Deconstructing the Volatility Smile

W. A. McGHEE† R. TRABALZINI**††

†The Royal Bank of Scotland, 135 Bishopsgate, London, EC2M 3UR, UK
‡‡Department of Mathematics, Imperial College London, London, SW7 2AZ, UK

(±v1.0 released 24 July 2014)

This paper investigates the relationship between the implied volatility smile and the underlying joint density of two quantities characterizing the stochastic volatility process - namely the mean integrated variance, \( \frac{1}{T} \int_0^T \sigma_s^2 ds \), and the stochastic integral \( \int_0^T \sigma_s dW_s \). A simple form of this joint density is proposed which, when fit to the zero correlation smile and a single non-zero correlation smile, will then generate to good agreement the smile for an arbitrarily chosen correlation. Further, the method complements and extends the work of Carr and Lee (2008) and Friz and Gatheral (2005) to non-zero correlation. In doing so, it allows for the study of volatility derivatives in the quanto case which is particularly relevant in the foreign exchange markets.

Keywords: Volatility Smile, SABR model, Heston model, volatility swap, quanto.

Introduction

The starting point in investigating the features and behaviour of a volatility process underlying asset dynamics is the construction of the volatility smile. In recent years, some investigation has taken place to understand the relationship between the particular choice of volatility process and the resulting implied volatility smile. Of particular interest is control of the wings of the volatility smile. Many models used in the financial sector - and in particular the widely deployed SABR Hagan et al. (2002) model - tend to show wing behaviour which deviates from that observed in many markets.

In this paper, a different approach is explored. On a filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{Q})\) satisfying the usual conditions where \(\mathbb{Q}\) is the equivalent martingale measure, let \(\{W_0^t, t \geq 0\}\) and \(\{W_\sigma^t \geq 0\}\) be two independent Brownian Motions, \(\mathcal{F}_0^t, \mathcal{F}_\sigma^t\) to be the natural filtrations of \(W_0^t\) and \(W_\sigma^t\), and \(\mathcal{F}_t = \mathcal{F}_0^t \cup \mathcal{F}_\sigma^t\). Letting \(\rho := \sqrt{1 - \rho^2}\), define

\[
W_t^S = \rho W_t^0 + \sqrt{1 - \rho^2} W_t^\sigma
\]

so that \(\langle W^S, W^\sigma \rangle_t = \rho t\). A generic conditionally lognormal forward process and an arbitrary stochastic volatility process are identified by the following system of SDEs,

\[
dF_t = \sigma_t F_t dW_t^S
\]

\[
d\sigma_t = g(\sigma_t) dt + h(\sigma_t) dW_t^\sigma
\]

---

* Email: william.mcghee@rbs.com ** Email: romano.trabalzini10@imperial.ac.uk

1Amongst recent attempts, Balland (2006) and later Andreasen and Hugue (2012) introduce a CEV power on the stochastic volatility process and McGhee (2013) introduces a third noise factor driving the volatility of volatility.

2The exact meaning of this expression will become clear in section 1.
where \( g(\cdot) : \mathbb{R} \to \mathbb{R} \) and \( h(\cdot) : \mathbb{R} \to \mathbb{R} \) are \( C^2 \) functions. In general, no Lipschitz conditions will be assumed on the functions \( g(\cdot) \) and \( h(\cdot) \) but the following assumptions will be made in order for the stochastic integral to be well defined:

**Assumption 1 (INTEGRABILITY)** \( \mathbb{E} \left[ \int_0^T |\sigma_s|^2 ds \right] < \infty, \) a.s.

**Assumption 2 (NON DEGENERACY)** (0.3) is a non-degenerate diffusion

By (0.1), equation (0.2) can be rewritten as

\[
dF_t = \rho \sigma_t F_t dW_t^\sigma + \bar{\rho} \sigma_t F_t dW_t^0 \\
d\sigma_t = g(\sigma_t)dt + h(\sigma_t)dW_t^\sigma
\]

Denoting by \( T \) the time to maturity, \( K \) the Strike, \( F_0 \) the initial forward level\(^1\) and \( \sigma_0 \) the initial volatility, the insight of Hull and White (1987)\(^2\) yields that a vanilla call option can be expressed as

\[
V_c(T, K, F_0, \sigma_0) = \mathbb{E}(F_T - K)^+
\]

\[
= \mathbb{E}[\mathbb{E}(F_T - K)^+ | \mathcal{F}_T^F] \quad (0.5)
\]

and a vanilla put as

\[
V_p(T, K, F_0, \sigma_0) = \mathbb{E}(K - F_T)^+
\]

\[
= \mathbb{E}[\mathbb{E}(K - F_T)^+ | \mathcal{F}_T^F] \quad (0.6)
\]

In the setup of this paper vanilla calls and puts are priced by averaging Black-Scholes kernels computed for randomized integrated variance levels, i.e. they are expressed as an integral over the joint density

\[
\psi(\vartheta_t, I_t)
\]

of the mean integrated realized instantaneous variance (\( \vartheta_t \)) (hereafter: MIV) and the volatility path integral (\( I_t \)) where

\[
\vartheta_t := \frac{1}{T-t} \int_t^T \sigma_s^2 ds \quad (0.7)
\]

and

\[
I_t := \int_t^T \sigma_s dW_s^\sigma \quad (0.8)
\]

In essence, this paper investigates the construction of the joint density \( \psi(\vartheta_t, I_t) \) via a parsimonious parametrization and then provides several examples where the recovered object can be used in actual pricing.

---

\(^1\)Throughout the paper the simplifying assumption of frictionless markets is made

\(^2\)Subsequently extended by Stein and Stein (1991) and Willard (1997), see below
0.1. **Literature**

The existing literature has approached the problem in an incomplete way. The following discussion will deal separately with the two cases of zero and non-zero correlation between asset process and volatility process.

### 0.1.1. The Zero Correlation Case.

As already discussed, the intuition behind treating a European vanilla payoff as a linear combination of Black-Scholes prices conditional upon different integrated variance levels comes originally from [Hull and White (1987)](HullWhite1987). Building on those ideas, the zero correlation case has subsequently been explored in [Carr and Lee (2008)](CarrLee2008) where the authors are interested in the variance market and in the probability density of the realized quadratic variation which they explicitly recover from a strip of European calls. Indeed if, as in this paper, the stock price is assumed to follow a continuous-path diffusion, the payoff of a *variance swap* corresponds to the *integrated instantaneous realized variance*, Section 5 contains a more detailed explanation and in Appendix A a simple extension to this approach is explored.

In [Friz and Gatheral (2005)](FrizGatheral2005), the authors show that expressing a European vanilla as an average of Black-Scholes prices is best done via direct inversion of the integral kernel (0.5) or (0.6). Although this procedure results in an highly unstable inverse problem they point towards regularization techniques that improve the fit in cases where parameter regimes are not benign.

### 0.1.2. The Non Zero Correlation Case.

The case of nonzero correlation between spot process and asset process is more complicated and has received less attention in literature. Notable contributions in this area are in order of appearance: [Willard (1997)](Willard1997), [Carr and Lee (2008)](CarrLee2008) and [McGhee (2011)](McGhee2011). [Willard (1997)](Willard1997) contributes the crucial intuition that in a stochastic volatility model, where the asset and volatility processes are correlated, the terminal stock price is lognormal conditional not only on the mean integrated variance (0.7) but on the entire path of the volatility process (0.8). Although most of the contributions in [Carr and Lee (2008)](CarrLee2008) are primarily targeted to the case where correlation between asset and volatility process is zero, as described above, Laplace transforms techniques are employed, the paper extends the methodology to the case where correlation is slightly perturbed away from $\rho = 0$, showing that a mixing formula is robust to the introduction of correlation.

The paper [McGhee (2011)](McGhee2011) gives a comprehensive analytic and numerical treatment for the above mentioned mixing formula in the context of the SABR model.

### 0.2. **Structure of the paper**

The remainder of this paper is organized as follows. Section 1 explains the general setup and gives some analytical explanation about how the low delta wings contain information about the right hand tail of the MIV distribution and each smile wing embodies the same information. Moreover, it points out that low volatility scenarios are concentrated around the AMTF whereas high volatility scenarios open up in the wings.

Section 2 deals with the case of zero correlation between asset and variance process. First, an analytic form for the MIV density is showed using research done in the context of Asian Options. Both an approximation (via Gram-Charlier expansion of density) and a pseudo-analytic close form (via Laplace Transform) are showed to make sense in the specialized lognormal case (taken as

---

1 Details about this can be accomplished will be given in the following
2 The analytic results of Willard (1997) form the skeleton of the treatment of the inverse problem in section 3
3 At-the-money forward is an option whose strike is at the same level as the prevailing market price of the underlying forward contract.
starting point). After that, a three parameter family of proxy distributions for the MIV density in the zero correlation case is fitted to market smile.

Section 3 explores the non zero correlation case, giving (in the non parametric case) an analytic framework to recover the joint density of MIV $\vartheta_t$ and the Volatility Path Integral $I_t$ as defined in equations (0.7) and (0.8) respectively. In order to recover the joint density conditioning is used: in particular, the proxy distribution for $\psi(\vartheta_t)$ is now used as building block for the joint density. The section shows that starting from any (single tenor) smile and using that proxy distribution, a stable parametrization for the joint density can be obtained. The key insight here is the following: since the joint density $\psi(\vartheta_t, I_t)$ has no dependence on correlation, it should be expected that a (parametric) form of it is recovered which is the same irrespective of the initial smile used in the inverse problem.

Section 4 presents two sets of conceptually different results. On the one hand section 4.1 shows how the joint density can recover either a SABR or a Heston model parametrization. Firstly, turning off correlation the stochastic volatility model is fitted with three out of four parameters in the joint density: subsequently, turning on correlation again, the additional parameter is used to match the correlated stochastic volatility model. On the other hand 4.2 independently fits the two stochastic volatility models (SABR and Heston) and the joint density to the same market smile, drawing comparisons between the fits.

Section 5 gives two applications of the work performed so far, in particular, it deals with the pricing of foreign denominated volatility and variance swaps when the correlation between spot and volatility is nonzero.

Section 6 concludes, outlining areas of future research and broadening the context of the current paper.

1. The abstract setup

1.1. Analytic setup

Using the setup explained above, composition of forward process (0.4), ie
\[
dF_t = \rho \sigma_t F_t dW_t^\sigma + \bar{\rho} \sigma_t F_t dW_t^0
\]
with $f : \mathbb{R}_+ \to \mathbb{R}$ function $f(\cdot) := \log(\cdot) \in C^2$ yields by Itô
\[
dx_t := d\log F_t = -\frac{1}{2} \sigma_t^2 dt + \rho \sigma_t dW_t^\sigma + \bar{\rho} \sigma_t dW_t^0
\]
which upon conditioning on $\mathcal{F}^\sigma_T \lor \mathcal{F}^0_T$ and rewriting in integral form yields
\[
X_T = X_t + A(\rho, \overline{\vartheta}_t, I_t) + B(\overline{\rho}, t)
\tag{1.1}
\]
where for $\overline{\vartheta}_t := \vartheta_t(T - t)$
\[
A(\rho, \overline{\vartheta}_t, I_t) := \rho I_t - \frac{1}{2} \overline{\vartheta}_t, \quad A(\rho, \overline{\vartheta}_t, I_t) \in m\mathcal{F}^\sigma_T
\]
and
\[
B(\overline{\rho}, t) := \overline{\rho} \int_t^T \sigma_s dW_s^0
\]
Conditioning on a path of \( \{W^\sigma_t, t \in [0, T]\} \) yields
\[
m_1 := E[X_T | \mathcal{F}_T^\sigma \vee \mathcal{F}_t^0] = X_t + \left( \rho I_t - \frac{1}{2} \langle \rho I_t \rangle_t \right) - \frac{\rho^2}{2} \bar{\vartheta}_t
\]
(1.2)

By Itô’s isometry and (1.2)
\[
m_2 := E[(X_T - m_1)^2 | \mathcal{F}_T^\sigma \vee \mathcal{F}_t^0] = \rho^2 \bar{\vartheta}_t
\]
(1.3)

Combining (1.2) and (1.3) yields:
\[
X_T \overset{d}{=} N \left( X_t + \rho I_t - \frac{1}{2} \left( \rho I_t \right)_t - \frac{\rho^2}{2} \bar{\vartheta}_t, \rho^2 \bar{\vartheta}_t \right)
\]

ie \( F_T := \exp(X_T) \) is conditionally lognormal

### 1.2. The conditional forward

First, define
\[
V^{BS}_{cp}(T, K, F, a) := \begin{cases} 
FN(d_1) - KN(d_2), & \text{if call} \\
KN(-d_2) - FN(-d_1), & \text{if put} 
\end{cases}
\]

where
\[
d_1 = \frac{1}{a \sqrt{T}} \ln \left( \frac{F}{K} \right) + \frac{a \sqrt{T}}{2}
\]
\[
d_2 = d_1 - a \sqrt{T}
\]

Following Willard (1997) and denoting with \( T \) the time to maturity, \( K \) the strike, \( X_0 \) the initial log-forward level and \( \sigma_0 \) the initial volatility, a vanilla call option\(^1\) in log forward coordinates \( V(T, K, e^{X_0}, \sigma_0) \) can be expressed (following (0.5)) as
\[
V_c(T, K, e^{X_0}, \sigma_0) = E \left[ E \left[ (e^{X_T} + A(\rho, \bar{\vartheta}_t, I_t) + B(\pi, 0) - K)^+ \bigg| \mathcal{F}_T^\sigma \right] \right]
\]

Rewriting the above yields\(^2\) in the case \( \rho = 0 \),
\[
E[(e^{X_T} - K)^+] = \int_{\mathbb{R}} d\vartheta \psi(\vartheta) V^{BS}_{cp}(T, K, e^{X_0}, \bar{\vartheta})
\]
(1.4)

Under assumption of integrability \( E[(e^{X_T} - K)^+] < \infty \), in the case \( \rho \neq 0 \), the integral equation

\(^1\)The actual calibration will use vanilla calls or puts depending on the region of the smile being fitted, but the put case being completely analogous it will not be analyzed further.

\(^2\)Dropping the time index, the two integration variables \( \vartheta_t \) and \( I_t \) will be expressed as \( \vartheta \) and \( I \).
can be rewritten (by applying Fubini’s theorem) as
\[
E[(e^{X_T} - K)^+] = \int_{\mathbb{R}_+ \times \mathbb{R}} d\vartheta dI \psi(\vartheta, I) V_{cp}^{BS}(T, K, e^{X_0}, \bar{\vartheta}_0, \rho)
\]
\[
= \int_{[K, \infty]} \left\{ \int_{\mathbb{R}} \left( e^{X_0 + \rho I_0 - \frac{1}{2} \langle \rho I_0 \rangle_t - \frac{\sigma_0^2}{2} T_0 + \int_0^T \sigma_0 dW_0^r - K \right) \psi(\vartheta, I) dI \right\} d\vartheta
\]

The question addressed in sections 2 and 3 will be how to capture the joint density \( \psi(\vartheta, I) \) of (0.7) and (0.8). To conclude this section, more insight is given about what information with respect to the distribution of MIV can be read from the wings of a given set of market prices, i.e. a smile.

1.3. An expression for the low delta wings

It is well understood that the SABR approximation of the underlying stochastic process (Hagan et al. [2002]) in the presence of a large skew, and especially in the long dated regime\(^2\), gives rise to regions of negative implied probability density. It will be shown that a non trivial relationship exists between low delta wings and the right hand tail of the MIV distribution, where each smile wing embodies the same information. Moreover, it is also pointed out that low volatility scenarios are concentrated around the AMTF whereas high volatility scenarios open up in the wings. The two cases of call and put prices are investigated separately with same notation as in section 1.2.

1.3.1. The call case.

**Proposition 1.1** In the zero correlation case, for sufficiently large \( K \), \( \forall \epsilon > 0, \exists \tilde{\vartheta}_0 : V_c^{BS}(T, K, e^{X_0}, \tilde{\vartheta}_0) < \epsilon \).

**Proof.** As \( V_c(T, K, 0, \sigma_0) = 0 \) and \( V_c(T, K, e^{X_0}, \sigma_0) \) is a continuous increasing function of \( \vartheta_0 \),
\[
V_c(T, K, e^{X_0}, \sigma_0) = \int_{[0, \tilde{\vartheta}_0]} d\vartheta \psi(\vartheta) V_c^{BS}(T, K, e^{X_0}, \vartheta_0) + \int_{[\tilde{\vartheta}_0, \infty]} d\vartheta \psi(\vartheta) V_c^{BS}(T, K, e^{X_0}, \vartheta_0)
\]
\[
= \int_{[\tilde{\vartheta}_0, \infty]} d\vartheta \psi(\vartheta) V_c^{BS}(T, K, e^{X_0}, \vartheta_0) + \zeta
\]
where \( \forall \epsilon > 0, \zeta < \epsilon \) as
\[
\zeta := \int_{[0, \tilde{\vartheta}_0]} d\vartheta \psi(\vartheta) V_c^{BS}(T, K, e^{X_0}, \vartheta_0) < \epsilon \int_{[0, \tilde{\vartheta}_0]} d\vartheta \psi(\vartheta) < \epsilon
\]

1.3.2. The put case.

---

1. Low delta refers to regions of the smile where strike is very high (for calls) or very small (for puts)
2. Recently also in low rates environment
Proposition 1.2 In the zero correlation case, for sufficiently small $K$, \[ \forall \epsilon > 0, \exists \tilde{\vartheta}_0 : V_{p}^{BS}(T, K, e^{X_0}, \tilde{\vartheta}_0) < \epsilon. \]

The proof is similar to the case above and will not be repeated.

Remark 1.3 Low delta wings and skewed distributions

Propositions (1.1) and (1.2) above give an analytic expression to the insight that low delta wings - i.e. OTM calls and puts - are functions of the right hand tail of MIV distribution. This is the reason why upon fitting the distribution of MIV a functional form will be chosen that is rich enough to accommodate the skew itself.

2. Zero correlation Case

In this section, firstly some (pseudo-)analytic attempts to specify a functional form for the two quantities of interest $\vartheta_t$ and $I_t$ are reviewed. Secondly, since the mean of the MIV determines the level of the smile, its standard deviation the curvature and its skew the steepness of the wings, a three parameter family of distributions is proposed that can capture the salient features of MIV during calibration.

2.1. Density of MIV in the lognormal volatility case

Firstly, (Yor 2001, Ch. 2) recovers formulae for the density of the integral of exponential Brownian motion. Their use in the context of the SABR model is demonstrated in [slah] (2009). Secondly, if a lognormal volatility process is assumed as a specialized case, then an explicit form for $\psi(\vartheta_t)$ can be displayed. In particular, an approximation will be given which is based on Gram-Charlier expansion on a lognormal distribution.

2.2. Yor formula for exponential functionals of Brownian motions

In the context of Asian Options pricing, where the relevant entity is an integral of exponential functional of Brownian Motion, (Yor 2001, Ch. 2-3) explicitly computes the joint density

$$
\psi(A_T^{(\mu)}, B_T^{(\mu)})
$$

where

$$
B_T^{(\mu)} \sim N(\mu T, T) \quad A_T^{(\mu)} = \int_0^T ds \exp (2B_s^{(\mu)})
$$

as

$$
\psi(A_T^{(\mu)} \in du, B_T^{(\mu)} \in dx) = \exp \left( \mu x - \frac{\mu^2 T}{2} \right) \exp \left( - \frac{1 + e^{2x}}{2u} \right) \theta(e^x; T) \frac{dudx}{u}
$$

$$
\theta(r; T) = \frac{r}{\sqrt{2\pi^3 T}} \exp \left( \frac{\pi^2}{2T} \right) \int_0^\infty d\xi \exp \left( -\frac{\xi^2}{2T} \right) \exp (-r \cosh(\xi)) \sinh(\xi) \sin \left( \frac{\pi \xi}{T} \right)
$$

More appropriately, semi-explicitly, as the expression contains an integral term which needs to be evaluated numerically.
with \( r > 0, T > 0 \). At a first look, the approach taken by Yor seems useful if an explicit expression for the joint density \( \psi(\vartheta_t, \sigma_T) \) is sought. In that case, under time change \( \phi(t) := \xi^2 t \)

\[
\sigma(\phi(t)) = \sigma_0 \exp \left( -\frac{1}{2} \phi(t) + W(\phi(t)) \right) =: \sigma_0 \exp \left( B_{\phi(t)}^{(-1/2)} \right)
\]

and

\[
\vartheta = \frac{\sigma_0^2}{\xi^2 T} \int_0^{\xi^2 T} d\phi(t) \exp \left( 2B_{\phi(t)}^{(-1/2)} \right) =: \frac{\sigma_0^2}{\xi^2 T} A_{\nu^2 T}^{(-1/2)}
\]

from which

\[
\psi(\vartheta \in da, \sigma_T \in db) := \psi \left( A_{\nu^2 T}^{(-1/2)} = da \frac{\xi^2 T}{\sigma_0^2}, B_{\phi(t)}^{(-1/2)} = \log \left( \frac{db}{\sigma_0} \right) \right)
\]

At a second look, from (2.1), following (Matsumoto and Yor 2005, Thm. 4.4) some algebraic manipulations lead to a semi explicit expression for the density of \( \psi(\vartheta_t) \), ie

\[ P(A_t \in du) = \frac{du}{\sqrt{2\pi u}^3} \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} d\eta \cosh(\eta) \exp \left( \frac{\cosh(\eta)^2}{2u} - \frac{(\eta + i\pi/2)^2}{2t} \right) \]  

(2.2)

**Remark 2.1 (Explicit density for MIV under a modified SABR process)** As an application of the above reasoning, Islah (2009) posits a modified SABR model whose system of SDEs is

\[
\begin{align*}
    dS_t &= \xi_s t S_t^\alpha dW_t^1 \\
    d\sigma_t &= \xi_s \sigma_t dW_t^2 + \frac{1}{2} \xi_s^2 \sigma_t dt, \quad \xi \in \mathbb{R}
\end{align*}
\]

\( \langle W^1, W^2 \rangle_t = \rho t. \) For it, using again (Yor 2001, Ch. 2-3), the density of the MIV

\[
\vartheta_0 T = \int_0^T dt \int_0^{\xi^2 T} du \exp \left( 2B_u \right) =: \frac{\sigma_0^2}{\xi^2 T} A_{\xi^2 T}
\]

is explicitly given by using (2.2).

### 2.3. Expansion over lognormal density

In McGhee (2008), and taking inspiration from the earlier Jarrow and Rudd (1982), a Gram Charlier expansion of the density is performed, where

\[
\psi(\vartheta) \sim \psi_{GC} := \psi_{LN}(\vartheta) \left[ 1 + \frac{k_3}{3! \beta^{3/2}} H_3 \left( \frac{\log \vartheta - \alpha}{\sqrt{\beta}} \right) + \frac{k_4}{4! \beta^2} H_4 \left( \frac{\log \vartheta - \alpha}{\sqrt{\beta}} \right) + \ldots \right]
\]

where \( H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} = \left( x - \frac{d}{dx} \right)^n 1 \) is the n-th Hermite polynomial, \( k_n \) is the n-th cumulant of the probability distribution and \( \psi_{LN}(\cdot) \) is the log-normal density.

**Remark 2.2** It is important to remark that additional moments in the above expansion can be added to achieve a finer control on the curvature/skewness

---

1Because of the additional drift
Remark 2.3 In [McGhee (2011)], for the case of a lognormal SABR model, the distribution of the mean integrated variance is expressed as an integral of lognormals over the distribution of the terminal instantaneous volatility, $\sigma(T)$. Each lognormal has mean and variance to match the first two moments of the mean integrated variance conditional on $\sigma(T)$. The resulting density is shown to match the first four moments well for a large range of expiries.

2.4. Calibration of MIV density to market smiles: the skewed (log) normal fit

Remark (1.3) pointed towards a connection between low delta wings and the skew of MIV distribution (particularly the positive or right hand tail skew). Put differently, OTM options depend on the right wing of integrated variance density. Taking this into consideration, the hypothesis is made that a three parameter family of continuous distributions can capture most of the relevant features of MIV distribution. A candidate for such a family of continuous distributions is the class of Skew normal distributions\footnote{See Azzalini (1985)} with density

\[ P(X \in dx) := \frac{2}{\omega} \phi \left( \frac{x - \mu}{\omega} \right) \Phi \left( \alpha \left( \frac{x - \mu}{\omega} \right) \right) dx, \quad x \in (-\infty, \infty) \tag{2.3} \]

where $\phi$ (resp. $\Phi$) are pdf (resp. cdf) of Standard Normal and $\mu \in \mathbb{R}$ a location parameter, $\omega \in \mathbb{R}^+$ a scale parameter and $\alpha \in \mathbb{R}$ a shape parameter.

An explanation of the heuristic behind this is now presented. It would be meaningful for the above distribution to be fitted only to symmetric market smiles, i.e. smiles which when expressed in the mapping given by a stochastic volatility model would result in zero correlation parameter. Those smiles are rarely if ever observed in the market and hence it would seem that no meaningful calibration can be performed when the smile is not symmetric. In reality, the abstraction represented by the proposed MIV distributional form does not lose its importance since it is only a tool to accommodate a marginal density when a joint is indeed required. The following section will demonstrate the importance of the above abstraction.

3. Non-zero correlation case

This section contains the main result of the paper, i.e. an analytical and numerical framework to recover the joint density of the asset process and the variance process in a stochastic volatility setup where asset process is lognormal and variance process has an arbitrary structure.

3.1. Building a proxy distribution

The joint density $\psi(\varphi_t, I_t)$ is now required. By the considerations above regarding the dependence of call (resp. put) prices in regions of smile with very high (resp. very low) strike from the right hand tail of the MIV distribution, a proxied functional form for MIV has been chosen that can accommodate the presence of such a tail. This form can now be used as a starting point in order to construct the density of $I_t$ conditional on MIV levels.

3.1.1. Simplest case: first moment and polynomial expansion. Using the fit obtained in 2.4 and by Bayes’ Theorem, the joint density $\psi(\varphi_t, I_t)$ can be expressed as

\[ \psi(\varphi_t, I_t) \overset{Bayes}{=} \psi(I_t | \varphi_t) \psi(\varphi_t). \]
To be more precise, and in order to attempt a generalization, a few cases can be explored where the terminal distribution of \( \psi(I_t|\vartheta_t) \) is given an analytical form.

Consider first specialization of (3.3) given by

\[
d\sigma_t = \xi \sigma_t dW^\sigma_t, \quad \xi \in \mathbb{R}, \ \sigma(0) = \sigma_0
\]  

Integration of (3.1) gives \( I_0 = \sigma_T - \sigma_0 \) which, given that \( \sigma_T \) is lognormal, is a shifted lognormal.

Consider another specialization of (0.3) given by

\[
dv_t = k(v_{\infty} - v_t)dt + \xi \sqrt{v_t} dW^v_t, \quad k, v_{\infty}, \xi \in \mathbb{R}, \ v(0) = v_0
\]  

Integration of (3.2) gives \( I_0 = v_T - v_0 - kT(v_{\infty} - \vartheta_0) \) whose density is known analytically.

In the following section, instead of the conditional density \( \psi(I_t|\vartheta_t) \) its expected value will be considered. Using Assumption (1) and Itô isometry the following holds (see appendix B for a discussion of this):

\[
E \left[ \left( \int_0^T \sigma_s dW^\sigma_s \right)^2 \right] = E \left[ \int_0^T \sigma_s^2 ds \right]
\]

As a way to build a parsimonious model, consider an expansion of \( I_t \) in a polynomial in \( \vartheta_1^{1/2} \)

\[
E[I_t|\vartheta_t] = \sum_{k=0}^{n} \lambda_k \vartheta_1^{k/2}
\]  

(3.3)

Taking into consideration the Martingale Property of Itô integrals and the isometry (B.1), ie

\[
E[I_t] = 0 \quad \quad E[I_t^2] = (T-t) E[\vartheta_t]
\]

a second order expansion of \( I_t \) in a polynomial in \( \vartheta_1^{1/2} \) can be performed. Thus, having

\[
E[I_t|\vartheta_t] = \lambda_0 + \lambda_1 \vartheta_1^{1/2} + \lambda_2 \vartheta_t + o(\vartheta_t),
\]  

(3.4)

upon solving quadratic equation in \( \lambda_1 \)

\[
\lambda_0 = -\lambda_1 E[\vartheta_1^{1/2}] - \lambda_2 E[\vartheta_t]
\]  

(3.5)

and

\[
\lambda_1 = \frac{-\lambda_2(E[\vartheta_1^{1/2}] E[\vartheta_t] - E[\vartheta_1^{3/2}]) \pm \sqrt{\Delta}}{E[\vartheta_1^{1/2}]^2 - E[\vartheta_t]}
\]  

(3.6)

with

\[
\Delta := \lambda_2^2 (E[\vartheta_1^{1/2}] E[\vartheta_t] - E[\vartheta_1^{3/2}])^2 + (E[\vartheta_t] - E[\vartheta_1^{1/2}])^2 (\lambda_2^2 E[\vartheta_t] + (T-t) E[\vartheta_t])
\]

**Remark 3.1** Coefficients \( \lambda_0, \lambda_1, \lambda_2 \) in Eq. (3.4) must be interpreted as time-dependent even if, for ease of notation, this dependence is not explicitly stressed.

**Remark 3.2** A second look at equation (1.1) shows that \( X_T \) is expressed as linear combination of
components each of which does not depend on the parameter \( \rho \). In other words, the joint density \( \psi(\vartheta_t, I_t) \) which is sought is an absolute object with respect to correlation, and therefore an arbitrary smile is expected to convey enough information to recover it.

4. Numerical results of calibration

This section will give two sets of numerical results coming from the implementation of the analytical model outlined in section 2 and 3. In particular, in section 4.1 the joint density will be fitted to a SABR and a Heston model. This fit of the parametric form to a particular stochastic volatility model will demonstrate that the proposed model structure is flexible enough to recover what other well established models can accomplish.

In section 4.2 real market data will be fitted respectively with SABR model, with Heston model and finally with the joint density model. The resulting calibration will be shown to give a good fit to the market observables, at least on a par with standardly used models. The advantage of using the recovered joint density will be fully realized in the examples of section 5.

Remark 4.1 In Appendix (E) a different way to validate the consistency of the proxy distribution and its parameterization is sketched. It consists in evolving the Kolmogorov Forward Equation for the recovered transition density in the case of lognormal SABR and in the case of Heston process. The analytic structure is investigated, as an alternative way of proceeding.

4.1. Fitting joint density to stochastic volatility models

4.1.1. Log normal SABR. Consider the following lognormal SABR model:

\[
\begin{align*}
    dF_t &= \sigma_t F_t^\beta dW_{F}^t \\
    d\sigma_t &= \xi \sigma_t dW_{\sigma}^t, \quad \xi \in \mathbb{R}, \quad \sigma(0) = \sigma_0
\end{align*}
\]  

(4.1)

for some \( \mathcal{F}_t \)-adapted Brownian Motions \( W_{F}^t \) and \( W_{\sigma}^t \) where \( \langle W_{F}^t, W_{\sigma}^t \rangle_t = \rho t, \ |\rho| \leq 1 \). Suppose now a market smile were given which in terms of the model (4.1) had the following representation:

<table>
<thead>
<tr>
<th>( T )</th>
<th>( F(0) )</th>
<th>( r_d )</th>
<th>( r_f )</th>
<th>( \sigma_0 )</th>
<th>( \xi )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>2.500</td>
<td>0.00%</td>
<td>0.00%</td>
<td>10.00%</td>
<td>50.00%</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 4.1. Smile Characterization: Lognormal SABR Model

The reason why the calibrated parameter \( \rho \) in the above fit is not displayed needs some clarification. The market observables, when fitted, do give a unique value for the parameter \( \rho \). Here that calibrated parameter is subsequently changed to arbitrarily produce different smiles, as first column of tables (C1) and (C2) shows. Choosing any of those arbitrarily produced smiles the joint density is subsequently fitted.

Table (4.3) shows the parameterization that has been obtained for the proxy distribution where the parameters \( \mu, \omega \) and \( \alpha \) come from a calibration to a smile given by the parameters in table (4.2):

<table>
<thead>
<tr>
<th>( T )</th>
<th>( F(0) )</th>
<th>( r_d )</th>
<th>( r_f )</th>
<th>( \sigma_0 )</th>
<th>( \xi )</th>
<th>( \beta )</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>2.500</td>
<td>0.00%</td>
<td>0.00%</td>
<td>10.00%</td>
<td>50.00%</td>
<td>1.0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.2. Zero-Correlation Smile Characterization: Lognormal SABR Model

\(^{1}\) See Hagan et al. (2002)
where fitted zero-correlation distributional form is given by (2.3) while parameter $\lambda_2$ (and by (3.5) and (3.6)) also $\lambda_0$ and $\lambda_1$ are given by a fit to the parameters in table (4.1) with an arbitrary $\rho \neq 0$ parameter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-2.5917</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.3931</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>1.6171</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Table 4.3. SABR fitted parameters

In Appendix (C) the figures C1, C2 and C3 display graphically the agreement of single smile calibration to all the other smiles as expected, while tables (C1) and (C2) show implied volatility and premium differences across all the fitted smiles. It appears possible to conclude that the regions where there is maximum discrepancy are also the ones carrying the least contribution to the total price.

4.1.2. **Heston square root process.** Consider the following Heston model:

$$
\begin{aligned}
\begin{cases}
    dS_t = (r_d - r_f)S_t dt + \sqrt{v_t}S_t dW^S_t \\
    dv_t = k(v_\infty - v_t)dt + \xi \sqrt{v_t}dW^v_t, \quad k, v_\infty \in \mathbb{R}, \quad v(0) = v_0
\end{cases}
\end{aligned}
$$

(4.2)

for some $\mathcal{F}_t$-adapted Brownian Motions $W^S_t$ and $W^v_t$ where $(W^S_t, W^v_t)_t = \rho t$, for some $|\rho| \leq 1$. Suppose now a market smile were given which in terms of the model (4.2) had the following representation:

<table>
<thead>
<tr>
<th>$T$</th>
<th>$S(0)$</th>
<th>$r_d$</th>
<th>$r_f$</th>
<th>$\sqrt{v_0}$</th>
<th>$\xi$</th>
<th>$\kappa$</th>
<th>$\sqrt{v_\infty}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>2.50</td>
<td>0.00%</td>
<td>0.00%</td>
<td>10.59%</td>
<td>9.89%</td>
<td>0.30</td>
<td>10.59%</td>
</tr>
</tbody>
</table>

Table 4.4. Smile Characterization: Heston Model

The procedure here is exactly the same as in section (4.1.1) and so it will not be commented any further. In Appendix (D) the figures D1, D2 and D3 display graphically the agreement of single smile calibration to all the other smiles as expected, while tables (D1) and (D2) show implied volatility and premium differences across all the fitted smiles. Again, it appears possible to conclude that the regions where there is maximum discrepancy are also the ones carrying less contribution to the total price.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>-2.0477</td>
</tr>
<tr>
<td>$\omega$</td>
<td>0.3627</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-1.9600</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 4.5. Heston fitted parameters

\(^{1}\text{See Heston (1993)}\)
4.2. **Fitting joint density to Market data**

This section documents the smile deconstruction with real market data, a USDJPY smile with 1 Year tenor.

<table>
<thead>
<tr>
<th>%Delta</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5C</td>
<td>8.7970%</td>
</tr>
<tr>
<td>10C</td>
<td>8.3675%</td>
</tr>
<tr>
<td>15C</td>
<td>8.0758%</td>
</tr>
<tr>
<td>20C</td>
<td>7.8797%</td>
</tr>
<tr>
<td>25C</td>
<td>7.7504%</td>
</tr>
<tr>
<td>30C</td>
<td>7.6697%</td>
</tr>
<tr>
<td>35C</td>
<td>7.6245%</td>
</tr>
<tr>
<td>40C</td>
<td>7.6067%</td>
</tr>
<tr>
<td>45C</td>
<td>7.6139%</td>
</tr>
<tr>
<td>ATM</td>
<td>7.6466%</td>
</tr>
<tr>
<td>45P</td>
<td>7.7067%</td>
</tr>
<tr>
<td>40P</td>
<td>7.7973%</td>
</tr>
<tr>
<td>35P</td>
<td>7.9236%</td>
</tr>
<tr>
<td>30P</td>
<td>8.0936%</td>
</tr>
<tr>
<td>25P</td>
<td>8.3179%</td>
</tr>
<tr>
<td>20P</td>
<td>8.6116%</td>
</tr>
<tr>
<td>15P</td>
<td>9.0037%</td>
</tr>
<tr>
<td>10P</td>
<td>9.5518%</td>
</tr>
<tr>
<td>5P</td>
<td>10.2956%</td>
</tr>
</tbody>
</table>

Table 4.6. USDJPY 1Y smile

Using standard FX market convention, the above smile is fitted with the two stochastic volatility models already mentioned, namely SABR and Heston, and also with the joint density model.

4.2.1. **Lognormal SABR.** Here are the calibration figures for a lognormal SABR model.

<table>
<thead>
<tr>
<th>$\sigma_0$</th>
<th>$\xi$</th>
<th>$\beta$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.38%</td>
<td>73.74%</td>
<td>1.0</td>
<td>-14.03%</td>
</tr>
</tbody>
</table>

Table 4.7. USDJPYTY Smile Calibration: Lognormal SABR Model

4.2.2. **Heston.** Here are the calibration figures for a Heston model.

<table>
<thead>
<tr>
<th>$\sqrt{v_0}$</th>
<th>$\xi$</th>
<th>$\kappa$</th>
<th>$\sqrt{v_\infty}$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.46%</td>
<td>12.95%</td>
<td>0.01%</td>
<td>8.89%</td>
<td>-15.12%</td>
</tr>
</tbody>
</table>

Table 4.8. USDJPYTY Smile Calibration: Heston Model

4.2.3. **Density calibration.** Here are the calibration figures for the joint density.

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\lambda_2$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.6449</td>
<td>0.4588</td>
<td>0.0257</td>
<td>13.00</td>
<td>-14%</td>
</tr>
</tbody>
</table>

Table 4.9. USDJPYTY Smile Calibration: joint density model

The figures below illustrate the fit to the market of each of the three models.
Figure 4.1. USDJPY Smile Calibration: SABR

Figure 4.2. USDJPY Smile Calibration: Heston
4.3. The problem of single observed market (or model) smile

The incremental calibration that has been performed in sections 2.4 and 3 is hardly an option when fitting market observables. Still the benefit of decoupling the two effects, the zero correlated (symmetric) smile on the one hand and the correlated smile on the other, contributes intuition to the solution of the problem. Another way to contemplate the same argument would be rephrasing it along the lines of Friz and Gatheral (2005) who argue that knowledge of the non-zero correlation smile is of no importance as in any stochastic volatility model the variance process is given by an autonomous equation with no dependence on spot process. The next section will explore this issue more in detail.

5. Applications: Quanto volatility derivatives in FX

Under the assumption that the stock price follows a continuous-path diffusion, a semi-static model-independent hedge is well known to perfectly replicate the realized variance

$$\int_0^T v(s, \omega)ds =: \langle x \rangle_T$$

$$= \vartheta_0 T$$

The paper Carr and Lee (2008) concentrates the investigation in the zero-correlation smile case where the construction of $\psi(\vartheta_t)$ allows for the calculation of theoretical (i.e. continuous rather than discretely monitored) volatility derivatives which are a function of the MIV. To be precise, for arbitrary $f(y)$ that can be represented or approximated by a linear combination of Laplace
functionals $y \mapsto \exp[\lambda y]$, $\lambda \in \mathbb{C}$, prices and hedges of contracts having payoff

$$\frac{1}{T} \mathbb{E}[f(\langle x \rangle_T)] = \mathbb{E}[f(\vartheta_0)] = \int_{\mathbb{R}_+} d\vartheta \psi(\vartheta)f(\vartheta)$$

are given explicitly in terms of associated European payoffs provided that correlation is zero. Friz and Gatheral (2005) argue that knowledge of the non-zero correlation smile is of no importance: in the context of stochastic volatility models the fair value of volatility derivatives does not depend on the correlation assumption. This remark can alternatively be phrased by saying that in a stochastic volatility model the variance process is given by an autonomous equation with no dependence on spot process. The hidden assumption here, common in equity markets, is that the contracts are denominated in the numeraire currency. In the FX market instead, a volatility derivative can pay in either the domestic or the foreign currency. Using same notation and assumptions as in section 1 and having $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_0^t]$, the time-$t$ domestic currency value of a foreign paying volatility derivative can be expressed as

$$\mathbb{E}_t[f(\vartheta_t)S_T] = \mathbb{E}_t[\mathbb{E}_t[f(\vartheta_t)S_T|\mathcal{F}_T^t]] = \mathbb{E}_t[f(\vartheta_t)\mathbb{E}[S_T|\mathcal{F}_T^t \vee \mathcal{F}_0^t]]$$

by Tower Property and for $f(\vartheta_t) \in m(\mathcal{F}_T^t \vee \mathcal{F}_0^t)$. For $L^2$-integrable $\sigma_t$ (see Assumption 1), Itô isometry gives Laplace Transform of Wiener integral $B_0 := \int_0^T \sigma_u dW_u^0$ and inner expectation in (5.1) can be written as

$$\mathbb{E}[S_T|\mathcal{F}_T^t \vee \mathcal{F}_0^t] := \mathbb{E}[e^{X_T}|\mathcal{F}_T^t \vee \mathcal{F}_0^t]$$

$$= e^{X_t - \frac{1}{2} \langle I_t \rangle_T + \frac{\rho^2}{2} \vartheta_t(T-t)} \mathbb{E}[e^{\rho I_t}|\mathcal{F}_T^t]$$

$$\approx e^{X_t - \frac{1}{2} \rho^2 \vartheta_t(T-t)+\rho(\lambda_0+\lambda_1\vartheta_{1/2}^1+\lambda_2\vartheta_{1})}$$

where (5.3) follows from (5.2) by measurability considerations and isometry above, while (5.4) follows from (5.3) by using the proxy previously given in (3.4).

To summarize, when $\rho = 0$ the contracts are simply related by the forward rate; however, when $\rho \neq 0$ a non-trivial relationship exists due to the relationship between the level of volatility and the FX rate.

Clearly the function $f(\cdot)$ can take several forms.1 Among the number of possible payoffs that can be priced with this method, in the next subsection actual results of this computations will be provided for Quanto Volatility and Variance Swaps in the two cases of SABR and Heston fit.

5.1. **Foreign denominated volatility and variance swaps under SABR model**

Starting from parametrization in table (4.1), the following is a comparison of foreign denominated volatility and variance swaps prices computed with different methods. In particular, results from three different procedures are highlighted for different correlation regimes. The

---

1An incomplete list would contain at least the following:

(i) Volatility Swap: $f : y \mapsto \sqrt{y}$; Variance Swap: $f : y \mapsto y$

(ii) Option on Volatility: $f : y \mapsto (\sqrt{y} - K)^+$; Option on Variance: $f : y \mapsto (y - K)^+$
**Additional Integration Method** refers to the procedure outlined in McGhee (2011), where instead of performing an expansion the SABR process is integrated directly. The **Joint Density Method** refers to the procedure which has been described in section (3). Finally the **MonteCarlo Method** is here performed with Log-Euler in forward coordinate and Euler in volatility coordinate, with (TimeSteps,Iterations)=(1e3,5e5).

<table>
<thead>
<tr>
<th>Corr</th>
<th>VolDom</th>
<th>VolFor</th>
<th>Diff</th>
<th>Joint Integration</th>
<th>Joint Density</th>
<th>MonteCarlo</th>
</tr>
</thead>
<tbody>
<tr>
<td>-80%</td>
<td>10.20%</td>
<td>9.94%</td>
<td>-0.26%</td>
<td>9.95%</td>
<td>9.98%</td>
<td>9.98%</td>
</tr>
<tr>
<td>-60%</td>
<td>10.20%</td>
<td>10.04%</td>
<td>-0.17%</td>
<td>10.01%</td>
<td>10.04%</td>
<td>10.04%</td>
</tr>
<tr>
<td>-40%</td>
<td>10.20%</td>
<td>10.09%</td>
<td>-0.11%</td>
<td>10.08%</td>
<td>10.09%</td>
<td>10.09%</td>
</tr>
<tr>
<td>-20%</td>
<td>10.20%</td>
<td>10.15%</td>
<td>-0.06%</td>
<td>10.14%</td>
<td>10.15%</td>
<td>10.15%</td>
</tr>
<tr>
<td>0%</td>
<td>10.20%</td>
<td>10.20%</td>
<td>0.00%</td>
<td>10.21%</td>
<td>10.20%</td>
<td>10.20%</td>
</tr>
<tr>
<td>20%</td>
<td>10.20%</td>
<td>10.26%</td>
<td>0.06%</td>
<td>10.28%</td>
<td>10.26%</td>
<td>10.26%</td>
</tr>
<tr>
<td>40%</td>
<td>10.20%</td>
<td>10.32%</td>
<td>0.12%</td>
<td>10.34%</td>
<td>10.32%</td>
<td>10.32%</td>
</tr>
<tr>
<td>60%</td>
<td>10.20%</td>
<td>10.38%</td>
<td>0.17%</td>
<td>10.41%</td>
<td>10.38%</td>
<td>10.38%</td>
</tr>
<tr>
<td>80%</td>
<td>10.20%</td>
<td>10.44%</td>
<td>0.24%</td>
<td>10.48%</td>
<td>10.44%</td>
<td>10.44%</td>
</tr>
</tbody>
</table>

**Table 5.2. Comparison of Variance Swaps Prices**

<table>
<thead>
<tr>
<th>Corr</th>
<th>VarDom</th>
<th>VarFor</th>
<th>Diff</th>
<th>Joint Integration</th>
<th>Joint Density</th>
<th>MonteCarlo</th>
</tr>
</thead>
<tbody>
<tr>
<td>-80%</td>
<td>10.66%</td>
<td>10.40%</td>
<td>-0.25%</td>
<td>10.39%</td>
<td>10.40%</td>
<td>10.40%</td>
</tr>
<tr>
<td>-60%</td>
<td>10.66%</td>
<td>10.47%</td>
<td>-0.19%</td>
<td>10.46%</td>
<td>10.46%</td>
<td>10.46%</td>
</tr>
<tr>
<td>-40%</td>
<td>10.66%</td>
<td>10.53%</td>
<td>-0.13%</td>
<td>10.54%</td>
<td>10.53%</td>
<td>10.53%</td>
</tr>
<tr>
<td>-20%</td>
<td>10.66%</td>
<td>10.59%</td>
<td>-0.07%</td>
<td>10.61%</td>
<td>10.59%</td>
<td>10.59%</td>
</tr>
<tr>
<td>0%</td>
<td>10.66%</td>
<td>10.66%</td>
<td>0.00%</td>
<td>10.69%</td>
<td>10.65%</td>
<td>10.65%</td>
</tr>
<tr>
<td>20%</td>
<td>10.66%</td>
<td>10.73%</td>
<td>0.07%</td>
<td>10.77%</td>
<td>10.72%</td>
<td>10.72%</td>
</tr>
<tr>
<td>40%</td>
<td>10.66%</td>
<td>10.79%</td>
<td>0.14%</td>
<td>10.85%</td>
<td>10.79%</td>
<td>10.79%</td>
</tr>
<tr>
<td>60%</td>
<td>10.66%</td>
<td>10.86%</td>
<td>0.21%</td>
<td>10.93%</td>
<td>10.86%</td>
<td>10.86%</td>
</tr>
<tr>
<td>80%</td>
<td>10.66%</td>
<td>10.94%</td>
<td>0.28%</td>
<td>11.01%</td>
<td>10.93%</td>
<td>10.93%</td>
</tr>
</tbody>
</table>

Tables 5.1 and 5.2 show in the first column the calibrated correlation parameter, in second column the fair prices of volatility (resp. variance) swap under the domestic measure, in third and fourth columns the fair prices of them in the foreign measure as obtained by the method of conditional integration, in fifth and sixth the fair prices of them in the foreign measure as obtained by integration of joint density, and in the last two the fair prices of them as obtained with a MonteCarlo simulation.

### 5.2. Foreign denominated volatility and variance swaps under Heston model

In case of a Heston process, following Gatheral (2006) a Milstein discretization scheme with absorbing boundary can substantially alleviate the problem of negative variances. Specifically, using the discretization below for the variance process

\[ v_{i+1} = v_i - k(v_i - v_\infty)\Delta t + \xi \sqrt{v_i} \sqrt{\Delta t} Z + \frac{\xi^2}{4} \sqrt{\Delta t} \left(Z^2 - 1\right) \]

\(^1\)See Kloeden and Platen (1999)
where $Z \sim \mathcal{N}(0, 1)$ the absorbed paths are zero for the parameter regime given by Table 4.4.

Tables 5.3 and 5.4 show the comparison figures: first column lists correlations, second column lists volatility (resp. variance) swap prices denominated in domestic measure, third and fourth list the prices obtained by sampling where Log-Euler has been used in spot coordinate and Milstein in variance coordinate, with $(\text{TimeSteps, Iterations}) = (1e3, 5e5)$ while last two show the foreign denominated prices as obtained by integration of joint density.

### Table 5.3. Comparison of Volatility Swaps Prices

<table>
<thead>
<tr>
<th>Corr</th>
<th>VolDom</th>
<th>VolFor</th>
<th>Diff</th>
<th>VolFor</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>-60%</td>
<td>10.29%</td>
<td>10.15%</td>
<td>-0.14%</td>
<td>10.13%</td>
<td>-0.16%</td>
</tr>
<tr>
<td>-40%</td>
<td>10.29%</td>
<td>10.20%</td>
<td>-0.09%</td>
<td>10.18%</td>
<td>-0.10%</td>
</tr>
<tr>
<td>-20%</td>
<td>10.29%</td>
<td>10.24%</td>
<td>-0.05%</td>
<td>10.24%</td>
<td>-0.05%</td>
</tr>
<tr>
<td>0%</td>
<td>10.29%</td>
<td>10.29%</td>
<td>0.00%</td>
<td>10.29%</td>
<td>0.00%</td>
</tr>
<tr>
<td>20%</td>
<td>10.29%</td>
<td>10.33%</td>
<td>0.05%</td>
<td>10.34%</td>
<td>0.05%</td>
</tr>
<tr>
<td>40%</td>
<td>10.29%</td>
<td>10.38%</td>
<td>0.09%</td>
<td>10.40%</td>
<td>0.11%</td>
</tr>
<tr>
<td>60%</td>
<td>10.29%</td>
<td>10.43%</td>
<td>0.14%</td>
<td>10.45%</td>
<td>0.16%</td>
</tr>
</tbody>
</table>

### Table 5.4. Comparison of Variance Swaps Prices

<table>
<thead>
<tr>
<th>Corr</th>
<th>VarDom</th>
<th>VarFor</th>
<th>Diff</th>
<th>VarFor</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>-60%</td>
<td>10.59%</td>
<td>10.45%</td>
<td>-0.14%</td>
<td>10.43%</td>
<td>-0.16%</td>
</tr>
<tr>
<td>-40%</td>
<td>10.59%</td>
<td>10.50%</td>
<td>-0.09%</td>
<td>10.49%</td>
<td>-0.11%</td>
</tr>
<tr>
<td>-20%</td>
<td>10.59%</td>
<td>10.54%</td>
<td>-0.05%</td>
<td>10.54%</td>
<td>-0.05%</td>
</tr>
<tr>
<td>0%</td>
<td>10.59%</td>
<td>10.59%</td>
<td>0.00%</td>
<td>10.59%</td>
<td>0.00%</td>
</tr>
<tr>
<td>20%</td>
<td>10.59%</td>
<td>10.64%</td>
<td>0.05%</td>
<td>10.65%</td>
<td>0.05%</td>
</tr>
<tr>
<td>40%</td>
<td>10.59%</td>
<td>10.69%</td>
<td>0.10%</td>
<td>10.70%</td>
<td>0.11%</td>
</tr>
<tr>
<td>60%</td>
<td>10.59%</td>
<td>10.74%</td>
<td>0.14%</td>
<td>10.76%</td>
<td>0.17%</td>
</tr>
</tbody>
</table>

---

Alternatives abound in literature: among the most prominent, the exact simulation of the system of SDE (see Broadie and Kaya (2006)). The scheme used to produce figures in Tables 5.3 and 5.4 is that suggested in Gatheral (2006) pg. 22ff.

---

2 Also by Feller Condition, see Clark (2011).
6. Conclusion

Starting from the seminal contribution of Stein and Stein (1991) where the concept of mixing distribution is defined and explored and using insights coming from Willard (1997) in this paper it is shown that from a single tenor market observed smile information can be recovered pertaining to two additional quantities which are relevant in a number of pricing exercises, namely the mean integrated variance, \( \frac{1}{T} \int_0^T \sigma_s^2 ds \), and the stochastic integral \( \int_0^T \sigma_s dW_s \).

A parsimonious fit is shown to be able to unlock most of the structure that links the second to the first. In recovering the joint density of the two mentioned quantities, some of the hidden assumptions that lie behind standardly used stochastic volatility models have become clearer. In doing so, the contributions of Carr and Lee (2008) and Friz and Gatheral (2005) have been extended to the more general case of non zero correlation, which becomes relevant in valuing Quanto products involving realized volatility (or variance).

It appears possible to perform the calibration of the joint density to a given market (or model) smile across a wide variety of stochastic volatility models. Besides the two already explored, ie the SABR and the Heston model, some others can be considered, in particular the ZABR model[1] and the exponential Ornstein-Uhlenbeck[2]. Those will be explored in future work. A subtler issue appears to be that of relaxing the lognormality assumption as regards the spot process: the case of CEV process, which is cognate to local volatility modeling, is also left for subsequent work.

Acknowledgements The authors[3] would like to thank Katia Babbar, Han Lee, Andrea Odetti, Stephen Smith and in particular Professor Mark H. A. Davis for reviewing and making valuable comments on this paper.

---

[3] The second author would like to acknowledge the financial support given by RBS plc
Appendix A: A note on a connection with power payoff: the Moment Problem

Under assumption of diffusion and zero correlation between spot moves and volatility moves, Carr and Lee (2008) give an expression for the Laplace transform of the quadratic variation \( \langle x \rangle_T \) in terms of the fair value of the power payoff (here \( \langle x \rangle_T \) corresponds to \( \vartheta_0 T \) in Eq. (0.7)). Under the assumption that \( E[\exp(\lambda \langle x \rangle_T)] < \infty \) and for some \( \lambda > 0 \), we write

\[
E[\exp(\lambda \langle x \rangle_T)] := E[\exp(p(\lambda)x_T)], \quad p(\lambda) := \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2\lambda} \tag{A1}
\]

Differentiation of LHS gives

\[
\frac{d^k}{d\lambda^k} (LHS) \mid_{\lambda=0} = E[\langle x \rangle_T^k]
\]

whilst same procedure on RHS

\[
\frac{d^k}{d\lambda^k} (RHS) \mid_{\lambda=0} = \left( \frac{d^k}{d\lambda^k} p(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^k] + \left( \frac{1}{2} \frac{d^k}{d\lambda^k} p^2(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^2] + \ldots + \left( \frac{1}{k!} \frac{d^k}{d\lambda^k} p^k(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^k] + \ldots \tag{A2}
\]

It is now interesting to notice that, by choosing, as we may, \( p(\lambda) := \frac{1}{2} - \sqrt{\frac{1}{4} + 2\lambda} \), we can simplify the previous expression and get an extremely concise form for moments of quadratic variation as linear combination of finitely many moments of power contract.

**Proposition A.1** Upon choosing \( p(\lambda) := \frac{1}{2} - \sqrt{\frac{1}{4} + 2\lambda} \), RHS of equation (A1) can be expressed via a finite collection of moments.

**Proof.** In formula (A2), \( \forall n,k: k < n \),

\[
\frac{d^k}{d\lambda^k} p^n(\lambda) \bigg|_{\lambda=0} = \alpha p^{n-k}(\lambda) \bigg|_{\lambda=0} = 0, \quad \alpha \in \mathbb{R}
\]

hence:

\[
\frac{d^k}{d\lambda^k} (RHS) \mid_{\lambda=0} = \left( \frac{d^k}{d\lambda^k} p(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^k] + \left( \frac{1}{2} \frac{d^k}{d\lambda^k} p^2(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^2] + \ldots + \left( \frac{1}{k!} \frac{d^k}{d\lambda^k} p^k(\lambda) \bigg|_{\lambda=0} \right) \cdot E[x_T^k] \]

\[\square\]
Appendix B: Itô Isometry

**Proposition B.1** Let:

\[
A := E \left[ \left( \int_0^T \sigma_s dW_s^\sigma \right)^2 \right] \quad B := E \left[ \int_0^T \sigma_s^2 ds \right] = \| \int_0^T \sigma_s^2 ds \|_{L^2}
\]

Then for driving SDE

\[
d\sigma_t = \xi \sigma_t dW_t^\sigma \quad \xi \in \mathbb{R} \tag{B1}
\]

and partition \(0 = t_0 < \ldots < t_k < T\), the following holds:

\[
\frac{A - B}{B} \to 0 \quad \text{as} \quad k \to \infty
\]

**Proof.** By (B1) and using Lebesgue Monotone Convergence Theorem,

\[
B = E \left[ \int_0^T ds \sigma_0^2 \exp \left( -\xi^2 s + 2\xi W_s^\sigma \right) \right] = \frac{\sigma_0^2}{\xi^2} (\exp (\xi^2 T) - 1)
\]

On the other hand, again using (B1),

\[
A' := \frac{A}{\sigma_0^2} = E \left[ \left( \int_0^T \exp \left( -\frac{\xi^2 s}{2} + \xi W_s^\sigma \right) dW_s^\sigma \right)^2 \right] \tag{B2}
\]

Define a partition \(0 = t_0 < \ldots < t_k < T\) and denote with \(A'_k\) the discretization of integral in RHS of Equation (B2).

\[
A'_k = E \left[ \sum_{i=0}^{k-1} \exp \left( -\frac{\xi^2 t_i}{2} + \xi W_{t_i}^\sigma \right) \left( W_{t_{i+1}}^\sigma - W_{t_i}^\sigma \right)^2 + \exp \left( -\frac{\xi^2 t_k}{2} + \xi W_{t_k}^\sigma \right) \left( W_{T}^\sigma - W_{t_k}^\sigma \right)^2 \right] \tag{B3}
\]

\[
= E \left[ \sum_{i=0}^{k-1} \exp \left( -\xi^2 t_i + 2\xi W_{t_i}^\sigma \right) \left( W_{t_{i+1}}^\sigma - W_{t_i}^\sigma \right)^2 + \exp \left( -\xi^2 t_k + 2\xi W_{t_k}^\sigma \right) \left( W_{T}^\sigma - W_{t_k}^\sigma \right)^2 \right] \tag{B4}
\]

\[
= \sum_{i=0}^{k-1} E \left[ \exp \left( 2\xi W_{t_i}^\sigma \right) \right] E \left[ \exp \left( -\xi^2 t_i \right) \left( W_{t_{i+1}}^\sigma - W_{t_i}^\sigma \right)^2 \right] + E \left[ \exp \left( 2\xi W_{t_k}^\sigma \right) \right] E \left[ \exp \left( -\xi^2 t_k \right) \left( W_{T}^\sigma - W_{t_k}^\sigma \right)^2 \right] \tag{B5}
\]

\[
= \sum_{i=0}^{k-1} \exp (2\xi^2 t_i) \exp (-\xi^2 t_i) \exp (-\xi^2 t_{i+1} - t_i) + \exp (2\xi^2 t_k) \exp (-\xi^2 t_k) \exp (-\xi^2 t_k) (T - t_k) \tag{B6}
\]

\[
k \equiv \int_0^T ds \exp (\xi^2 s)
\]

where (B4) from (B3) as disjoint partitions yield zero expectation for cross terms, (B5) from (B4) by Brownian independent increments, (B6) from (B5) by Itô isometry. \(\square\)
Appendix C: Numerical Results of Calibration: SABR Case

Figure C1. Smile Calibration: $\rho = 0$ case

Figure C2. Smile Calibration: $\rho = 50\%$ case
Figure C3. Smile Calibration: $\rho = -50\%$ case

**Remark C.1 Robustness of Fit**

Figures C2 and C3 reveal regions of the smile where there is some difference between the implied volatilities resulting on the one hand from SABR calibration and on the other hand from the joint density fit. Upon comparing the relevant entries in Tables C2 and C1, it is clear that the areas where the implied volatilities deviate are at the same time insignificant in terms of price, i.e. carry very little PV.
<table>
<thead>
<tr>
<th>Strike Space</th>
<th>2.00</th>
<th>2.10</th>
<th>2.20</th>
<th>2.31</th>
<th>2.41</th>
<th>2.51</th>
<th>2.61</th>
<th>2.71</th>
<th>2.82</th>
<th>2.92</th>
<th>3.02</th>
<th>3.10</th>
<th>3.17</th>
<th>3.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>0.06%</td>
<td>0.24%</td>
<td>0.84%</td>
<td>2.43%</td>
<td>5.65%</td>
<td>9.73%</td>
<td>6.29%</td>
<td>4.07%</td>
<td>2.66%</td>
<td>1.77%</td>
<td>1.20%</td>
<td>0.90%</td>
<td>0.69%</td>
<td>0.53%</td>
</tr>
<tr>
<td></td>
<td>0.05%</td>
<td>0.20%</td>
<td>0.78%</td>
<td>2.37%</td>
<td>5.63%</td>
<td>9.73%</td>
<td>6.30%</td>
<td>4.07%</td>
<td>2.66%</td>
<td>1.76%</td>
<td>1.18%</td>
<td>0.89%</td>
<td>0.67%</td>
<td>0.52%</td>
</tr>
<tr>
<td>25%</td>
<td>0.17%</td>
<td>0.46%</td>
<td>1.20%</td>
<td>2.83%</td>
<td>5.91%</td>
<td>9.73%</td>
<td>6.08%</td>
<td>3.75%</td>
<td>2.31%</td>
<td>1.45%</td>
<td>0.92%</td>
<td>0.67%</td>
<td>0.49%</td>
<td>0.36%</td>
</tr>
<tr>
<td></td>
<td>0.16%</td>
<td>0.44%</td>
<td>1.16%</td>
<td>2.80%</td>
<td>5.89%</td>
<td>9.74%</td>
<td>6.11%</td>
<td>3.78%</td>
<td>2.34%</td>
<td>1.47%</td>
<td>0.94%</td>
<td>0.68%</td>
<td>0.50%</td>
<td>0.37%</td>
</tr>
<tr>
<td>0%</td>
<td>0.31%</td>
<td>0.70%</td>
<td>1.53%</td>
<td>3.18%</td>
<td>6.12%</td>
<td>9.70%</td>
<td>5.82%</td>
<td>3.37%</td>
<td>1.93%</td>
<td>1.11%</td>
<td>0.65%</td>
<td>0.44%</td>
<td>0.30%</td>
<td>0.21%</td>
</tr>
<tr>
<td></td>
<td>0.31%</td>
<td>0.70%</td>
<td>1.53%</td>
<td>3.18%</td>
<td>6.12%</td>
<td>9.70%</td>
<td>5.82%</td>
<td>3.37%</td>
<td>1.93%</td>
<td>1.11%</td>
<td>0.65%</td>
<td>0.44%</td>
<td>0.30%</td>
<td>0.21%</td>
</tr>
<tr>
<td>-25%</td>
<td>0.46%</td>
<td>0.93%</td>
<td>1.83%</td>
<td>3.47%</td>
<td>6.28%</td>
<td>9.62%</td>
<td>5.51%</td>
<td>2.95%</td>
<td>1.51%</td>
<td>0.77%</td>
<td>0.39%</td>
<td>0.24%</td>
<td>0.15%</td>
<td>0.09%</td>
</tr>
<tr>
<td></td>
<td>0.46%</td>
<td>0.94%</td>
<td>1.85%</td>
<td>3.50%</td>
<td>6.29%</td>
<td>9.61%</td>
<td>5.47%</td>
<td>2.89%</td>
<td>1.46%</td>
<td>0.72%</td>
<td>0.36%</td>
<td>0.22%</td>
<td>0.13%</td>
<td>0.08%</td>
</tr>
<tr>
<td>-50%</td>
<td>0.62%</td>
<td>1.14%</td>
<td>2.09%</td>
<td>3.71%</td>
<td>6.39%</td>
<td>9.50%</td>
<td>5.15%</td>
<td>2.48%</td>
<td>1.07%</td>
<td>0.44%</td>
<td>0.17%</td>
<td>0.09%</td>
<td>0.05%</td>
<td>0.02%</td>
</tr>
<tr>
<td></td>
<td>0.61%</td>
<td>1.14%</td>
<td>2.09%</td>
<td>3.73%</td>
<td>6.41%</td>
<td>9.50%</td>
<td>5.11%</td>
<td>2.40%</td>
<td>0.99%</td>
<td>0.37%</td>
<td>0.13%</td>
<td>0.06%</td>
<td>0.03%</td>
<td>0.02%</td>
</tr>
</tbody>
</table>

† For each correlation, first row is SABR fitted smile, second is calibrated joint density.
Table C2. Implied Volatilities Differences (lognormal SABR)

<table>
<thead>
<tr>
<th>$\rho$ †</th>
<th>2.00</th>
<th>2.10</th>
<th>2.20</th>
<th>2.31</th>
<th>2.41</th>
<th>2.51</th>
<th>2.61</th>
<th>2.71</th>
<th>2.82</th>
<th>2.92</th>
<th>3.02</th>
<th>3.10</th>
<th>3.17</th>
<th>3.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>9.39%</td>
<td>9.19%</td>
<td>9.19%</td>
<td>9.39%</td>
<td>9.76%</td>
<td>10.25%</td>
<td>10.80%</td>
<td>11.37%</td>
<td>11.96%</td>
<td>12.54%</td>
<td>13.10%</td>
<td>13.51%</td>
<td>13.91%</td>
<td>14.30%</td>
</tr>
<tr>
<td></td>
<td>9.06%</td>
<td>8.91%</td>
<td>9.00%</td>
<td>9.29%</td>
<td>9.73%</td>
<td>10.24%</td>
<td>10.80%</td>
<td>11.38%</td>
<td>11.95%</td>
<td>12.51%</td>
<td>13.06%</td>
<td>13.47%</td>
<td>13.84%</td>
<td>14.22%</td>
</tr>
<tr>
<td>25%</td>
<td>10.76%</td>
<td>10.35%</td>
<td>10.08%</td>
<td>9.98%</td>
<td>10.04%</td>
<td>10.25%</td>
<td>10.57%</td>
<td>10.97%</td>
<td>11.41%</td>
<td>11.87%</td>
<td>12.35%</td>
<td>12.70%</td>
<td>13.05%</td>
<td>13.40%</td>
</tr>
<tr>
<td></td>
<td>10.65%</td>
<td>10.24%</td>
<td>10.00%</td>
<td>9.93%</td>
<td>10.02%</td>
<td>10.26%</td>
<td>10.60%</td>
<td>11.01%</td>
<td>11.45%</td>
<td>11.92%</td>
<td>12.39%</td>
<td>12.73%</td>
<td>13.09%</td>
<td>13.43%</td>
</tr>
<tr>
<td>0%</td>
<td>11.85%</td>
<td>11.29%</td>
<td>10.82%</td>
<td>10.47%</td>
<td>10.27%</td>
<td>10.21%</td>
<td>10.29%</td>
<td>10.48%</td>
<td>10.76%</td>
<td>11.09%</td>
<td>11.46%</td>
<td>11.75%</td>
<td>12.04%</td>
<td>12.34%</td>
</tr>
<tr>
<td></td>
<td>11.85%</td>
<td>11.28%</td>
<td>10.82%</td>
<td>10.48%</td>
<td>10.27%</td>
<td>10.21%</td>
<td>10.29%</td>
<td>10.49%</td>
<td>10.76%</td>
<td>11.09%</td>
<td>11.45%</td>
<td>11.74%</td>
<td>12.04%</td>
<td>12.33%</td>
</tr>
<tr>
<td>-25%</td>
<td>12.75%</td>
<td>12.05%</td>
<td>11.43%</td>
<td>10.88%</td>
<td>10.45%</td>
<td>10.13%</td>
<td>9.96%</td>
<td>9.92%</td>
<td>10.00%</td>
<td>10.17%</td>
<td>10.41%</td>
<td>10.61%</td>
<td>10.84%</td>
<td>11.07%</td>
</tr>
<tr>
<td></td>
<td>12.77%</td>
<td>12.09%</td>
<td>11.46%</td>
<td>10.91%</td>
<td>10.46%</td>
<td>10.12%</td>
<td>9.91%</td>
<td>9.84%</td>
<td>9.89%</td>
<td>10.05%</td>
<td>10.27%</td>
<td>10.48%</td>
<td>10.70%</td>
<td>10.93%</td>
</tr>
<tr>
<td>-50%</td>
<td>13.49%</td>
<td>12.68%</td>
<td>11.92%</td>
<td>11.21%</td>
<td>10.57%</td>
<td>10.02%</td>
<td>9.57%</td>
<td>9.26%</td>
<td>9.09%</td>
<td>9.05%</td>
<td>9.12%</td>
<td>9.23%</td>
<td>9.37%</td>
<td>9.53%</td>
</tr>
<tr>
<td></td>
<td>13.45%</td>
<td>12.67%</td>
<td>11.93%</td>
<td>11.23%</td>
<td>10.60%</td>
<td>10.01%</td>
<td>9.52%</td>
<td>9.15%</td>
<td>8.90%</td>
<td>8.78%</td>
<td>8.79%</td>
<td>8.85%</td>
<td>8.97%</td>
<td>9.12%</td>
</tr>
</tbody>
</table>

† For each correlation, first row is SABR fitted smile, second is calibrated joint density
Appendix D: Numerical Results of Calibration: Heston Case

Figure D1. Smile Calibration: $\rho = 0$ case

Figure D2. Smile Calibration: $\rho = 50\%$ case
Figure D3. Smile Calibration: $\rho = -50\%$ case

Remark D.1 Robustness of Procedure
Same consideration applies here as for Lognormal SABR. See Remark (C.1)
Table D1. Premium Differences (Heston)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>2.00</th>
<th>2.10</th>
<th>2.20</th>
<th>2.31</th>
<th>2.41</th>
<th>2.51</th>
<th>2.61</th>
<th>2.71</th>
<th>2.82</th>
<th>2.92</th>
<th>3.02</th>
<th>3.10</th>
<th>3.17</th>
<th>3.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>0.06%</td>
<td>0.26%</td>
<td>0.91%</td>
<td>2.57%</td>
<td>5.84%</td>
<td>9.88%</td>
<td>6.34%</td>
<td>4.01%</td>
<td>2.52%</td>
<td>1.57%</td>
<td>0.98%</td>
<td>0.69%</td>
<td>0.48%</td>
<td>0.34%</td>
</tr>
<tr>
<td></td>
<td>0.04%</td>
<td>0.22%</td>
<td>0.85%</td>
<td>2.53%</td>
<td>5.84%</td>
<td>9.92%</td>
<td>6.38%</td>
<td>4.04%</td>
<td>2.54%</td>
<td>1.59%</td>
<td>0.99%</td>
<td>0.70%</td>
<td>0.49%</td>
<td>0.35%</td>
</tr>
<tr>
<td>25%</td>
<td>0.14%</td>
<td>0.44%</td>
<td>1.21%</td>
<td>2.90%</td>
<td>6.00%</td>
<td>9.83%</td>
<td>6.11%</td>
<td>3.70%</td>
<td>2.21%</td>
<td>1.30%</td>
<td>0.77%</td>
<td>0.51%</td>
<td>0.35%</td>
<td>0.23%</td>
</tr>
<tr>
<td></td>
<td>0.13%</td>
<td>0.42%</td>
<td>1.17%</td>
<td>2.86%</td>
<td>5.99%</td>
<td>9.85%</td>
<td>6.15%</td>
<td>3.75%</td>
<td>2.24%</td>
<td>1.33%</td>
<td>0.79%</td>
<td>0.53%</td>
<td>0.36%</td>
<td>0.24%</td>
</tr>
<tr>
<td>0%</td>
<td>0.24%</td>
<td>0.63%</td>
<td>1.48%</td>
<td>3.19%</td>
<td>6.17%</td>
<td>9.78%</td>
<td>5.88%</td>
<td>3.37%</td>
<td>1.89%</td>
<td>1.03%</td>
<td>0.56%</td>
<td>0.35%</td>
<td>0.22%</td>
<td>0.14%</td>
</tr>
<tr>
<td></td>
<td>0.24%</td>
<td>0.63%</td>
<td>1.49%</td>
<td>3.19%</td>
<td>6.18%</td>
<td>9.78%</td>
<td>5.87%</td>
<td>3.38%</td>
<td>1.89%</td>
<td>1.03%</td>
<td>0.56%</td>
<td>0.35%</td>
<td>0.22%</td>
<td>0.14%</td>
</tr>
<tr>
<td>-25%</td>
<td>0.35%</td>
<td>0.81%</td>
<td>1.74%</td>
<td>3.44%</td>
<td>6.33%</td>
<td>9.73%</td>
<td>5.61%</td>
<td>3.02%</td>
<td>1.54%</td>
<td>0.76%</td>
<td>0.36%</td>
<td>0.21%</td>
<td>0.12%</td>
<td>0.07%</td>
</tr>
<tr>
<td></td>
<td>0.36%</td>
<td>0.83%</td>
<td>1.76%</td>
<td>3.48%</td>
<td>6.36%</td>
<td>9.73%</td>
<td>5.59%</td>
<td>2.97%</td>
<td>1.49%</td>
<td>0.72%</td>
<td>0.34%</td>
<td>0.19%</td>
<td>0.11%</td>
<td>0.06%</td>
</tr>
<tr>
<td>-50%</td>
<td>0.46%</td>
<td>0.98%</td>
<td>1.97%</td>
<td>3.68%</td>
<td>6.47%</td>
<td>9.68%</td>
<td>5.34%</td>
<td>2.63%</td>
<td>1.16%</td>
<td>0.48%</td>
<td>0.19%</td>
<td>0.09%</td>
<td>0.04%</td>
<td>0.02%</td>
</tr>
<tr>
<td></td>
<td>0.47%</td>
<td>1.00%</td>
<td>1.99%</td>
<td>3.72%</td>
<td>6.53%</td>
<td>9.72%</td>
<td>5.35%</td>
<td>2.58%</td>
<td>1.09%</td>
<td>0.41%</td>
<td>0.15%</td>
<td>0.07%</td>
<td>0.03%</td>
<td>0.01%</td>
</tr>
</tbody>
</table>

† For each correlation, first row is Heston fitted smile, second is calibrated joint density.
<table>
<thead>
<tr>
<th>( \rho )</th>
<th>2.00</th>
<th>2.10</th>
<th>2.20</th>
<th>2.31</th>
<th>2.41</th>
<th>2.51</th>
<th>2.61</th>
<th>2.71</th>
<th>2.82</th>
<th>2.92</th>
<th>3.02</th>
<th>3.10</th>
<th>3.17</th>
<th>3.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>9.36%</td>
<td>9.30%</td>
<td>9.38%</td>
<td>9.60%</td>
<td>9.96%</td>
<td>10.39%</td>
<td>10.85%</td>
<td>11.31%</td>
<td>11.73%</td>
<td>12.14%</td>
<td>12.51%</td>
<td>12.77%</td>
<td>13.03%</td>
<td>13.27%</td>
</tr>
<tr>
<td>25%</td>
<td>10.49%</td>
<td>10.27%</td>
<td>10.12%</td>
<td>10.07%</td>
<td>10.15%</td>
<td>10.34%</td>
<td>10.61%</td>
<td>10.91%</td>
<td>11.23%</td>
<td>11.55%</td>
<td>11.87%</td>
<td>12.09%</td>
<td>12.31%</td>
<td>12.51%</td>
</tr>
<tr>
<td>0%</td>
<td>11.38%</td>
<td>11.02%</td>
<td>10.72%</td>
<td>10.49%</td>
<td>10.33%</td>
<td>10.29%</td>
<td>10.35%</td>
<td>10.49%</td>
<td>10.68%</td>
<td>10.90%</td>
<td>11.13%</td>
<td>11.31%</td>
<td>11.49%</td>
<td>11.66%</td>
</tr>
<tr>
<td>-25%</td>
<td>12.11%</td>
<td>11.67%</td>
<td>11.24%</td>
<td>10.84%</td>
<td>10.50%</td>
<td>10.24%</td>
<td>10.07%</td>
<td>10.01%</td>
<td>10.05%</td>
<td>10.15%</td>
<td>10.28%</td>
<td>10.39%</td>
<td>10.51%</td>
<td>10.63%</td>
</tr>
<tr>
<td>-50%</td>
<td>12.76%</td>
<td>12.22%</td>
<td>11.70%</td>
<td>11.17%</td>
<td>10.66%</td>
<td>10.19%</td>
<td>9.78%</td>
<td>9.47%</td>
<td>9.28%</td>
<td>9.21%</td>
<td>9.21%</td>
<td>9.23%</td>
<td>9.29%</td>
<td>9.36%</td>
</tr>
</tbody>
</table>

For each correlation, first row is Heston fitted smile, second is calibrated joint density.
Appendix E: Construction of joint density via finite difference

E.1. Fwd Kolmogorov for the Lognormal Volatility specialization

As hinted in Remark (4.1), performing a consistency analysis of the recovered density by transitioning its Forward Kolmogorov equation, implies having a clear structure of coefficients in the infinitesimal generator, i.e. implies specifying a priori the SDE that is assumed to generate the smile. This consistency analysis in order to test the adequacy of the recovered density can only be performed on a case-by-case basis, i.e. after specializing the generic SDE to a particular form. In the following, the lognormal case will primarily be explored. References to more general processes will be given later on.

Following (Øksendal 2003, Chap. 8), for \( X_t \) an Itô process with infinitesimal generator

\[
\mathcal{L} f(y) := \sum_{i,j} a_{i,j}(y) \frac{\partial^2 f}{\partial y_i \partial y_j} + \sum_i b_{i,j}(y) \frac{\partial f}{\partial y_i}, \quad f \in \mathcal{C}_0^2
\]

letting

\[
\mathcal{L}^* \psi(t,x,y) := \sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j} (a_{i,j} \psi_t) - \sum_i \frac{\partial f}{\partial y_i} (b_i(y) \psi_t)
\]

be the adjoint operator to \( \mathcal{L} \) in a suitable defined space, where as usual inner product relationship holds, i.e

\[
\langle \mathcal{L} x, y \rangle = \langle x, \mathcal{L}^* y \rangle
\]

the IVP satisfied by the transition density \( \psi \) is

\[
\begin{align*}
\frac{d}{dt} \psi(x,y,t) dt = \mathcal{L}^* \psi(x,y,t) \\
\psi(x,y,t = 0) = \delta(x - x_0)\delta(y - y_0)
\end{align*}
\]

(E1)

In the following, given a lognormal volatility process

\[
d\sigma_t = a(\sigma_t, t) dt + b(\sigma_t, t) dW_t = \xi \sigma_t dW_t
\]

(E2)

the Forward Equation satisfied by the transition density \( \psi(\sigma_t, I_t) \) will be evolved by means of a finite difference schema. To be more specific, given the specific form (E2) for the volatility process the Forward Equation satisfied by the transition density \( \psi(\sigma_T, \sigma_t, I_t) \) is evolved. Prior to doing that, two preliminary steps are performed, which should give confidence about each individual density, and so give confidence about the joint. For lognormal volatility the evolution of

(i) \( \psi(\sigma_T, \sigma_t) \) and
(ii) \( \psi(\sigma_T, I_t) \)

will now be explored in turn.

E.1.1. Evolving \( \psi(\sigma, \sigma_t) \) for lognormal volatility. The transition density of joint

\[
\psi(x,y,t) := \psi(\sigma, \sigma_t, t)
\]
can be evolved forward. It must obey (E1) where
\[ \mathcal{L}^* \psi := \mathcal{L}^* \psi(x, y, t) \]
\[
= \left( -\frac{1}{2} \frac{\partial}{\partial x} \left( \mathbb{E}[(dx)] \right) - \frac{\partial}{\partial y} \left( \mathbb{E}[(dy)] \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( \mathbb{E}[(dx)^2] \right) + \frac{\partial^2}{\partial x \partial y} \left( \mathbb{E}[dxdy] \right) + \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( \mathbb{E}[(dy)^2] \right) \right) \psi(x, y, t) \]

By Eq. (E2), \( \mathbb{E}[dx] = 0, \mathbb{E}[(dx)^2] = \xi^2 \sigma(t)^2 dt, \mathbb{E}[dy] = \sigma(t)^2 dt, \mathbb{E}[(dy)^2] = 0, \mathbb{E}[dxdy] = 0 \) and so in this case (E1) becomes
\[
\begin{cases}
\frac{\partial}{\partial t} \psi(x, y, t) = \left( -\frac{\partial}{\partial y} (\xi^2 \sigma(t)^2) \right) \psi(x, y, t) \\
\psi(x, y, t = 0) = \delta(x - x_0)\delta(y - y_0) 
\end{cases}
\]

**E.1.2. Evolving \( \psi(\sigma, I_{\sigma dw}) \) for lognormal volatility.** The transition density of joint
\[ \psi(x, y, t) := \psi(\sigma, I_t, t) \]
can be evolved forward. It must obey (E1) where
\[ \mathcal{L}^* \psi := \mathcal{L}^* \psi(x, y, t) \]
\[
= \left( -\frac{1}{2} \frac{\partial}{\partial x} \left( \mathbb{E}[(dx)] \right) - \frac{\partial}{\partial y} \left( \mathbb{E}[(dy)] \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( \mathbb{E}[(dx)^2] \right) + \frac{\partial^2}{\partial x \partial y} \left( \mathbb{E}[dxdy] \right) + \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( \mathbb{E}[(dy)^2] \right) \right) \psi(x, y, t) \]

By Eq. (E2), \( \mathbb{E}[dx] = 0, \mathbb{E}[(dx)^2] = \xi^2 \sigma(t)^2 dt, \mathbb{E}[dy] = 0, \mathbb{E}[(dy)^2] = \sigma(t)^2 dt, \mathbb{E}[dxdy] = \xi \sigma(t)^2 dt \) and so in this case (E1) becomes
\[
\begin{cases}
\frac{\partial}{\partial t} \psi(x, y, t) = \left( \frac{\xi^2}{2} \frac{\partial^2}{\partial x^2} (\sigma(t)^2) + \frac{\xi}{\frac{\partial}{\partial y}} (\sigma(t)^2) \right) \psi(x, y, t) \\
\psi(x, y, t = 0) = \delta(x - x_0)\delta(y - y_0) 
\end{cases}
\]

**E.2. Fwd Kolmogorov for the Joint Density**
The transition density of joint
\[ \psi(x, y, t) := \psi(\partial_t, I_t, t) \]
can be evolved forward. It must obey (E1) where
\[ \mathcal{L}^* \psi := \mathcal{L}^* \psi(x, y, t) \]
\[
= \left( -\frac{1}{2} \frac{\partial}{\partial x} \left( \mathbb{E}[(dx)] \right) - \frac{\partial}{\partial y} \left( \mathbb{E}[(dy)] \right) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left( \mathbb{E}[(dx)^2] \right) + \frac{\partial^2}{\partial x \partial y} \left( \mathbb{E}[dxdy] \right) + \frac{1}{2} \frac{\partial^2}{\partial y^2} \left( \mathbb{E}[(dy)^2] \right) \right) \psi(x, y, t) \]

By Eq. (E2), \( \mathbb{E}[dx] = \sigma(t)^2 dt, \mathbb{E}[(dx)^2] = 0, \mathbb{E}[dy] = \mathbb{E}[\int_0^t \sigma(s) dW_s] = 0, \mathbb{E}[(dy)^2] = \sigma(t)^2 dt, \mathbb{E}[dxdy] = \mathbb{E}[\int_0^t ds \sigma(s)^2 \left( \int_0^t \sigma(s) dW_s \right)] = 0 \) and so in this case (E1) becomes
\[
\begin{cases}
\frac{\partial}{\partial t} \psi(x, y, t) = \left( -\frac{\partial}{\partial x} (\sigma(t)^2) + \frac{\partial^2}{\partial y^2} (\sigma(t)^2) \right) \psi(x, y, t) \\
\psi(x, y, t = 0) = \delta(x - x_0)\delta(y - y_0) 
\end{cases}
\]
References


Yor, M., Exponential Functionals of Brownian Motion and Related Processes, 2001 (Springer: Berlin).