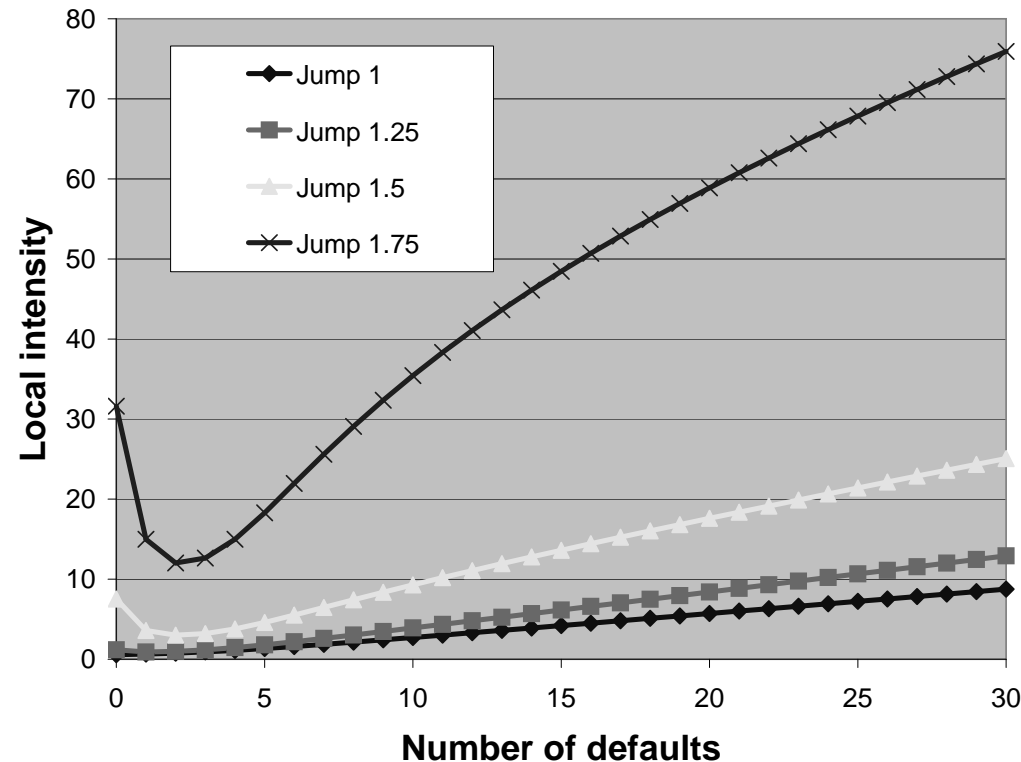


Figure 5: Local intensity from stochastic intensity for different values of the intensity jump upon default. Other parameters are $\kappa = 1$, $\rho = 0.3$, $\lambda_0=1$, $h_0 = 0$, $h_1 = 1$, maturity 5y.

Applicability of top-down affine models is problematic

Local intensity in affine models typically grows slower than linear, hence it will be difficult to count on a good calibration to the tranches.

We assumed a deterministic loss-given-default (LGD). Stochastic LGD might improve the situation only if it is correlated with the loss and or/intensity.

Alternatively, it makes sense to try going beyond the class of affine models.

A minimal non-affine model (Lopatin and Misirpashaev, 2007)

$$d\lambda_t = \kappa(\rho(L_t, t) - \lambda_t)dt + \sigma\sqrt{\lambda_t}dW_t$$

We now have sufficient freedom to calibrate the entire surface of loss distribution by adjusting the free function $\rho(L, t)$ for any volatility σ . (Other forms of the diffusion term are possible, also we could add jumps.)

Calibration of $\rho(L, t)$ to the surface of loss (and the tranches) can be done without simulation.

Instruments dependent on the dynamics can be computed either by a forward simulation or by backward induction.

Conclusions

Dupire's theory of effective volatility gave birth to many non-trivial applications and extensions, including

- closed-form results for a projection on a displaced diffusion
- multi-component generalization and projections on stochastic volatility models
- projection on a Markov chain in the top-down credit basket modeling. The counterpart of the local volatility is local intensity, $\Lambda(N, t) = \mathbb{E}[\lambda_t | N_t = N]$.

References

A. Antonov and T. Misirpashaev (2006a) “Markovian Projection onto a Displaced Diffusion: Generic Formulas with Applications,” available at SSRN: <http://ssrn.com/abstract=937860>.

A. Antonov and T. Misirpashaev (2006b) “Efficient Calibration to FX Options by Markovian Projection in Cross-Currency LIBOR Market Models,” available at SSRN: <http://ssrn.com/abstract=936087>.

M. Avellaneda, D. Boyer-Olson, J. Busca, and P. Friz (2002) “Reconstructing volatility,” *Risk*, October, 87–91.

B. Dupire (1997) “A Unified Theory of Volatility,” Banque Paribas working paper, reprinted in “Derivatives Pricing: the Classic Collection,” edited by P. Carr, Risk Books, London, 2004.

E. Errais, K. Giesecke and L. Goldberg (2007) “Pricing credit from the top down with affine point processes,” working paper, available at defaultrisk.com.

K. Giesecke and L. Goldberg (2005) “A top down approach to multi-name credit,” working paper, available at defaultrisk.com.

I. Gyöngy (1986) “Mimicking the One-Dimensional Marginal Distributions of Processes Having an Ito Differential” *Probability Theory and Related Fields*, 71, 501–516.

A. V. Lopatin and T. Misirpashaev (2007) “Two-Dimensional Markovian Model for Dynamics of Aggregate Credit Loss,” working paper, available at defaultRisk.com

V. Piterbarg (2005) “Time to smile,” *Risk*, May

V. Piterbarg (2006) “Markovian Projection Method for Volatility Calibration,” available at SSRN: <http://ssrn.com/abstract=906473>.