INVERSION OF CONVEX ORDERING: LOCAL VOLATILITY DOES NOT MAXIMIZE THE PRICE OF VIX FUTURES

BEATRICE ACCIAIO AND JULIEN GUYON

Abstract. It has often been stated that, within the class of continuous stochastic volatility models calibrated to vanillas, the price of a VIX future is maximized by the Dupire local volatility model. In this article we prove that this statement is incorrect: we build a continuous stochastic volatility model in which a VIX future is *strictly more expensive* than in its associated local volatility model. More generally, in this model, strictly convex payoffs on a squared VIX are strictly cheaper than in the associated local volatility model. This corresponds to an *inversion of convex ordering* between local and stochastic variances, when moving from instantaneous variances to squared VIX, as convex payoffs on instantaneous variances are always cheaper in the local volatility model. We thus prove that this inversion of convex ordering, which is observed in the SPX market for short VIX maturities, can be produced by a continuous stochastic volatility model. We also prove that the model can be extended so that, as suggested by market data, the convex ordering is preserved for long maturities.

1. Introduction

For simplicity, let us assume zero interest rates, repos, and dividends. Let $\mathcal{F}_t$ denote the market information available up to time t. We consider continuous stochastic volatility models on the SPX index of the form

$$\frac{dS_t}{S_t} = \sigma_t \, dW_t, \qquad S_0 = s_0, \tag{1.1}$$

where $W = (W_t)_{t\geq 0}$ denotes a standard one-dimensional $(\mathcal{F}_t)$-Brownian motion, $\sigma = (\sigma_t)_{t\geq 0}$ is an $(\mathcal{F}_t)$-adapted process such that $\int_0^t \sigma_s^2 \, ds < \infty$ a.s. for all $t \geq 0$, and $s_0 > 0$ is the initial SPX price. By continuous model we mean that the SPX has no jump, while the volatility process σ may be discontinuous. The local volatility function associated to Model (1.1) is the function σ_{loc} defined by

$$\sigma_{\mathrm{loc}}^2(t,x) := \mathbb{E}[\sigma_t^2 | S_t = x]. \tag{1.2}$$

The associated local volatility model is defined by:

$$\frac{dS_t^{\mathrm{loc}}}{S_t^{\mathrm{loc}}} = \sigma_{\mathrm{loc}}(t, S_t^{\mathrm{loc}}) \, dW_t, \qquad S_0^{\mathrm{loc}} = s_0.$$

From [11], the marginal distributions of the processes $(S_t)_{t\geq 0}$ and $(S_t^{\mathrm{loc}})_{t\geq 0}$ agree:

$$\forall t \geq 0, \quad S_t^{\mathrm{loc}} \stackrel{(d)}{=} S_t. \tag{1.3}$$

Let $T \geq 0$. By definition, the (idealized) VIX at time T is the implied volatility of a 30 day log-contract on the SPX index starting at T. For continuous models (1.1), this translates into

$$\mathrm{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2 \, dt \middle| \mathcal{F}_T\right] = \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}\left[\sigma_t^2 \middle| \mathcal{F}_T\right] dt, \tag{1.4}$$

where $\tau = \frac{30}{365}$ (30 days). In the associated local volatility model, since by the Markov property of $(S_t^{\mathrm{loc}})_{t\geq 0}$, $\mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|\mathcal{F}_T] = \mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|S_T^{\mathrm{loc}}]$, the VIX, denoted by $\mathrm{VIX}_{\mathrm{loc},T}$, satisfies

$$\mathrm{VIX}_{\mathrm{loc},T}^2 = \frac{1}{\tau}\int_T^{T+\tau} \mathbb{E}[\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})|S_T^{\mathrm{loc}}] \, dt = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}}) \, dt \middle| S_T^{\mathrm{loc}}\right]. \tag{1.5}$$

Department of Statistics, London School of Economics and Political Science
Quantitative Research, Bloomberg L.P.
E-mail address: b.acciaio@lse.ac.uk, jguyon2@bloomberg.net.
Date: December 15, 2019.
Key words and phrases. VIX, VIX futures, stochastic volatility, local volatility, convex order, inversion of convex ordering.

Note that VIX_T^2 and $\mathrm{VIX}_{\mathrm{loc},T}^2$ have the same mean:

$$\mathbb{E}\left[\mathrm{VIX}_T^2\right] = \mathbb{E}\left[\mathrm{VIX}_{\mathrm{loc},T}^2\right] = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2\, dt\right]. \tag{1.6}$$

It has often been stated that, within the class of continuous stochastic volatility models calibrated to vanillas, the price of a VIX future is maximized by Dupire's local volatility model. For example, in a general discussion in the introduction of [4] about the difficulty of jointly calibrating a stochastic volatility model to both SPX and VIX smiles, De Marco and Henry-Labordère approximate the VIX by the instantaneous volatility, i.e., $\mathrm{VIX}_T \approx \sigma_T$ and $\mathrm{VIX}_{\mathrm{loc},T} \approx \sigma_{\mathrm{loc}}(T, S_T^{\mathrm{loc}})$, and, using Jensen's inequality and (1.3), they conclude that "within [the] class of continuous stochastic volatility models calibrated to vanillas, the VIX future is bounded from above by the Dupire local volatility model":

$$\begin{aligned}\mathbb{E}[\mathrm{VIX}_T] \approx \mathbb{E}[\sigma_T] = \mathbb{E}\left[\sqrt{\sigma_T^2}\right] &= \mathbb{E}\left[\mathbb{E}\left[\sqrt{\sigma_T^2}\middle| S_T\right]\right] \\ &\le \mathbb{E}\left[\sqrt{\mathbb{E}\left[\sigma_T^2 | S_T\right]}\right] = \mathbb{E}\left[\sigma_{\mathrm{loc}}(T, S_T)\right] = \mathbb{E}\left[\sigma_{\mathrm{loc}}(T, S_T^{\mathrm{loc}})\right] \approx \mathbb{E}\left[\mathrm{VIX}_{\mathrm{loc},T}\right].\end{aligned}$$

Similarly, one would conclude that within the class of continuous stochastic volatility models calibrated to vanillas, the price of convex options on the *squared* VIX is *minimized* by the local volatility model: for any convex function f, such as the call or put payoff function,

$$\begin{aligned}\mathbb{E}\left[f\left(\mathrm{VIX}_T^2\right)\right] &\approx \mathbb{E}\left[f\left(\sigma_T^2\right)\right] = \mathbb{E}\left[\mathbb{E}\left[f\left(\sigma_T^2\right)\middle| S_T\right]\right] \\ &\ge \mathbb{E}\left[f\left(\mathbb{E}\left[\sigma_T^2 | S_T\right]\right)\right] = \mathbb{E}\left[f\left(\sigma_{\mathrm{loc}}^2(T, S_T)\right)\right] = \mathbb{E}\left[f\left(\sigma_{\mathrm{loc}}^2(T, S_T^{\mathrm{loc}})\right)\right] \approx \mathbb{E}\left[f\left(\mathrm{VIX}_{\mathrm{loc},T}^2\right)\right].\end{aligned} \tag{1.7}$$

(The (correct) fact that $\mathbb{E}\left[f\left(\sigma_T^2\right)\right] \ge \mathbb{E}\left[f\left(\sigma_{\mathrm{loc}}^2(T, S_T^{\mathrm{loc}})\right)\right]$ had already been noticed by Dupire in [6].)

In this article, we prove that these statements are in fact incorrect. Even if 30 days is a relatively short horizon, it cannot be harmlessly ignored. VIX are implied volatilities (of SPX options maturing 30 days later), not instantaneous volatilities. We can actually build continuous stochastic volatility models, i.e., processes (σ_t), such that

$$\mathbb{E}[\mathrm{VIX}_T] > \mathbb{E}\left[\mathrm{VIX}_{\mathrm{loc},T}\right] \tag{1.8}$$

and, more generally, such that for any strictly convex function f,

$$\mathbb{E}\left[f\left(\mathrm{VIX}_T^2\right)\right] < \mathbb{E}\left[f\left(\mathrm{VIX}_{\mathrm{loc},T}^2\right)\right]. \tag{1.9}$$

(Our counterexample actually works for any $\tau > 0$.) Not only do we find one convex function f such that (1.9) holds, we actually build a model in which (1.9) holds for any strictly convex function f. Actually, we prove an *inversion of convex ordering*: Despite the fact that $\sigma_{\mathrm{loc}}^2(t, S_t^{\mathrm{loc}})$ is smaller than σ_t^2 in convex order for all $t \in [T, T+\tau]$ (see (1.7)), $\mathrm{VIX}_{\mathrm{loc},T}^2$ is strictly *larger* than VIX_T^2 in convex order. Interestingly, Guyon [8, 10] has reported that for short maturities T, the market exhibits this inversion of convex ordering: the distribution of $\mathrm{VIX}_{\mathrm{loc},T}^2$ (computed with the market-implied Dupire local volatility) is strictly *larger* than the distribution of VIX_T^2 (implied from the market prices of VIX options) in convex order.[1]

Guyon [9, 10] has shown that when the (typically negative) spot-vol correlation is large enough in absolute value, (a) traditional stochastic volatility models with large mean reversion, and (b) rough volatility models with small Hurst exponent, do exhibit this inversion of convex ordering. The fast mean reversion or small Hurst exponent ensures that $\mathcal{F}_T$ contains little information on σ_t, $t > T$, so VIX_T^2 is almost constant, while the large spot-vol correlation yields a non-flat local volatility, and as a result it can be numerically checked that VIX_T^2 is strictly smaller than $\mathrm{VIX}_{\mathrm{loc},T}^2$ in convex order for short maturities. However, it is very difficult to mathematically prove the inversion of convex ordering in these models.[2] In order to reproduce this inversion, our idea is to choose a more extreme model in which the volatility process σ is such that $(\sigma_t)_{t\in[T,T+\tau]}$ is independent of $\mathcal{F}_T$, so that $\mathcal{F}_T$, instead of revealing little information on $(\sigma_t)_{t\in[T,T+\tau]}$, contains *no information at all* on it. This is easily achieved by choosing σ_t deterministic (e.g., constant) on $[0, T]$. In this extreme case, since VIX_T^2 and $\mathrm{VIX}_{\mathrm{loc},T}^2$ have the same mean (recall (1.6)), to prove that VIX_T^2 is strictly smaller than

[1] To be precise, this convex ordering holds for reasonable, consistent extrapolations of the SPX and VIX market smiles (see [10]). For a definition and characterization of *strict* convex ordering, see [10, Appendix C].

[2] Strictly speaking, those models may not satisfy the inversion of convex ordering at very low strikes, see [10, Remark 6].

$\mathrm{VIX}^2_{\mathrm{loc},T}$ in convex order, it is enough that $\mathrm{VIX}^2_{\mathrm{loc},T}$ be not a.s. constant. Then (1.9) holds for any strictly convex function f, and in particular applying (1.9) with $f(y) = -\sqrt{y}$ yields (1.8).

Clearly, in order for $\mathrm{VIX}^2_{\mathrm{loc},T}$ to be non-constant, the local volatility cannot be constant as a function of S, dt-a.e. in $[T, T+\tau]$. There are many ways to achieve this, e.g., through volatility of volatility, and it is easy to numerically verify that $\mathrm{VIX}^2_{\mathrm{loc},T}$ (estimated from (1.5), e.g., using kernel regressions) is non-constant. However, the main mathematical difficulty here is to *prove* this result. To this end, we will consider models where the non-constant local volatility can be derived in closed form. The simplest such model is the one where, after some given time $t_0 \in (T, T+\tau)$, the volatility can take only two values depending on a coin toss independent of $\mathcal{F}_T$ (see Section 2). This counterexample, which is inspired by [2] where Beiglböck, Friz, and Sturm use a similar model to prove that local volatility does not minimize the price of options on realized variance, will be generalized in Section 3. Even though the resulting models may not be realistic, the proofs emphasize the fact that, like in the realistic models described above, the inversion of convex ordering in the VIX market most likely results from the loss of information that comes with projecting the future realized variance $\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2\, dt$ onto $\mathcal{F}_T$, rather than from averaging the instantaneous variance over 30 days (see [10] for detailed numerical experiments).

Interestingly enough, the realistic models (a) and (b) mentioned above reproduce another characteristic of the SPX/VIX markets: that for longer maturities (typically, 3-6 months) the distributions of $\mathrm{VIX}^2_{\mathrm{loc},T}$ and VIX^2_T become non-rankable in convex order, and for even longer maturities the latter becomes strictly larger than the former in convex order, i.e., the inversion of convex ordering vanishes as T increases. As explained in [10, Remark 4], in those models this results from the fact that, like the market local volatility, the associated local volatility flattens over time as a function of S. Once again, it is not clear how this could be mathematically proved. We are therefore reproducing this behavior with a simplistic, extreme stochastic local volatility model in which, by construction, beyond some fixed maturity, the associated local volatility is constant, i.e., perfectly flat.

The remainder of the article is structured as follows. In Section 2 we derive the simple counterexample described above, about inversion of convex ordering for short maturities. This counterexample is then generalized in Section 3. Eventually in Section 4 we explain how the model can be extended so that the convex ordering is preserved for long maturities.

2. A simple counterexample

Inspired by [2], we fix $T > 0$ and consider the following volatility process:

$$\sigma_t = \begin{cases} \sigma_0 & \text{if } t < T + \frac{\tau}{2} \\ \overline{\sigma} & \text{if } t \geq T + \frac{\tau}{2} \text{ and } U = 1 \\ \underline{\sigma} & \text{if } t \geq T + \frac{\tau}{2} \text{ and } U = -1 \end{cases}, \tag{2.1}$$

where $\underline{\sigma} < \sigma_0 < \overline{\sigma}$ are three positive constants and U denotes the result of a fair coin toss, independent of $\mathcal{F}_T$ (e.g., known only at a time $t \in (T, T + \frac{\tau}{2}]$).

Proposition 1. *The stochastic volatility model in* (1.1)*, with volatility process described in* (2.1)*, satisfies* (1.9)*. In particular, VIX futures are strictly more expensive than in their associated local volatility model.*

Proof. Let us denote

$$\sigma_+(t) = \begin{cases} \sigma_0 & \text{if } t < T + \frac{\tau}{2} \\ \overline{\sigma} & \text{if } t \geq T + \frac{\tau}{2} \end{cases}, \qquad \sigma_-(t) = \begin{cases} \sigma_0 & \text{if } t < T + \frac{\tau}{2} \\ \underline{\sigma} & \text{if } t \geq T + \frac{\tau}{2} \end{cases},$$

so that σ_t is given by $\sigma_+(t)$ or $\sigma_-(t)$ depending on the coin toss U. As $(\sigma_t)_{t\in[T,T+\tau]}$ is independent of $\mathcal{F}_T$, VIX^2_T is a.s. constant:

$$\mathrm{VIX}^2_T = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2\, dt\right] = \frac{1}{2}\left(\sigma_0^2 + \frac{\overline{\sigma}^2 + \underline{\sigma}^2}{2}\right).$$

Since this is also the mean of $\mathrm{VIX}^2_{\mathrm{loc},T}$, in order to prove (1.9), it is enough to prove that $\mathrm{VIX}^2_{\mathrm{loc},T}$ is not a.s. constant.

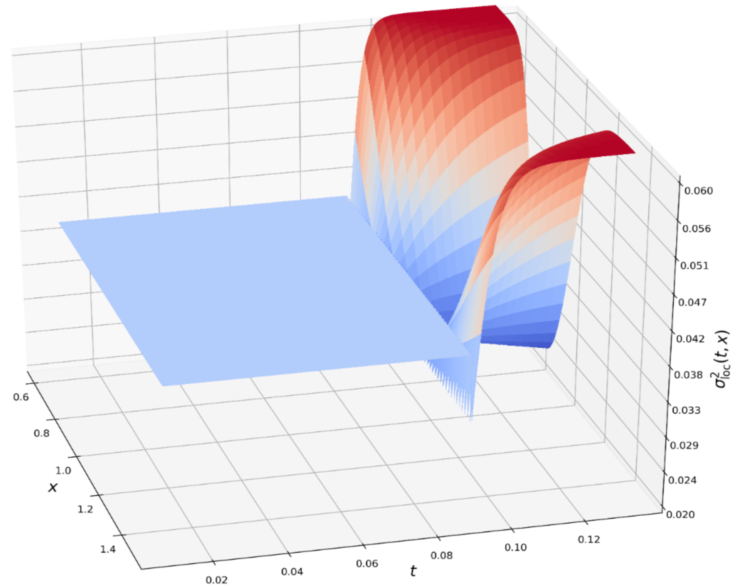

FIGURE 2.1. Graph of $(t,x) \mapsto \sigma^2_{\text{loc}}(t,x)$ for $T = 0.05$, $\sigma_0 = 0.2$, $\overline{\sigma} = 0.25$, $\underline{\sigma} = 0.02$, $s_0 = 1$

Due to the very simple form of Model (2.1), we know the local volatility in closed form:

$$\sigma^2_{\text{loc}}(t,x) = \frac{p_+(t,x)\sigma_+^2(t) + p_-(t,x)\sigma_-^2(t)}{p_+(t,x) + p_-(t,x)}, \tag{2.2}$$

where $p_\pm(t,\cdot)$ is the density of the process $(S_t^\pm)_{t\geq 0}$ with dynamics $\frac{dS_t^\pm}{S_t^\pm} = \sigma_\pm(t)\,dW_t$, $S_0 = s_0$, i.e.,

$$p_\pm(t,x) = \frac{1}{x\sqrt{2\pi\Sigma_\pm(t)}} \exp\left(-\frac{1}{2}\left(\frac{\ln\frac{x}{s_0}}{\sqrt{\Sigma_\pm(t)}} + \frac{1}{2}\sqrt{\Sigma_\pm(t)}\right)^2\right), \qquad \Sigma_\pm(t) = \int_0^t \sigma_\pm^2(s)\,ds.$$

Figure 2.1 shows the shape of σ^2_{loc}. Note in particular that σ_{loc} takes values in $(\underline{\sigma}, \overline{\sigma})$ and that

$$\forall t \in \left(T + \frac{\tau}{2}, T\right], \qquad \lim_{x\to+\infty} \sigma_{\text{loc}}(t,x) = \overline{\sigma}. \tag{2.3}$$

Let us define

$$\psi(x) := \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\text{loc}}(t, S_t^{\text{loc}})\,dt \middle| S_T^{\text{loc}} = x\right] \tag{2.4}$$

so that $\text{VIX}^2_{\text{loc},T} = \psi\left(S_T^{\text{loc}}\right)$. Note that

$$\forall x > 0, \qquad \psi(x) < \ell := \frac{1}{2}\left(\sigma_0^2 + \overline{\sigma}^2\right).$$

Since S_T^{loc} has support $\mathbb{R}_+$, in order to prove that $\text{VIX}^2_{\text{loc},T}$ is not a.s. constant, it is enough to prove that ψ tends to ℓ when x tends to $+\infty$. This follows from the next lemma. $\square$

Lemma 2. *In Model (2.1), the function ψ defined by (2.4) satisfies*

$$\lim_{x\to+\infty} \psi(x) = \ell.$$

Proof. Note that $\psi(x) = \frac{1}{2}\left(\sigma_0^2 + \varphi(x)\right)$, where

$$\varphi(x) := \mathbb{E}\left[\frac{2}{\tau}\int_{T+\frac{\tau}{2}}^{T+\tau} \sigma^2_{\text{loc}}(t, S_t^{\text{loc}})\,dt \middle| S_T^{\text{loc}} = x\right],$$

so it is enough to prove that $\varphi(x)$ tends to $\overline{\sigma}^2$ when x tends to $+\infty$. Let $\varepsilon > 0$ and $\varepsilon' := \left(1 + 2\left(\overline{\sigma}^2 - \underline{\sigma}^2\right)\right)^{-1}\varepsilon$. Let us denote $L_t := \ln(S_t^{\mathrm{loc}})$, whose dynamics is given by

$$dL_t = -\frac{1}{2}\sigma_{\mathrm{loc}}^2\left(t, e^{L_t}\right)\, dt + \sigma_{\mathrm{loc}}\left(t, e^{L_t}\right)\, dW_t, \qquad L_0 = \ln s_0.$$

Since σ_{loc} is bounded, it is easily checked that $c := \sup_{t\in[T,T+\tau],x\in\mathbb{R}} \mathbb{E}[(L_t - L_T)^2 | L_T = x] < +\infty$. Let $\Delta := \sqrt{\frac{c}{\varepsilon'}}$. Then

$$\forall t \in [T, T+\tau],\ \ \forall x \in \mathbb{R}, \quad \mathbb{P}(|L_t - L_T| \geq \Delta | L_T = x) \leq \frac{\mathbb{E}[(L_t - L_T)^2 | L_T = x]}{\Delta^2} \leq \frac{c}{\Delta^2} = \varepsilon'. \tag{2.5}$$

We have

$$\overline{\sigma}^2 - \varphi(e^x) = \frac{2}{\tau}\int_{T+\frac{\tau}{2}}^{T+\tau} \mathbb{E}\left[\left(\overline{\sigma}^2 - \sigma_{\mathrm{loc}}^2(t, e^{L_t})\right)\big| L_T = x\right] dt = I_1(x) + I_2(x) + I_3(x),$$

where

$$\begin{aligned}
I_1(x) &:= \frac{2}{\tau}\int_{T+\frac{\tau}{2}}^{T+\tau} \mathbb{E}\left[\left(\overline{\sigma}^2 - \sigma_{\mathrm{loc}}^2(t, e^{L_t})\mathbf{1}_{L_t \leq L_T - \Delta}\right)\big| L_T = x\right] dt,\\
I_2(x) &:= \frac{2}{\tau}\int_{T+\frac{\tau}{2}}^{T+\frac{\tau}{2}(1+\varepsilon')} \mathbb{E}\left[\left(\overline{\sigma}^2 - \sigma_{\mathrm{loc}}^2(t, e^{L_t})\mathbf{1}_{L_t > L_T - \Delta}\right)\big| L_T = x\right] dt,\\
I_3(x) &:= \frac{2}{\tau}\int_{T+\frac{\tau}{2}(1+\varepsilon')}^{T+\tau} \mathbb{E}\left[\left(\overline{\sigma}^2 - \sigma_{\mathrm{loc}}^2(t, e^{L_t})\mathbf{1}_{L_t > L_T - \Delta}\right)\big| L_T = x\right] dt.
\end{aligned}$$

Recall that σ_{loc} takes values in $(\underline{\sigma}, \overline{\sigma})$. From (2.5), $0 \leq I_1(x) \leq \left(\overline{\sigma}^2 - \underline{\sigma}^2\right)\varepsilon'$ for all $x \in \mathbb{R}$. Obviously, $0 \leq I_2(x) \leq \left(\overline{\sigma}^2 - \underline{\sigma}^2\right)\varepsilon'$ for all $x \in \mathbb{R}$. Moreover, it is easy to check that the convergence (2.3) is uniform w.r.t. $t \in [T + \frac{\tau}{2}(1+\varepsilon'), T+\tau]$: there exists x^* such that

$$\forall t \in \left[T + \frac{\tau}{2}(1+\varepsilon'), T+\tau\right], \quad \forall x \geq x^*, \qquad 0 \leq \overline{\sigma}^2 - \sigma_{\mathrm{loc}}^2(t, e^x) \leq \varepsilon'.$$

As a consequence, for all $x \geq x^* + \Delta$, $0 \leq I_3(x) \leq \varepsilon'$. Finally,

$$\forall x \geq x^* + \Delta, \qquad 0 \leq \overline{\sigma}^2 - \varphi(e^x) \leq \left(1 + 2\left(\overline{\sigma}^2 - \underline{\sigma}^2\right)\right)\varepsilon' = \varepsilon.$$

We have thus proved that $\varphi(e^x)$, hence $\varphi(x)$, tends to $\overline{\sigma}^2$ when x tends to $+\infty$. □

Remark 3. Note that, if we fix $t_1 \in (0, \tau)$ and define

$$\sigma_t = \begin{cases} \sigma_0 & \text{if } t < t_1\\ \overline{\sigma} & \text{if } t \geq t_1 \text{ and } U = 1\\ \underline{\sigma} & \text{if } t \geq t_1 \text{ and } U = -1 \end{cases},$$

with U only known at time t_1, then we have built a model where the inversion of convex ordering holds for every short maturity $T \in (0, t_1)$.

3. Generalization

In this section, we generalize the example presented in Section 2, to show that the desired inversion of convex ordering can be obtained with a more interesting structure for the volatility. We fix $0 < t_1 < \tau < t_2 = t_1 + \tau$, and define a càdlàg process σ on $[0, t_2)$, which is independent of $\mathcal{F}^W$, the filtration generated by W. We start by setting σ_t constant equal to $\sigma_0 > 0$ for $t \in [0, t_1)$. This ensures that $\mathcal{F}_T = \mathcal{F}_T^{S,W} = \mathcal{F}_T^W$, and as a consequence $(\sigma_t)_{t\in[T,T+\tau]}$ is independent of $\mathcal{F}_T$ for all $0 < T < t_1$, thus VIX_T^2 is constant:

$$\mathrm{VIX}_T^2 = \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma_t^2 dt\right].$$

We shall now define σ in $[t_1, t_2)$, with the aim of having $\mathrm{VIX}_{\mathrm{loc},T}^2$ not constant. Let $0 < \underline{v} \leq \sigma_0 \leq \overline{v}$, and $(\sigma_t)_{t\in[t_1,t_2)}$ take values in $[\underline{v}, \overline{v}]$. We denote by Λ the law of $(\sigma_t)_{t\in[0,t_2)}$ on $\mathcal{D} = \mathcal{D}[0, t_2)$, the space of càdlàg functions on $[0, t_2)$, and we assume it is not degenerate, that is, there is no $g \in \mathcal{D}$ such that $\Lambda = \delta_g$, where

δ_g denotes the Dirac mass at g. Note that, for $\Lambda = \frac{1}{2}(\delta_{\sigma_+} + \delta_{\sigma_-})$, and $T = t_1 - \tau/2$, we recover the example of Section 2.

For every path $g \in \mathcal{D}$, we denote by S^g the evolution of the stock price for this realization of σ, that is

$$\frac{dS_t^g}{S_t^g} = g(t)\, dW_t, \quad 0 \le t < t_2, \quad S_0^g = s_0,$$

and by $p_g(t,.)$ the density of the process S_t^g, that is

$$p_g(t,x) = \frac{1}{x\sqrt{2\pi\Sigma_g(t)}} \exp\left(-\frac{1}{2}\left(\frac{\ln \frac{x}{s_0}}{\sqrt{\Sigma_g(t)}} + \frac{1}{2}\sqrt{\Sigma_g(t)} \right)^2 \right), \qquad \Sigma_g(t) = \int_0^t g(s)^2\, ds.$$

The local volatility then takes the form

$$\sigma_{\mathrm{loc}}^2(t,x) = \int_{\mathcal{D}} g(t)^2 q_g(t,x)\, d\Lambda(g), \quad t \in [0,t_2), \tag{3.1}$$

where

$$q_g(t,x) = \frac{p_g(t,x)}{\int_{\mathcal{D}} p_h(t,x)\, d\Lambda(h)}.$$

Lemma 4. *The following limit holds for the local volatility:*

$$\lim_{x\to\infty} \sigma_{\mathrm{loc}}^2(t,x) = \frac{1}{\Lambda(A^{[t]})} \int_{A^{[t]}} g(t)^2\, d\Lambda(g) =: \overline{\sigma}(t)^2, \tag{3.2}$$

where

$$A^{[t]} := \{g \in \mathcal{D} : \Sigma_g(t) = \Lambda\text{-}\operatorname{ess\,sup}_{h\in\mathcal{D}} \Sigma_h(t)\}.$$

Proof. To study the limit of $\sigma_{\mathrm{loc}}^2(t,x)$ for $x \to \infty$, thanks to (3.1) and dominated convergence, we are reduced to consider the limit of $q_g(t,x)$. Note that

$$q_g(t,x)^{-1} = \int_{\mathcal{D}} F(g,h,t,x)\, d\Lambda(h), \tag{3.3}$$

where

$$F(g,h,t,x) := \sqrt{\tfrac{\Sigma_g(t)}{\Sigma_h(t)}} \exp\left\{ -\tfrac{1}{2}\left[\left(\ln \tfrac{x}{s_0}\right)^2 \left(\tfrac{1}{\Sigma_h(t)} - \tfrac{1}{\Sigma_g(t)} \right) + \tfrac{1}{4}(\Sigma_h(t) - \Sigma_g(t)) \right] \right\}.$$

By Fatou's lemma, $\lim_{x\to\infty} \int_{\mathcal{D}} F(g,h,t,x)\, d\Lambda(h) = +\infty$ as soon as $\Lambda(\mathcal{D}_{g,t}) > 0$, where

$$\mathcal{D}_{g,t} := \{h \in \mathcal{D} : \Sigma_h(t) > \Sigma_g(t)\},$$

which in turn implies $\lim_{x\to\infty} q_g(t,x) = 0$. On the other hand, if $\Lambda(\mathcal{D}_{g,t}) = 0$, then by dominated convergence $\lim_{x\to\infty} \int_{\mathcal{D}} F(g,h,t,x)\, d\Lambda(h) = \int_{\mathcal{D}} \lim_{x\to\infty} F(g,h,t,x)\, d\Lambda(h)$, being σ bounded and bounded away from zero. Now $\lim_{x\to\infty} F(g,h,t,x)$ equals zero when $\Sigma_h(t) < \Sigma_g(t)$, and one when $\Sigma_h(t) = \Sigma_g(t)$. This concludes the proof, noticing that

$$A^{[t]} = \{g \in \mathcal{D} : \Lambda(\mathcal{D}_{g,t}) = 0\}.$$

□

Proposition 5. *Consider the stochastic volatility model* (1.1)*. Let σ be constant equal to $\sigma_0 > 0$ in $[0,t_1)$, non-degenerate and independent of $\mathcal{F}^W$ in $[t_1,t_2)$, admitting only finitely many paths, so that*

$$\Lambda = \sum_{n=1}^N u_n \delta_{g_n}, \quad \text{with } N \in \mathbb{N},\ g_n \in \mathcal{D},\ u_n \ge 0,\ \textstyle\sum_{n=1}^N u_n = 1. \tag{3.4}$$

Then, for all maturities $T < t_1$, VIX futures are strictly more expensive than in their associated local volatility model.

With an abuse of notation, below we will write $q_n, \Sigma_n, F(n,m,t,x)$ instead of $q_{g_n}, \Sigma_{g_n}, F(g_n,g_m,t,x)$, to ease readability.

Proof. As in the example of Section 2, we consider the function

$$\psi(x) := \mathbb{E}\left[\frac{1}{\tau}\int_T^{T+\tau} \sigma^2_{\mathrm{loc}}(t, S^{\mathrm{loc}}_t)\, dt \middle| S^{\mathrm{loc}}_T = x\right]$$

and note that

$$\forall x > 0, \quad \psi(x) < \frac{1}{\tau}\int_T^{T+\tau} \overline{\sigma}(t)^2\, dt =: \ell,$$

being Λ non-degenerate. To prove the inversion of convex ordering for all maturities $T < t_1$, we will show that $\lim_{x\to\infty}\psi(x) = \ell$, that is,

$$\int_{t_1}^{T+\tau} \mathbb{E}[\sigma^2_{\mathrm{loc}}(t, S^{\mathrm{loc}}_t)|S^{\mathrm{loc}}_T = x]\ dt \ \xrightarrow[x\to\infty]{}\ \int_{t_1}^{T+\tau} \overline{\sigma}(t)^2\ dt. \tag{3.5}$$

Since Λ is discrete, and the functions $t \mapsto \Sigma_n(t)$ are continuous and bounded in $[t_1, t_2)$, this interval divides in countably many intervals $I_k = [a_k, b_k), k \in \mathbb{N}$, such that, in each open interval (a_k, b_k), the function $\overline{\sigma}$ defined in (3.2) coincides with one or more paths of σ. To be more precise, for every $k \in \mathbb{N}$, the sets $A^{[t]}$ coincide for every $t \in (a_k, b_k)$, say to a set A^k, and

$$\overline{\sigma}(t) = g_n(t) \ \text{ for } t \in (a_k, b_k), \quad \text{for all } g_n \in A^k. \tag{3.6}$$

To show the convergence in (3.5), we split the interval $[t_1, T+\tau]$ into subintervals $\widetilde{I}_k := [t_1, T+\tau]\cap I_k, k \in \mathbb{N}$, thus reducing ourselves to prove

$$\lim_{x\to\infty}\sum_{k\in\mathbb{N}}\int_{\widetilde{I}_k} \mathbb{E}[\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, S^{\mathrm{loc}}_t)|S^{\mathrm{loc}}_T = x]\ dt = \sum_{k\in\mathbb{N}}\lim_{x\to\infty}\int_{\widetilde{I}_k} \mathbb{E}[\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, S^{\mathrm{loc}}_t)|S^{\mathrm{loc}}_T = x]\ dt = 0, \tag{3.7}$$

by dominated convergence.

Fix $k \in \mathbb{N}$ and $\varepsilon_k > 0$, and set $\epsilon'_k := \min\{\epsilon_k(b_k - a_k + 3(\bar{v}^2 - \underline{v}^2))^{-1}, (b_k - a_k)/3\}$. We split the interval I_k into three subintervals

$$J'_k := [a_k, a_k + \epsilon'_k], \quad J_k := (a_k + \epsilon'_k, b_k - \epsilon'_k), \quad J''_k := [b_k - \epsilon'_k, b_k), \tag{3.8}$$

and we are going to show that $\sigma^2_{\mathrm{loc}}(t, x)$ converges uniformly to $\overline{\sigma}(t)^2$ w.r.t. $t \in J_k$, for $x \to \infty$.

Let $N^k := \{n \in \{1, ..., N\} : g_n \in A^k\}$ and note that the function $F(n, m, t, x)$ depends on the paths g_n and g_m only through $\Sigma_n(t)$ and $\Sigma_m(t)$, respectively. Therefore, from (3.3), we have

$$\frac{1}{q_n(t, x)} = \sum_{m=1}^{N} F(n, m, t, x)u_m = F(n, m_k, t, x)\Lambda(A^k) + \sum_{m\notin N^k} F(n, m, t, x)u_m, \quad t \in J_k, \tag{3.9}$$

for any $m_k \in N^k$, which reduces to $\Lambda(A^k) + \sum_{m\notin N^k} F(m_k, m, t, x)u_m$ for $n \in N^k$. Now, it is easy to verify that $F(n, m, t, x)$ converges to zero uniformly w.r.t. $t \in J_k$ whenever $n \in N^k$ and $m \notin N^k$. In particular, there is x_k such that, for all $x \geq x_k, t \in J_k$ and $n \in N^k$, $\sum_{m\notin N^k} F(n, m, t, x)u_m \leq \epsilon'_k\Lambda(A^k)^2\bar{v}^{-2}$, thus

$$\left|q_n(t, x) - \frac{1}{\Lambda(A^k)}\right| = \frac{\sum_{m\notin N^k} F(n, m, t, x)u_m}{\Lambda(A^k)(\Lambda(A^k) + \sum_{m\notin N^k} F(n, m, t, x)u_m)} \leq \epsilon'_k\bar{v}^{-2}. \tag{3.10}$$

Similarly, it is easily checked that $F(n, m, t, x)$ converges to $+\infty$ uniformly w.r.t. $t \in J_k$ whenever $n \notin N^k$ and $m \in N^k$. This gives the existence of y_k such that, for all $x \geq y_k, t \in J_k$ and $n \notin N^k, m \in N^k$,

$$F(n, m, t, x) \geq \bar{v}^2(\epsilon'_k\Lambda(A^k))^{-1},$$

which by (3.9) implies

$$q_n(t, x) \leq \epsilon'_k\bar{v}^{-2}. \tag{3.11}$$

Note that in the present setting we have

$$\sigma^2_{\mathrm{loc}}(t, x) = \sum_{n=1}^{N} u_n g_n(t)^2 q_n(t, x) \qquad \text{and} \qquad \overline{\sigma}(t)^2 = \frac{1}{\Lambda(A^k)}\sum_{n\in N^k} u_n g_n(t)^2,$$

from (3.6). Now (3.10) and (3.11) imply

$$|\sigma^2_{\mathrm{loc}}(t, x) - \overline{\sigma}(t)^2| \leq \epsilon'_k, \quad \text{for all } x \geq z_k := \max\{x_k, y_k\} \text{ and } t \in J_k, \tag{3.12}$$

which shows the claimed uniform convergence.

As in the proof of Lemma 2, we consider the log-price process $L_t = \ln(S_t^{\mathrm{loc}})$, and we have $c_k := \sup_{t\in I_k, x\in\mathbb{R}} \mathbb{E}[(L_t - L_T)^2 | L_T = x] < +\infty, k \in \mathbb{N}$, since σ is bounded. Setting $\Delta_k := \sqrt{c_k/\varepsilon'_k}$, we again obtain

$$\mathbb{P}(|L_t - L_T| \geq \Delta_k | L_T = x) \leq \varepsilon'_k, \quad \text{for all } t \in I_k \text{ and } x \in \mathbb{R}. \tag{3.13}$$

We are going to show that

$$\int_{\widetilde{I}_k} \left(\overline{\sigma}(t)^2 - \mathbb{E}[\sigma^2_{\mathrm{loc}}(t, S_t^{\mathrm{loc}}) | S_T^{\mathrm{loc}} = e^x]\right) dt = \int_{\widetilde{I}_k} \mathbb{E}[\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, e^{L_t}) | L_T = x]\, dt$$

converges to zero for $x \to \infty$, by proving that for x big enough this is smaller than the arbitrarily chosen ε_k. This in turn implies (3.7), being $k \in \mathbb{N}$ arbitrary, and concludes the proof of (3.5). In order to do that, we divide $\widetilde{I}_k$ in three subintervals :

$$\widetilde{J}_k := [t_1, T+\tau] \cap J_k, \quad \widetilde{J}'_k := [t_1, T+\tau] \cap J'_k, \quad \widetilde{J}''_k := [t_1, T+\tau] \cap J''_k, \quad k \in \mathbb{N},$$

where we used the notation introduced in (3.8). Note that, since σ takes values in $[\underline{v}, \overline{v}]$,

$$\int_{\widetilde{J}'_k} \mathbb{E}[\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, e^{L_t}) | L_T = x]\; dt \leq (\overline{v}^2 - \underline{v}^2)\varepsilon'_k,$$

and the same bound holds when taking the integral over $\widetilde{J}''_k$. On the other hand, (3.13) implies

$$\int_{\widetilde{J}_k} \mathbb{E}[(\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, e^{L_t}))\mathbf{1}_{L_t \leq L_T - \Delta_k} | L_T = x]\; dt \leq (\overline{v}^2 - \underline{v}^2)\varepsilon'_k,$$

and (3.12) implies

$$\int_{\widetilde{J}_k} \mathbb{E}[(\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, e^{L_t}))\mathbf{1}_{L_t > L_T - \Delta_k} | L_T = x]\; dt \leq \varepsilon'_k |\widetilde{J}_k|,$$

for all $x \geq \ln(z_k) + \Delta_k$. Altogether, for $x \geq \ln(z_k) + \Delta_k$ we have

$$\int_{\widetilde{I}_k} \mathbb{E}[\overline{\sigma}(t)^2 - \sigma^2_{\mathrm{loc}}(t, e^{L_t}) | L_T = x]\; dt \leq (b_k - a_k + 3(\overline{v}^2 - \underline{v}^2))\varepsilon'_k \leq \varepsilon_k.$$

This concludes the proof. □

4. Term-structure of convex ordering

In this section, we extend the model built in Section 3 in order to have the convex ordering preserved for long maturities, as market data suggests. To this end, we set

$$\frac{dS_t}{S_t} = \sigma_0 \frac{Y}{\sqrt{\mathbb{E}[Y^2|S_t]}}\, dW_t, \qquad t \geq t_2, \tag{4.1}$$

where $\sigma_0 \in \mathbb{R}_+$, and Y is a Bernoulli random variable known in t_2, and independent of anything else. Say Y takes the value y_- with probability q_- and y_+ with probability $q_+ = 1 - q_-$, for some $0 < y_- < y_+$ and $0 < q_- < 1$. By Jourdain and Zhou [12], the stochastic differential equation (SDE) (4.1) admits a weak solution $(\Omega, (\mathcal{F}_t), \mathbb{P}, W, (S_t), Y)$, which may not be unique. In the following we use the subscript or superscript $\mathbb{P}$ to emphasize that a priori the corresponding quantities depend on the weak solution of (4.1).

Note that (4.1) implies that, whatever the weak solution, $\sigma^2_{\mathrm{loc},\mathbb{P}}(t, S_t) = \sigma_0^2$ for $t \geq t_2$. Therefore $\mathrm{VIX}^2_{\mathrm{loc},T}$ does not depend on the weak solution and is constant equal to σ_0^2 for all maturities $T \geq t_2$. We now want to show that, on the other hand, for any weak solution of (4.1), this is not true for $\mathrm{VIX}^2_{\mathbb{P},T}$. This will imply that $\mathrm{VIX}^2_{\mathrm{loc},T}$ is strictly smaller than $\mathrm{VIX}^2_{\mathbb{P},T}$ in convex order for $T \geq t_2$, thus there is no inversion of convex ordering for long maturities.

For any weak solution of (4.1), we set

$$F_{\mathbb{P}}(s, t, x) := \mathbb{E}^{\mathbb{P}}[Y^2 | S_{t_2} = s, S_t = x].$$

Since Y is independent of W, the conditional law of $(S_t)_{t\geq t_2}$ given $Y = y_\pm$ and $S_{t_2} = s$ under $\mathbb{P}$ agrees with the (unique) law of a weak solution to the SDE

$$\frac{dS_t^{s,\mathbb{P},\pm}}{S_t^{s,\mathbb{P},\pm}} = \sigma_0 \frac{y_\pm}{\sqrt{F_\mathbb{P}\left(s,t,S_t^{s,\mathbb{P},\pm}\right)}}\, d\widetilde{W}_t, \qquad t \geq t_2, \qquad S_{t_2}^{s,\mathbb{P},\pm} = s,$$

living possibly on a different probability space (the weak uniqueness of the solution follows from [14, Theorem 3], given that [3, Proposition 5.1] ensures the existence of a measurable version of $F_\mathbb{P}$). Being $F_\mathbb{P}$ bounded and bounded away from zero, we deduce that, for $t > t_2$, the conditional law of S_t given $Y = y_\pm$ and $S_{t_2} = s$ under $\mathbb{P}$ admits a density $p_\pm^\mathbb{P}(s,t,x)$, and that $p_\pm^\mathbb{P}(s,t,x) > 0$ for all $x \in \mathbb{R}_+$, which in turn implies that

$$F_\mathbb{P}(s,t,x) = \frac{q_- y_-^2 p_-^\mathbb{P}(s,t,x) + q_+ y_+^2 p_+^\mathbb{P}(s,t,x)}{q_- p_-^\mathbb{P}(s,t,x) + q_+ p_+^\mathbb{P}(s,t,x)} \in (y_-^2, y_+^2), \qquad t > t_2. \tag{4.2}$$

Then, for $T \geq t_2$, we have

$$\mathrm{VIX}^2_{\mathbb{P},T} = \sigma_0^2 Y^2 \frac{1}{\tau} \int_T^{T+\tau} \mathbb{E}^\mathbb{P}\left[\frac{1}{F_\mathbb{P}(S_{t_2},t,S_t)} \middle| \mathcal{F}_T \right] dt =: \sigma_0^2 Y^2 \Psi_\mathbb{P}.$$

Now, having $\mathrm{VIX}^2_{\mathbb{P},T}$ constant (thus necessarily equal to σ_0^2) corresponds to having $Y^2\Psi_\mathbb{P} \equiv 1$, which is not possible given that $\Psi_\mathbb{P}$ takes values in $\left(\frac{1}{y_+^2}, \frac{1}{y_-^2}\right)$, by (4.2). This shows that $\mathrm{VIX}^2_{\mathbb{P},T}$ cannot be constant for any $T \geq t_2$.

Remark 6. To the best of our knowledge, uniqueness of a weak solution of (4.1) is still an open question. More generally, partial results on the existence of a weak solution of a calibrated stochastic local volatility (SLV) model of the form

$$\frac{dS_t}{S_t} = \sigma_{\mathrm{Dup}}(t,S_t) \frac{f(Y_t)}{\sqrt{\mathbb{E}[f(Y_t)^2|S_t]}}\, dW_t \tag{4.3}$$

have been obtained in [1, 12], but uniqueness has not been addressed. Note that Lacker *et al.* [13] have recently proved the weak existence and uniqueness of a *stationary* solution of a similar nonlinear SDE with drift, under some conditions. However, their result does not apply to the calibration of SLV models. Indeed, market-implied risk neutral distributions $(\mathcal{L}(S_t))_{t\geq 0}$ are strictly increasing in convex order and therefore no stationary solution $(S_t, Y_t)_{t\geq 0}$ can be a calibrated SLV model.

The possible absence of uniqueness of a weak solution of (4.1) or (4.3) is problematic, not only theoretically but also practically. It means that the price of a derivative in the calibrated SLV model may not be well defined. For example, in our case, the VIX may depend on $\mathbb{P}$. More generally, existence and uniqueness of (4.3) for general processes $(Y_t)_{t\geq 0}$ such as Itô processes remain a very challenging, open problem, despite the fact that these models are widely used in the financial industry, in particular thanks to the particle method of Guyon and Henry-Labordère [7].

Acknowledgements. We would like to thank Bruno Dupire, Vlad Bally, and the two anonymous referees for interesting discussions and helpful comments.

References

[1] Abergel, F., Tachet, R. : *A nonlinear partial integrodifferential equation from mathematical finance*, Discrete Cont. Dynamical Systems, Serie A, 27(3):907–917, 2010.
[2] Beiglböck, M., Friz, P., Sturm, S.: *Is the minimum value of an option on variance generated by local volatility?*, SIAM J. Finan. Math. 2:213–220, 2011.
[3] Brunick, G., Shreve, S.: *Mimicking an Itô process by a solution of a stochastic differential equation*, The Annals of Applied Probability, 23(4), 1584-1628, 2013.
[4] De Marco, S., Henry-Labordère, P.: *Linking vanillas and VIX options: A constrained martingale optimal transport problem*, SIAM J. Finan. Math., 6:1171–1194, 2015.
[5] Dupire, B.: *Pricing with a smile*, Risk, January, 1994.
[6] Dupire, B.: *Exploring Volatility Derivatives: New Advances in Modelling*, presentation at Global Derivatives, Paris, 2005.
[7] Guyon, J., Henry-Labordère, P., *Being Particular About Calibration*, Risk, January, 2012. Long version *The smile calibration problem solved*, preprint available at `ssrn.com/abstract=1885032`, 2011.

[8] Guyon, J: *On the Joint Calibration of SPX and VIX Options*, presentation at Jim Gatheral's 60th Birthday Conference, NYU, October 14, 2017.

[9] Guyon, J: *On the Joint Calibration of SPX and VIX Options*, presentation at the Finance and Stochastics seminar, Imperial College, London, March 28, 2018.

[10] Guyon J.: *Inversion of Convex Ordering in the VIX Market*, preprint available at `ssrn.com/abstract=3504022`, 2019.

[11] Gyöngy, I.: *Mimicking the One-Dimensional Marginal Distributions of Processes Having an Itô Differential*, Probability Theory and Related Fields, 71, 501-516, 1986.

[12] Jourdain, B., Zhou, A.: *Existence of a calibrated regime switching local volatility model*, Mathematical Finance, 2019.

[13] Lacker, D., Shkolnikov, M., Zhang, J.: *Inverting the Markovian projection, with an application to local stochastic volatility models*, preprint available at `arxiv.org/abs/1905.06213`, 2019.

[14] Veretennikov, A Yu: *On the strong solutions of stochastic differential equations*, Theory of Probability & Its Applications, 24(2), 354-366, 1980.