Towards a multi-stochastic volatility model

for CMS spread exotics

Vladimir Piterbarg
Barclays Capital

1 Plan

- Anything CMS-spread-linked is currently very popular
- Interested in modelling for CMS spread exotics
  - Term structure models
  - Not copulas
- Are current models for CMS spread exotics adequate?
- If not, what features need to be added?
- **How?**

2 CMS spread options

- Let $S_i(t)$, $i = 1, 2$, be two forward swap rates. Both fix at time T but have different tenors (for example, a 10y and a 2y)
- For example paying
$$(S_1(T) - S_2(T) - K)^+$$
or
$$1_{\{S_1(T)-S_2(T)>K\}}$$
at time T_p, $T_p \geq T$.
- Also very popular as part of exotics ie callable CMS spread swaps, CMS spread range accrual TARNs, etc
- One-dimensional distributions of individual swap rates are easily implied from prices of swaptions.
- Main driver of value of spread-linked securities – dependence between swap rates. Not observable in the market.

3 Implied normal (basis point) volatilities

- It is convenient to express prices of options in implied volatilities
- Spread can go negative, so cannot use implied Black volatility. Use implied Normal (Basis Point, or BP) volatility
- The "implied spread volatility" is the volatility σ that needs to be used in a Normal model in which the spread is treated as the underlying to recover a spread option price,
$$d\left(S^1 - S^2\right) = \sigma\, dW.$$
- "Spread smile" is the dependence of spread volatility σ on the strike of the spread option.
- Spread smile is a way to describe the distribution of the spread

4 Libor market model

- LMM a standard choice for exotics, in particular spread-based for which low-dimensional Markovian models are hard to use
- Recall skew-extended LMM, here $\{L_n(t)\}$ are spanning Libor rates,
$$dL_n(t) = \ldots dt + \varphi_n(L_n(t))\,\sigma_n(t)\,dW_n(t), \quad n = 1, \ldots, N.$$
- Here
$$\langle dW_n(t), dW_m(t)\rangle = \rho_{nm}\,dt$$
- (Libor) volatilities are implied from market prices of caps and swaptions, correlations – usually historically estimated

5 Asset-based model for the spread

- In LMM, swap rates (approximately) follow the same dynamics as the Libor rates
- Hence, the implied dynamics of swap rates in LMM is given by
$$dS_i(t) = \ldots dt + \varphi_i(S_i(t))\,dW_i(t), \quad i = 1, 2, \quad \langle dW_1, dW_2\rangle = \rho$$
(drifts because no measure under which both swap rates are martingales)
- Can also be used as a simple, stand-alone "asset-based" model (local volatility)
- Can estimate ρ historically
- Simple model: implied volatilities of swap rates + historically estimated correlation

6 Smile extensions

- The importance of incorporating volatility smiles well-understood
- Typical choice (for exotics), LMM with stochastic volatility ([ABR01], [AA02], [Pit04])
$$dL_n(t) = \ldots dt + \varphi_n(L_n(t))\sqrt{z(t)}\sigma_n(t)\,dW_n(t), \quad n = 1, \ldots, N,$$
where
$$dz(t) = \theta(1 - z(t))\,dt + \eta\sqrt{z(t)}\,dZ(t).$$
- Implies the following dynamics for swap rates
$$dS_i(t) = \ldots dt + \varphi_i(S_i(t))\sqrt{z(t)}\,dW_i(t), \quad i = 1, 2, \quad \langle dW_1, dW_2\rangle = \rho.$$
- The latter can also be used as a stand-alone asset-based model
- Can choose SV parameters and skew functions $\varphi_i(\cdot)$ to match swaption prices across strikes for both swap rates.
- Parameter choice same as before – everything but ρ is implied from the market, ρ historically estimated
- Are smile effects fully accounted for with such a model?

7 Variance decorrelation

- Recall
$$dS_i(t) = \ldots dt + \varphi_i(S_i(t))\sqrt{z(t)}\,dW_i(t), \quad \langle dW_1, dW_2\rangle = \rho.$$
- Notice that the same SV process, $z(\cdot)$, is used for both swap rates. A problem?
- Let us look at an extension
$$\begin{aligned} dz_i(t) &= \theta_i(1 - z_i(t))\,dt + \eta_i\sqrt{z(t)}\,dZ_i(t), \quad z_i(0) = 1, \qquad (1)\\ dS_i(t) &= \varphi_i(S_i(t))\sqrt{z_i(t)}\,dW_i(t), \\ \langle dW_1, dW_2\rangle &= \rho, \quad \langle dZ_1, dZ_2\rangle = \chi, \quad \langle dZ_i, dW_j\rangle = \gamma_{ij}, \\ \lambda_i &= \varphi_i\left(S^i(0)\right)/S^i(0), \quad \beta_i = \varphi_i'(S_i(0))/\lambda_i. \end{aligned}$$
- What is the effect of "variance decorrelation" ($\chi \neq 0$) on CMS spread options?

8 Variance correlation effect

- Naively, variance correlation χ will change the curvature of the spread smile, as it seems that χ affects the volatility of the variance of the difference of the two underlyings
- Expecting the effect of χ to be relatively minor
- **However**, both these "conclusions" are wrong!
  - Main impact of χ is on the overall level of the spread smile, much like ρ
  - The effect has the same order of magnitude as ρ, ie not minor at all

9 Variance correlation effect, cont

- Parameters typical for CMS 10y - CMS 2y in 10y years
- Calibrated stochastic volatility parameters, linear skew

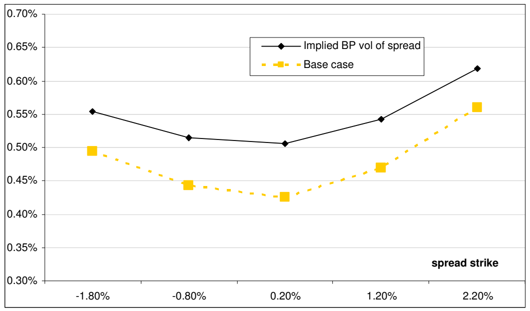
$$\varphi_i\left(S_i\left(t\right)\right)=\lambda_i\left(\beta_i S_i\left(t\right)+\left(1-\beta_i\right)S_i\left(0\right)\right).$$

Udl	Fwd $S_i(0)$	Vol λ_i	Skew β_i	Mean rev θ	Vol of var η_i	Spot/vol correl γ_{ii}
S_1	4.60%	17%	100%	10%	90%	-25%
S_2	4.30%	15%	70%	10%	80%	-25%

- $\rho = 80\%$

10 Variance correlation effect, cont

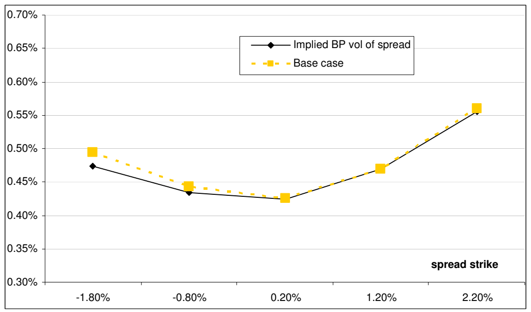
- Base case $\chi = 100\%$, vs. $\chi = 80\%$



- significant move in implied BP vols of the spread
- Equivalent to keeping χ at 100% but moving ρ from 80% to 70%

11 Variance correlation effect, cont

- Again consider the case $\chi = 100\%$, vs. $\chi = 80\%$
- If we compensated for the overall level (by adjusting ρ to 90%), what is the effect of χ?



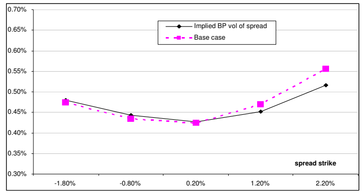
- (Smallish) shape impact of χ

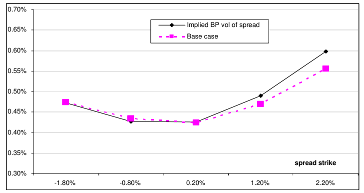
12 Effect of other parameters

- With χ having such an impact, what about other parameters?
- Use $\rho = 90\%$, $\chi = 80\%$ $\gamma_{ij} = -25\%$, $i, j = 1, 2$ as new "base case".
- Always adjust ρ to match ATM volatility of the spread (to study pure shape effects)

13 Spot 1 / Vol 1 correlation

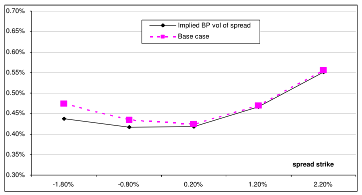
- γ_{11} is implied from the options on S_1 so technically not a "free" parameter. But can "compensate" with skew β_1
- $\gamma_{11} = -35\%$ ($\rho = 88\%$) and $\gamma_{11} = -15\%$ ($\rho = 91\%$)

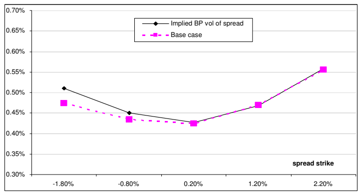




14 Spot 2 / Vol 2 correlation

- $\gamma_{22} = -35\%$ ($\rho = 90\%$) and $\gamma_{22} = -15\%$ ($\rho = 90\%$)

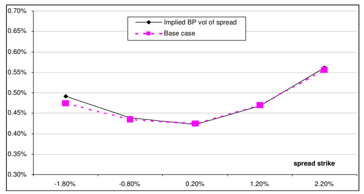


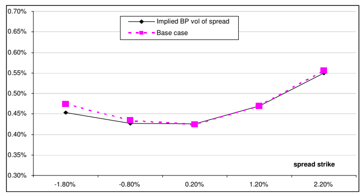


- Interestingly, γ_{11} affects high strikes of the spread, γ_{22} affects low strikes. In a sense, they affect the slope and *curvature* of the smile

15 Spot 1 / Vol 2 correlation

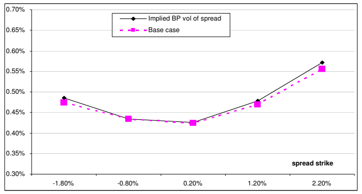
- γ_{12} and γ_{21} are "free" parameters, ie not implied by European swaption markets. Can be used to purely control the smile of the spread
- $\gamma_{12} = -35\%$ ($\rho = 91\%$) and $\gamma_{12} = -15\%$ ($\rho = 89\%$)

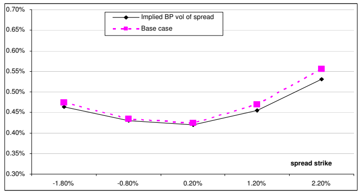




16 Spot 2 / Vol 1 correlation

- $\gamma_{21} = -35\%$ ($\rho = 90\%$) and $\gamma_{21} = -15\%$ ($\rho = 90\%$)

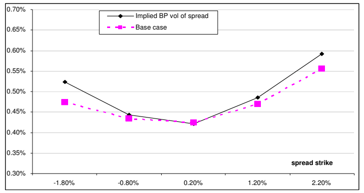


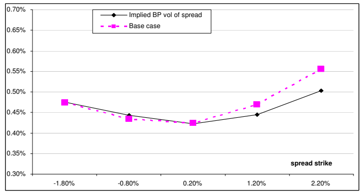


- Affects both low and high strikes, ie curvature effect

17 Spot 2 / Vol 1 correlation

- The effects of γ_{12} and γ_{21} are generally mild. Let us consider some extreme cases, $\gamma_{21} = -50\%$ and $\gamma_{21} = 0\%$ (min and max value allowed in the correlation matrix)
- $\gamma_{21} = -50\%$ ($\rho = 89\%$) and $\gamma_{21} = 0\%$ ($\rho = 90\%$)





18 Summary of observations

- Impact of unobservable parameters
  - Vol/Vol correlation (χ) mostly affects the level of the spread smile. Important for any spread option-linked exotic
  - Spot 1/Vol 2 (γ_{12}) and Spot 2/Vol 1 (γ_{21}) correlations affect slope and curvature of the spread smile. Important for non-ATM and (especially!) digital options on CMS spread
- A "simple" SV LMM (when a "projection" on two swap rates is considered) does not have these parameters
  - Single variance process $z(\cdot)$,
  - Also typically spot/vol correlation is 0
- Goals of term structure modelling for CMS spread exotics:
  - Account for variance decorrelation effects in pricing CMS spread exotics (ie have $\chi \neq 1$ in a model)
  - Have "knobs" to tweak CMS spread smiles
- LMM with multiple stochastic variance factors?

19 Approximation methods for multi-SV models

- To have a usable multi-SV model, need efficient approximation methods for calibration
- Let us start with a "simple" case. Consider the simple model as above and try to derive analytic approximations to values of (European) CMS spread options
- Rewrite the model in more uniform notation

$$dS_i(t) = \varphi_i(S_i(t))\sqrt{z_i(t)}\,dW_i(t), \tag{2}$$
$$dz_i(t) = \theta(1 - z_i(t))\,dt + \eta_i\sqrt{z_i(t)}\,dW_{2+i}(t), \quad z_i(0) = 1,$$
$$i = 1, 2,$$

  with the correlations given by

$$\langle dW_i(t), dW_j(t)\rangle = \rho_{ij} \quad i, j = 1, \ldots, 4.$$

- Denote

$$p_i = \varphi_i(S_i(0)), \quad q_i = \varphi_i'(S_i(0)).$$

- Use "almost linear" $\varphi_i(\cdot)$, ie linear or CEV

20 The idea

- Main idea:
  - Write $S(t) = S_1(t) - S_2(t)$ for the spread. We want to approximate the dynamics of $S(\cdot)$ by a model of the type (2), ie
$$dS(t) = \varphi(S(t))\sqrt{z(t)}\,dW(t), \tag{3}$$
    where $\varphi(\cdot)$, $W(\cdot)$, and $z(t)$ are some, to be identified, skew function, Brownian motion, and stochastic variance process *of the spread.*
  - The process $z(\cdot)$ to be written in the mean-reverting square-root form, ie like Heston
  - Then options on $S(t)$ can be valued by the shifted Heston formula (after linearizing $\varphi(\cdot)$)

21 The idea, cont

Theorem (Gyongy 86, Dupire 97) Let $X(t)$ be given by
$$dX(t) = \alpha(t)\,dt + \beta(t)\,dW(t), \tag{4}$$
where $\alpha(\cdot)$, $\beta(\cdot)$ are adapted bounded stochastic processes such that (4) admits a unique solution. Define $a(t,x)$, $b(t,x)$ by
$$a(t,x) = \mathsf{E}(\alpha(t)|\, X(t) = x),$$
$$b^2(t,x) = \mathsf{E}\left(\beta^2(t)\middle|\, X(t) = x\right),$$
Then the SDE
$$dY(t) = a(t, Y(t))\,dt + b(t, Y(t))\,dW(t), \qquad Y(0) = X(0), \tag{5}$$
admits a weak solution $Y(t)$ that has the same one-dimensional distributions as $X(t)$.

- See [Gyö86], [Dup97]

22 The idea, cont

Remark 1 Since $X(\cdot)$ and $Y(\cdot)$ have the same one-dimensional distributions, the prices of European options on $X(\cdot)$ and $Y(\cdot)$ for all strikes K and expiries T will be the same. Thus, for the purposes of European option valuation and/or calibration to European options, we can replace a potentially very complicated process $X(\cdot)$ with a much simpler Markov process $Y(\cdot)$, which we call the *Markovian projection* of $X(\cdot)$.

Remark 2 The process $Y(\cdot)$ follows what is known as a "local volatility" process. The function $b(t,x)$ is often called "Dupire's local volatility"

23 The idea, cont

- Any process (including a stochastic volatility one) can be replaced by a local volatility process for the purposes of European option valuation
- Requires calculations of conditional expected values. This is the hard bit. Approximations often necessary
- In approximations, better to replace "like for like". Replace a (complicated) SV model with a (simpler) SV model.
  - Approximations to conditional expected values may be simpler
  - Errors of approximations will tend to "cancel out"
- Gyongy-Dupire theorem still works

Corollary If two processes have the same Dupire's local volatility, the European option prices on both are the same for all strikes and expiries

24 The idea, cont

- Let $X_1(t)$ follow
$$dX_1(t) = b_1(t, X_1(t)) \sqrt{\zeta_1(t)}\, dW(t),$$
where $\zeta_1(t)$ is some variance process.
- We would like to match the European option prices on $X_1(\cdot)$ (for all expiries and strikes) in a model of the form
$$dX_2(t) = b_2(t, X_2(t)) \sqrt{\zeta_2(t)}\, dW(t),$$
where $\zeta_2(t)$ is a different, and potentially simpler, variance process.
- Then the Corollary and the Theorem imply that we need to set
$$b_2^2(t, x) = b_1^2(t, x) \frac{\mathsf{E}(\zeta_1(t)|\, X_1(t) = x)}{\mathsf{E}(\zeta_2(t)|\, X_2(t) = x)}. \tag{6}$$
- Error cancellation – whatever approximations are used for conditional expected values in (6), hopefully they will tend to cancel when we take the ratio
- For CMS spread, X_1 is the actual spread process $S(\cdot)$ (implied by (2)), and X_2 is the approximation (3)

25 The method

- Write down $dS(\cdot)$ for $S = S_1 - S_2$ given by (2)
- Identify a suitable "spread variance" process $z(\cdot)$
- Compute the skew function $\varphi(\cdot)$ of the spread using the Markovian projection ideas above
- "Massage" $z(\cdot)$ into the Heston form

26 Process for the spread

- We have
$$dS_i(t) = \varphi_i(S_i(t)) \sqrt{z_i(t)}\, dW_i(t),$$
- $S = S_1 - S_2$, then
$$dS(t) = \sigma(t)\; dW(t), \tag{7}$$
where
$$\begin{aligned}
\sigma^2(t) &= (\varphi_1(S_1(t))\, u_1(t))^2 - 2(\varphi_1(S_1(t))\, u_1(t))(\varphi_2(S_2(t))\, u_2(t))\,\rho_{12} \\
&\quad + (\varphi_2(S_2(t))\, u_2(t))^2, \\
dW(t) &= \frac{1}{\sigma(t)}(\varphi_1(S_1(t))\, u_1(t)\; dW_1(t) - \varphi_2(S_2(t))\, u_2(t)\; dW_2(t)), \\
u_i(t) &= \sqrt{z_i(t)}, \quad i = 1, 2.
\end{aligned}$$

27 Process for the variance of the spread

- Try to find a stochastic volatility process $z(\cdot)$ such that the curvature of the smile of the spread $S(\cdot)$ is explained by it, and the local volatility function is only used to induce the volatility skew
- To identify a suitable candidate for $z(\cdot)$, consider what the expression for $\sigma^2(t)$ would be if $\varphi_i(x)$, $i = 1, 2$, were constant functions.
- In this case, the expression for $\sigma^2(t)$ above would not involve the processes $S_i(\cdot)$, $i = 1, 2$ and this is a good candidate for the stochastic variance process.
- We define (the division by $\sigma^2(0)$ is to preserve the scaling $z(0) = 1$)
$$z(t) = \frac{1}{p^2}\left((p_1 u_1(t))^2 - 2 p_1 p_2 u_1(t)\, u_2(t)\, \rho_{12} + (p_2 u_2(t))^2\right), \tag{8}$$
where
$$p = \sigma(0) = \left(p_1^2 - 2 p_1 p_2 \rho_{12} + p_2^2\right)^{1/2}. \tag{9}$$

28 Skew function of the spread

- By Corollary,

$$\varphi^2(t,x) = \frac{\mathsf{E}\left(\sigma^2(t)\middle|\, S(t)=x\right)}{\mathsf{E}\left(z(t)\middle|\, S(t)=x\right)}. \tag{10}$$

- The expression for $\mathsf{E}\left(\sigma^2(t)\middle|\, S(t)=x\right)$ is a linear combinations of the conditional expected values of the terms

$$\varphi_i(S_i(t))\,\varphi_j(S_j(t))\,u_i(t)\,u_j(t),$$

- Approximate to the first order by

$$p_ip_j\left(1+\frac{q_i}{p_i}(S_i(t)-S_i(0))+\frac{q_j}{p_j}(S_j(t)-S_j(0))+(u_i(t)-1)+(u_j(t)-1)\right).$$

- Use Gaussian approximation to compute conditional expected values

29 Gaussian approximation

- Use $\bar{X}$ to denote a Gaussian approximation to X for a generic X, then

$$\begin{aligned}\mathsf{E}\left(S_i(t)-S_i(0)\middle|\, S(t)=x\right) &\approx \mathsf{E}\left(\bar{S}_i(t)-\bar{S}_i(0)\middle|\, \bar{S}(t)=x\right)\\ \mathsf{E}\left(u_i(t)-1\middle|\, S(t)=x\right) &\approx \mathsf{E}\left(\bar{u}_i(t)-1\middle|\, \bar{S}(t)=x\right),\end{aligned}$$

- Here (we ignore dt terms for du, although they may be included for more accurate approximations)

$$\begin{aligned}d\bar{S}_i(t) &= p_i\,dW_i(t), \quad d\bar{S}(t)=p\,d\bar{W}(t),\\ d\bar{u}_i(t) &= \frac{\eta_i}{2}\,dW_{2+i}(t), \quad d\bar{W}(t)=\frac{1}{p}\left(p_1\,dW_1(t)-p_2\,dW_2(t)\right).\end{aligned} \tag{11}$$

- Then

$$\begin{aligned}\mathsf{E}\left(\bar{S}_i(t)-\bar{S}_i(0)\middle|\, \bar{S}(t)=x\right) &= \frac{p_i\rho_i}{p}(x-S(0)),\\ \mathsf{E}\left(\bar{u}_i(t)-1\middle|\, \bar{S}(t)=x\right) &= \frac{\eta_i\rho_{2+i}}{2p}(x-S(0)),\end{aligned}$$

- We have denoted $\rho_i \triangleq \left\langle d\bar{W}(t), dW_i(t)\right\rangle/dt$, so that $\rho_i=\frac{1}{p}(p_i\rho_{i1}-p_2\rho_{i2})$, $i=1,\ldots,4$.

30 Skew function of the spread

- Combining the results, we get the following approximation to the spread dynamics,

$$dS(t)=\varphi(S(t))\sqrt{z(t)}\,dW(t),$$

- Here $\varphi(x)$ is a function of the same type as $\varphi_i(x)$ (linear or CEV) with

$$\varphi(S(0))=p, \quad \varphi'(S(0))=q,$$

where

$$\begin{aligned}p &= \left(p_1^2-2p_1p_2\rho_{12}+p_2^2\right)^{1/2}\\ q &\triangleq \frac{1}{p}\left(p_1\rho_1^2q_1-p_2\rho_2^2q_2\right).\end{aligned}$$

31 Variance process for the spread

- The process for S is in a nice form. But z is not:

$$z(t)=\frac{1}{p^2}\left(p_1^2z_1(t)-2p_1p_2\sqrt{z_1(t)\,z_2(t)}\rho_{12}+p_2^2z_2(t)\right).$$

- Compute dz,

$$\begin{aligned}dz(t) &= \delta_1(t)\;dt+\delta_2(t)\;dt+\delta_3(t)\;dt\\ &\quad+\xi_1(t)\;dW_3(t)+\xi_2(t)\;dW_4(t),\end{aligned}$$

- dW terms

$$\begin{aligned}\xi_1(t) &= \eta_1\frac{p_1^2}{p^2}\left(\sqrt{z_1(t)}-\frac{p_2}{p_1}\rho_{12}\sqrt{z_2(t)}\right),\\ \xi_2(t) &= \eta_2\frac{p_2^2}{p^2}\left(\sqrt{z_2(t)}-\frac{p_1}{p_2}\rho_{12}\sqrt{z_1(t)}\right).\end{aligned}$$

- dt terms

$$
\begin{aligned}
\delta_1(t) &= \theta \frac{p_1^2}{p^2}\left(1-\frac{p_2}{p_1}\rho_{12}\sqrt{\frac{z_2(t)}{z_1(t)}}\right)(1-z_1(t)),\\
\delta_2(t) &= \theta \frac{p_2^2}{p^2}\left(1-\frac{p_1}{p_2}\rho_{12}\sqrt{\frac{z_1(t)}{z_2(t)}}\right)(1-z_2(t)),\\
\delta_3(t) &= \frac{p_1 p_2 \rho_{12}}{4p^2}\left(\sqrt{\frac{z_2(t)}{z_1(t)}}\eta_1^2 - 2\eta_1\eta_2\rho_{34} + \sqrt{\frac{z_1(t)}{z_2(t)}}\eta_2^2\right).
\end{aligned}
$$

- Complicated expression, Not "closed" in $z(\cdot)$

32 Variance process for the spread

- The curvature of the volatility smile (of options on $S(\cdot)$) is driven by the variance of the stochastic variance
- It is preserved under the Markovian projection of $z(\cdot)$ so can apply the Theorem again, now to the process for $z(\cdot)$!
- Formulas getting unwieldy: need to compute conditional expected values of the type $\mathsf{E}\left(\sqrt{z_i(t)\,z_j(t)}\,\middle|\, z(t)=x\right)$ and $\mathsf{E}\left(\sqrt{z_i(t)/z_j(t)}\,\middle|\, z(t)=x\right)$, for which we would apply the Gaussian approximations
- Try something simpler:
  - replace $\sqrt{z_1(t)}$, $\sqrt{z_2(t)}$ in the dW terms with $\sqrt{z(t)}$;
  - replace $\sqrt{\frac{z_2(t)}{z_1(t)}}$, $\sqrt{\frac{z_1(t)}{z_2(t)}}$ in dt terms with 1.

33 Variance process for the spread, simple approximation

- $\delta_1(t)+\delta_2(t)$ becomes $\theta(1-z)$,
- $\delta_3(t)$ becomes

$$
\gamma \triangleq \frac{p_1 p_2 \rho_{12}}{4p^2}\left(\eta_1^2 - 2\eta_1\eta_2\rho_{34} + \eta_2^2\right). \qquad (12)
$$

- The dW terms can be re-written as $\eta\sqrt{z(t)}\,dB(t)$, where

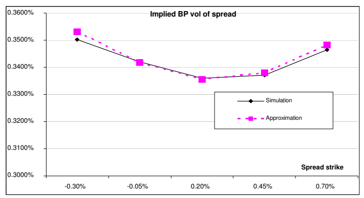
$$
\begin{aligned}
\eta^2 &= \frac{1}{p^2}\left((p_1\eta_1\rho_1)^2 - 2(p_1\eta_1\rho_1)(p_2\eta_2\rho_2)\rho_{34} + (p_2\eta_2\rho_2)^2\right),\\
dB(t) &= \frac{1}{\eta}\left(p_1\eta_1\rho_1\, dW_3(t) - p_2\eta_2\rho_2\, dW_4(t)\right).
\end{aligned}
$$

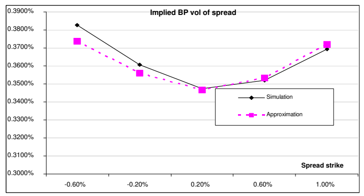
- Altogether

$$
\begin{aligned}
dS(t) &= \varphi(S(t))\sqrt{z(t)}\; dW(t),\\
dz(t) &= \theta\left(1+\frac{\gamma}{\theta} - z(t)\right) dt + \eta\sqrt{z(t)}\, dB(t).
\end{aligned}
$$

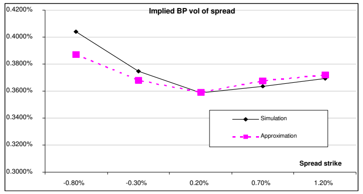
- Linearize φ and apply Heston valuation formula to options on the spread S!

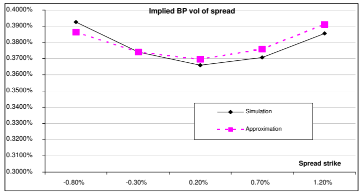
34 Test results, expiry 1y and 3y





35 Test results, expiry 5y and 10y





36 Improving the approximations

- Improve by approximating various terms more accurately.
- Recall that we approximated $\sqrt{\frac{z_2(t)}{z_1(t)}}$, $\sqrt{\frac{z_1(t)}{z_2(t)}}$, $\sqrt{z_1(t)\, z_2(t)}$ with 1. Something more accurate?
- Use lognormal approximation,

$$z_i \approx \exp\left(\xi_i - \frac{1}{2}\langle \xi_i \rangle\right), \quad i = 1, 2,$$

  where the covariance matrix of the Gaussian vector (ξ_1, ξ_2) is chosen to match variances and covariances of (z_1, z_2).
- Then computing $\mathsf{E}\sqrt{z_1(t)\, z_2(t)}$, etc is trivial, eg

$$\mathsf{E}\sqrt{z_1(t)\, z_2(t)} = \exp\left(-\frac{1}{8}\langle \xi_1 - \xi_2 \rangle\right).$$

- Handle time-dependence in coefficients by averaging techniques (see [Pit05b], [Pit05a])
- Can also use Avellaneda's approach (see [ABOBF02]) to get the local volatility component more accurately

37 Forward Libor model with multiple stochastic volatilities

- Have shown that to fully account for smile effects in CMS spread exotics, need to relax the "single stochastic variance" feature
- Also have shown that we can actually do that, ie have tools to handle multiple stochastic volatility drivers
- How to incorporate this in a term structure model?
- "Standard" SV-FLM

$$dL_n(t) = \ldots dt + \varphi_n(L_n(t))\sqrt{z(t)}\sigma_n(t)\, dW_n(t), \quad n = 1, \ldots, N,$$

  where

$$\langle dW_n(t), dW_m(t)\rangle = \rho_{nm}\, dt.$$

- Call the model "multi-stochastic-variance", or MSV, FLM

38 Choices

- Each Libor rate has its own stochastic variance process,

$$dL_n(t) = \ldots\sqrt{z_n(t)}\sigma_n(t)\, dW_n(t)$$

  - Pros: very flexible
  - Cons: very flexible
- Two- or multi-factor process for $z(\cdot)$ (eg a two-factor affine process)
  - May not give the desired variance decorrelation effect if z is the same for all rates (even if z is multi-factor)
- "Factor" structure imposed on stochastic variance
  - Conceptually, can do a PCA on the implied swaption volatility matrix, and use first few components to write a model
  - Libor rate volatility is a linear combination of z_i's (volatility factors) with some weights (loadings)
  - Reasonably parsimonious. Some parameters (loadings?) can potentially be historically estimated
  - In practice will use two volatility factors

39 MSV-FLM, specification

- Define two SV processes

$$dz^i(t) = \theta^i\left(1 - z^i(t)\right)\, dt + \varepsilon^i\sqrt{z^i(t)}\, dZ^i(t), \quad i = 1, 2. \qquad (13)$$

- The Brownian motions assumed independent,

$$\left\langle dZ^1(t), dZ^2(t)\right\rangle = 0.$$

- In general, in the MSV-FLM model the Brownian motions driving the stochastic variance processes will no longer be uncorrelated to the Brownian motions driving the rates
- Important to specify which measure the dynamics are specified under. In particular, we assume that (13) is specified under the spot Libor measure P.

40 MSV-FLM, specification

- Should we use

$$dL_n(t)/\varphi(L_n(t)) = \sigma_n(t)\sqrt{a_n^1(t)\, z^1(t) + a_n^2(t)\, z^2(t)}\, dW_n(t) \qquad (14)$$

  for Libor dynamics?
- Inconvenient technically.
  - Recall that swaption volatility approximations are typically based on representing a swap rate as a linear combination of Libors.
  - The functional form $\sqrt{a_n^1(t)\, z^1 + a_n^2(t)\, z^2}$ is not "closed" under "quadratic form". Typically like to have swap rates follow the same type of an SDE as the Libor rate.
  - Also, (14) is not affine (cross terms not linear in z_1, z_2), Useful to have
- DO NOT use (14)

41 MSV-FLM, specification

- Use affine form of the factor approach
- Two indep copies of Brownian motions, $\left\{W_n^i(t)\right\}$, $i = 1, 2$, with correlations

$$\begin{aligned}\left\langle dW_n^i(t), dW_m^i(t)\right\rangle &= \hat{\rho}_{nm}\, dt, \quad n, m = 1, \ldots, N-1, \quad i = 1, 2\\ \left\langle dW_n^1(t), dW_m^2(t)\right\rangle &= 0.\end{aligned}$$

- Define MSV-FLM by

$$\begin{aligned}dL_n(t)/\varphi(L_n(t)) = \sqrt{z^1(t)}\sigma_n^1(t)\left(dW_n^1(t) + \mu_n^1(t)\, dt\right)\\ + \sqrt{z^2(t)}\sigma_n^2(t)\left(dW_n^2(t) + \mu_n^2(t)\, dt\right),\end{aligned}$$

- Here $\left\{\sigma_n^i(t)\right\}$, $i = 1, 2$ are the two volatility structures that correspond to the two stochastic volatility factors $z^1(t)$ and $z^2(t)$,
- $\mu_n^i(t)$, $i = 1, 2$, are the no-arbitrage drifts.
- The following correlations are imposed between dZ's and dW's,

$$\left\langle dW_n^i(t), dZ^i(t)\right\rangle = \chi_n^i, \quad n = 1, \ldots, N-1, \quad i = 1, 2,$$

- Also (affine)

$$\left\langle dW_n^1(t), dZ^2(t)\right\rangle = \left\langle dW_n^2(t), dZ^1(t)\right\rangle = 0.$$

42 Connection to SV-FLM

- MSV FLM a proper extension of the SV-FLM model: if we set $\lambda_n^2(t) \equiv 0$ and $\chi_n^1 \equiv 0$, then the SV-FLM model is recovered,

$$\begin{aligned}dL_n(t)/\varphi(L_n(t)) &= \cdots + \sqrt{z(t)}\sigma_n(t)\, dW_n(t),\\ \langle dW_n(t), dW_m(t)\rangle &= \rho_{nm}\, dt.\end{aligned}$$

- Consider the instantaneous covariance structure of the two models. In the SV-FLM,

$$\begin{aligned}c_{nm}^{\mathrm{SV}} &\triangleq \langle dL_n(t)/\varphi(L_n(t)), dL_m(t)/\varphi(L_m(t))\rangle\\ &= z(t)\,\sigma_n(t)\,\sigma_m(t)\,\rho_{nm}\, dt,\end{aligned}$$

  and in the MSV-FLM,

$$\begin{aligned}c_{nm}^{\mathrm{MSV}} &\triangleq \langle dL_n(t)/\varphi(L_n(t)), dL_m(t)/\varphi(L_m(t))\rangle\\ &= \left(z_1(t)\,\sigma_n^1(t)\,\sigma_m^1(t) + z_2(t)\,\sigma_n^2(t)\,\sigma_m^2(t)\right)\hat{\rho}_{nm}\, dt.\end{aligned}$$

- In the zero-stochastic-volatility case $(\eta, \eta^1, \eta^2 = 0)$

$$\begin{aligned}c_{nm}^{\mathrm{SV}} &= \sigma_n(t)\,\sigma_m(t)\,\rho_{nm}\, dt,\\ c_{nm}^{\mathrm{MSV}} &= \left(\sigma_n^1(t)\,\sigma_m^1(t) + \sigma_n^2(t)\,\sigma_m^2(t)\right)\hat{\rho}_{nm}\, dt.\end{aligned}$$

43 Connection to SV-FLM

- To match the instantaneous variance of each Libor we must choose
$$\left(\sigma_n^1(t)\right)^2 + \left(\sigma_n^2(t)\right)^2 = (\sigma_n(t))^2, \tag{15}$$
- To match instantaneous correlations of Libor rates we much choose
$$\frac{\sigma_n^1(t)\,\sigma_m^1(t) + \sigma_n^2(t)\,\sigma_m^2(t)}{\sigma_n(t)\,\sigma_m(t)}\hat{\rho}_{nm} = \rho_{nm}. \tag{16}$$
- $\hat{\rho}_{nm}$ should be set *higher* than ρ_{nm} to achieve the same instantaneous Libor correlations, as a certain amount of decorrelation of forward Libor rates is already achieved by using two independent sets of Brownian motions
- With (15), (16), the SV and MSV models are identical in the zero-stochastic-volatility case $(\eta, \eta^1, \eta^2 = 0)$. Not so in general case

44 Drifts of Libor rates

- Define $n(t)$ by the condition
$$T_{n(t)-1} \leq t < T_{n(t)}.$$
- Under the spot Libor measure, the drifts $\mu_n^i(t)$, $n = 1, \ldots, N-1$, $i = 1, 2$, are given by
$$\mu_n^i(t) = \sqrt{z^i(t)} \sum_{j=n(t)}^{n} \frac{\tau_j \sigma_j^i(t)\,\hat{\rho}_{jn}\varphi(L_j(t))}{1 + \tau_j L_j(t)}.$$

45 Drifts of SV processes

- Under the T_{n+1} forward measure, the processes $z^i(t)$, $i = 1, 2$, follow the dynamics
$$dz^i(t) = \theta^i\left(1 - z^i(t)\right)\,dt - \varepsilon^i \nu^{i,n+1}(t, \mathbf{L}(t))\,z^i(t)\,dt + \varepsilon^i\sqrt{z^i(t)}\,dZ^{i,T_{n+1}}(t),$$
where
$$\nu^{i,n+1}(t, \mathbf{L}(t)) = \sum_{j=n(t)}^{n} \frac{\tau_j \sigma_j^i(t)\,\chi_n^i \varphi(L_j(t))}{1 + \tau_j L_j(t)}.$$
- Under swap measure $\mathsf{P}^{n,m}$ (for a swap rate that fixes at T_n and covers m periods), the processes $z^i(t)$, $i = 1, 2$, follow the dynamics
$$dz^i(t) = \theta^i\left(1 - z^i(t)\right)\,dt - \varepsilon^i \nu^{i,n,m}(t, \mathbf{L}(t))\,z^i(t)\,dt + \varepsilon^i\sqrt{z^i(t)}\,dZ^{i,n,m}(t),$$
where
$$\nu^{i,n,m}(t, \mathbf{L}(t)) = \sum_{k=n}^{n+m} \frac{\tau_k P(t, T_{k+1})}{A_{n,m}(t)}\nu^{i,k+1}(t, \mathbf{L}(t)), \tag{17}$$

46 Swap rate dynamics

- $S_{n,m}(\cdot)$ follows (under swap measure $\mathsf{P}^{n,m}$)
$$\begin{aligned} dS_{n,m}(t)/\varphi(S_{n,m}(t)) &= \sqrt{z^1(t)}\sigma_{n,m}^1(t)\,dW_{n,m}^{1,n,m}(t) \\ &\quad + \sqrt{z^2(t)}\sigma_{n,m}^2(t)\,dW_{n.m}^{2,n,m}(t), \end{aligned} \tag{18}$$
- $\sigma_{n,m}^i(t)$, $i = 1, 2$, are defined by
$$\left(\sigma_{n,m}^i(t)\right)^2 = \sum_{k,k'=n}^{n+m} w_{n,m}^k(t)\,w_{n,m}^{k'}(t)\,\sigma_k^i(t)\,\sigma_{k'}^i(t)\,\hat{\rho}_{kk'}, \tag{19}$$
- $dW_{n,m}^{i,n,m}$, $i = 1, 2$, are defined by
$$dW_{n,m}^{i,n,m}(t) = \frac{1}{\sigma_{n,m}^i(t)} \sum_{k=n}^{n+m} w_{n,m}^k(t)\,\sigma_k^i(t)\,dW_k^{i,n,m}(t),$$

47 Swap rate dynamics

- Exact result: with stochastic weights $w^k_{n,m}(t)$

$$w^k_{n,m}(t) = \frac{\varphi(L_k(t))}{\varphi(S_{n,m}(t))}\frac{\partial S_{n,m}(t)}{\partial L_k(t)}. \tag{20}$$

- For swaption pricing – the usual trick is to compute weights along the forwards,

$$w^k_{n,m} = \frac{\varphi(L_k(0))}{\varphi(S_{n,m}(0))}\frac{\partial S_{n,m}(0)}{\partial L_k(0)}. \tag{21}$$

- Same type of SDE as for Libor rates

48 CMS spread in MSV-FLM

- Let us demonstrate that the CMS spread dynamics in MSV-FLM have the desired features
- Without loss of generality look at two forward Libor rates $L_n(t)$ and $L_m(t)$, $n \neq m$.
- For simplicity assume $\varphi(x) \equiv 1$ and $\sigma^i_k(t) \equiv \sigma^i_k$, $i = 1, 2$, $k = n, m$.
- Ignoring drifts,

$$\begin{aligned} dL_k(t) &= \ldots dt + \sqrt{z^1(t)}\sigma^1_k\, dW^1_k(t) + \sqrt{z^2(t)}\sigma^2_k\, dW^2_k(t), \quad k = n, m, \\ dz^i(t) &= \ldots dt + \varepsilon^i\sqrt{z^i(t)}\, dZ^i(t), \quad i = 1, 2. \end{aligned}$$

49 CMS spread in MSV-FLM

- Rewrite the dynamics as

$$\begin{aligned} dL_k(t) &= \ldots dt + \sqrt{u^k(t)}\, dU_k(t), \\ du^k(t) &= \ldots dt + \sqrt{\eta^k(t)}\, dX^k(t), \end{aligned}$$

- Here

$$\begin{aligned} u^k(t) &= z^1(t)\left(\sigma^1_k\right)^2 + z^2(t)\left(\sigma^2_k\right)^2, \\ \eta^k(t) &= \left(\left(\sigma^1_k\right)^4\left(\varepsilon^1\right)^2 z^1(t) + \left(\sigma^2_k\right)^4\left(\varepsilon^2\right)^2 z^2(t)\right), \end{aligned}$$

- Also

$$\begin{aligned} dU_k(t) &= \frac{1}{\sqrt{u^k(t)}}\left(\sqrt{z^1(t)}\sigma^1_k\, dW^1_k(t) + \sqrt{z^2(t)}\sigma^2_k\, dW^2_k(t)\right), \\ dX^k(t) &= \frac{1}{\sqrt{\eta^k(t)}}\left(\left(\sigma^1_k\right)^2\varepsilon^1\sqrt{z^1(t)}\, dZ^1(t) + \left(\sigma^2_k\right)^2\varepsilon^2\sqrt{z^2(t)}\, dZ^2(t)\right). \end{aligned}$$

50 CMS spread in MSV-FLM

- Interested in various correlations between u's and L's. "At the forward": $z^1(t) = z^2(t) = 1$. (up to a scaling)

$$\begin{aligned} d\bar{U}_k(t) &= \sigma^1_k\, dW^1_k(t) + \sigma^2_k\, dW^2_k(t), \\ d\bar{X}^k(t) &= \left(\sigma^1_k\right)^2\varepsilon^1\, dZ^1(t) + \left(\sigma^2_k\right)^2\varepsilon^2\, dZ^2(t). \end{aligned}$$

- Correlations

$$\begin{aligned} du^n(t)\cdot du^m(t) &= \cos\left[\left(\left(\sigma^1_n\right)^2\varepsilon^1, \left(\sigma^2_n\right)^2\varepsilon^2\right) \hat{}\ \left(\left(\sigma^1_m\right)^2\varepsilon^1, \left(\sigma^2_m\right)^2\varepsilon^2\right)\right], \\ dL_n(t)\cdot dL_m(t) &= \hat{\rho}_{nm}\cos\left[\left(\sigma^1_n, \sigma^2_n\right) \hat{}\ \left(\sigma^1_m, \sigma^2_m\right)\right], \\ dL_n(t)\cdot du^m(t) &= \sigma^1_n\left(\sigma^1_m\right)^2\varepsilon^1\chi^1_n + \sigma^2_n\left(\sigma^2_m\right)^2\varepsilon^2\chi^2_m. \end{aligned}$$

- Correlation between SVs for two Libor rates is determined by the relative sizes of $\left(\sigma_n^1/\sigma_n^2\right)^2\left(\varepsilon^1/\varepsilon^2\right)$ and $\left(\sigma_m^1/\sigma_m^2\right)^2\left(\varepsilon^1/\varepsilon^2\right)$. These ratios could be used to calibrate to stochastic variance correlations
- Once the ratios are set, the Libor (or swap) rate correlations $dL_n(t)\cdot dL_m(t)$ can be matched by choosing $\hat{\rho}_{nm}$ appropriately
- "Cross" correlations $dL_n(t)\cdot du^m(t)$ are controlled by χ_k^i, $i=1,2$, $k=n,m$, and by the ratio $\varepsilon^1/\varepsilon^2$
- To control correlations, 6 parameters altogether, just like in the simple model before

51 Fast European swaption pricing, freeze the drift

- Need a fast method for pricing European swaptions for calibration
- Fix a particular swap rate $S(t)=S_{n,m}(t)$, drop the subscripts n,m
- Recall under the swap measure $\mathsf{P}^{n,m}$,

$$dS(t)/\varphi(S(t)) = \sqrt{z^1(t)}\sigma^1(t)\;dW^1(t)+\sqrt{z^2(t)}\sigma^2(t)\;dW^2(t), \quad (22)$$
$$dz^i(t)=\theta^i\left(1-z^i(t)\right)\,dt-\varepsilon^i\nu^i(t,\mathbf{L}(t))\,z^i(t)\;dt+\varepsilon^i\sqrt{z^i(t)}\,dZ^i(t).$$

- Step 1. Freeze the SV drift:

$$dz^i(t)=\theta^i\left(1-z^i(t)\right)\,dt-\varepsilon^i\nu^i(t,\mathbf{L}(0))\,z^i(t)\;dt+\varepsilon^i\sqrt{z^i(t)}\,dZ^i(t). \quad (23)$$

- Note that (22), (23) define an *affine* system of SDEs (if $\varphi(\cdot)$ is linearized)
- Idea 1. Effective FFT methods for pricing options on $S(t)$ are available (an extension of [AA02]). Fast enough for calibration?
- Idea 2. Extend averaging techniques of [Pit04], [Pit05b] to handle two stochastic variance processes

52 Fast European swaption pricing, single SV

- Idea 3. Approximate the dynamics of $S(t)$ with a single stochastic variance process (as we did before for the CMS spread option)
- Rewrite

$$dS(t)/\varphi(S(t))=\sqrt{u(t)}\,dU(t), \quad (24)$$

where

$$u(t) = z^1(t)\left(\sigma^1(t)\right)^2+z^2(t)\left(\sigma^2(t)\right)^2, \quad (25)$$
$$dU(t) = \frac{1}{\sqrt{u(t)}}\left(\sqrt{z^1(t)}\sigma^1(t)\,dW^1(t)+\sqrt{z^2(t)}\sigma^2(t)\;dW^2(t)\right).$$

- For du,

$$du(t)=\left(a^1(t)+b^1(t)\,z^1(t)\right)\,dt+\left(a^2(t)+b^2(t)\,z^2(t)\right)\,dt+\sqrt{\eta(t)}\,dX(t), \quad (26)$$

where

$$\eta(t) = \left(\left(\sigma^1(t)\right)^4\left(\varepsilon^1\right)^2z^1(t)+\left(\sigma^2(t)\right)^4\left(\varepsilon^2\right)^2z^2(t)\right),$$
$$dX(t) = \frac{1}{\sqrt{\eta(t)}}\left(\left(\sigma^1(t)\right)^2\varepsilon^1\sqrt{z^1(t)}\,dZ^1(t)+\left(\sigma^2(t)\right)^2\varepsilon^2\sqrt{z^2(t)}\,dZ^2(t)\right).$$

53 Fast European swaption pricing, Markovian projection

- As before, for any $\xi(t)$, for European options in the model (24), (25) can use

$$dS(t)/\psi(t,S(t))=\sqrt{\xi(t)}\,dU(t) \quad (27)$$

- where new local volatility function $\psi(t,x)$ is given by

$$\psi^2(t,x)=\varphi^2(x)\frac{\mathsf{E}\left(u(t)|\,S(t)=x\right)}{\mathsf{E}\left(\xi(t)|\,S(t)=x\right)}. \quad (28)$$

- Should keep $\xi(t)$ as close as possible to $u(t)$ to improve the quality of approximations in computing $\frac{\mathsf{E}(u(t)|S(t)=x)}{\mathsf{E}(\xi(t)|S(t)=x)}$
- Use Gaussian approximations

54 Fast European swaption pricing, projection of the SV process

- What to use for $\xi(t)$? Use $\tilde{u}(t)$, the Markovian projection of $u(\cdot)$. Then

$$\mathsf{E}\tilde{u}(t) = \mathsf{E}u(t), \quad \mathsf{Var}(\tilde{u}(t)) = \mathsf{Var}(u(t)),$$

- The overall level and the curvature of the smile are preserved.
- Using Gaussian approximations we obtain a process for $\tilde{u}(t)$,

$$d\tilde{u}(t) = (\tilde{\mu}_1(t) + \tilde{\mu}_2(t)\tilde{u}(t))\,dt + \sqrt{\tilde{\eta}_1(t) + \tilde{\eta}_2(t)\tilde{u}(t)}\,dX(t). \quad (29)$$

- Convenient to work with
- Use parameter averaging techniques to relate to a constant-parameter model

References

[AA02] Leif B.G. Andersen and Jesper Andreasen. Volatile volatilities. *Risk*, 15(12):163–168, December 2002.

[ABOBF02] Marco Avellaneda, Dash Boyer-Olson, Jérôme Busca, and Peter Friz. Reconstructing volatility. *Risk*, 15(10), October 2002.

[ABR01] Leif B.G. Andersen and Rupert Brotherton-Ratcliffe. Extended libor market models with stochastic volatility. Working paper, 2001.

[Dup97] Bruno Dupire. A unified theory of volatility. Banque Paribas working paper, 1997.

[Gyö86] I. Gyöngy. Mimicking the one-dimensional distributions of processes having an Ito differential. *Prob. Th. Rel. Fields*, 71:501 – 516, 1986.

[Pit04] Vladimir V. Piterbarg. A stochastic volatility forward Libor model with a term structure of volatility smiles. SSRN Working paper, 2004.

[Pit05a] Vladimir V. Piterbarg. Stochastic volatility model with time-dependent skew. *Applied Mathematical Finance*, 12(2):147–185,

[Pit05b] Vladimir V. Piterbarg. Time to smile. *Risk Magazine*, 18(5):71–75, May 2005.