EXTREME-STRIKE COMPARISONS AND STRUCTURAL BOUNDS FOR SPX AND VIX OPTIONS

A. PAPANICOLAOU*

Abstract. This article explores relationships between the SPX and VIX options markets. High-strike VIX call options are used to hedge tail risk in the SPX, which means that SPX options are a reflection of the extreme-strike asymptotics of VIX options, and vice versa. This relationship can be quantified using moment formulas in a model-free way. Comparison formulas are presented along with various examples of stochastic volatility models.

1. Introduction. The S&P500 (SPX) index and it's volatility have been shown to have strong negative correlation. For this reason there is a great deal of interest in the VIX, the options based volatility index and its derivatives for hedging tail risk. The Chicago Board Options Exchange (CBOE) has, in particular, designed an index, the VIX Tail Hedge (VXTH), based on and SPX and VIX trading strategy, which has historically (backdated on data) performed better than the SPX when there has been a crisis event. Part of the VXTH strategy is to buy high-strike European call options on VIX to insure against losses in SPX, as the subsequent rise in the VIX index results in VIX call options posting positive returns. Prior to VIX derivatives, a similar insurance strategy might have been to buy low-strike European put options on SPX. This similarity means that there is information on the risk-neutral distribution for VIX that is implied by low-strike put options. With so much liquidity in both SPX and VIX options markets at present, it is useful to have structural bounds that quantify the relationship between these two markets. In particular, Lemma 3.2 in this paper will show that for SPX price S_t being a super-martingale, the moment generating function (MGF) of VIX_T^2 with $\xi \in \mathbb{R}$ satisfies

$$\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} \le \frac{1}{q}\mathbb{E}_t S_{T+\tau}^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t S_T^{\frac{2\xi p}{\tau}} \qquad \text{for times } 0 \le t \le T, \tag{1.1}$$

where $\mathbb{E}_t$ denotes risk-neutral expectation conditional on the market at time t, $\tau = 30$ days, and both $p \ge 1$ and $q \ge 1$ are Hölder-conjugate exponents with $\frac{1}{p} + \frac{1}{q} = 1$. If there exists $\xi > 0$ for which the right-hand side of (1.1) is finite, then the SPX market is saying that the VIX's distribution is not heavy tailed. Conversely, if the VIX's distribution is heavy tailed, then the right-hand side of (1.1) is infinite for all $\xi > 0$ and there do not exists any negative moments for $S_{T+\tau}$.

Stochastic volatility and Lévy processes for jumps (or a combination of these two) have been the state-of-the-art in option pricing since the 1990's, and at that time it may have seemed as if volatility derivatives could be priced and hedged from a well-calibrated model. However, data from various days throughout the 2000's exhibit VIX options trading at prices with implied-volatility smiles that reject standards such as the Heston model. From the perspective of someone searching for the "right model", the availability of VIX-options data is useful because it gives new information in addition to the data from SPX options. Nonetheless, model selection remains an important question because it is still a nontrivial task to fit a single model both to the SPX and VIX implied-volatility surfaces. Hence, it would be quite useful if there were a general theory to explain the bearing that one option surface has on the other. This paper presents the beginnings of such a theory in a model-free context.

Of further interest is the understanding of the implied volatility from VIX options. It is certainly true that every asset class requires tailored expert analysis of implied volatility, but VIX option

*School of Mathematics & Statistics, The University of Sydney, NSW 2006, Australia *alpapani@maths.usyd.edu.au*.

implied volatility is special because it is really the implied *volatility-of-volatility* for the SPX, and hence it is saying something about SPX options. In particular, implied volatilities from VIX options are sometimes very high (i.e. in the range of 80% for high-strike VIX call options), but there is not yet a standard for making comparisons to implied volatilities observed from SPX options. It would be a significant contribution if implied volatiles from VIX call options could be used to make definitive statements about the no-arbitrage range for SPX implied volatilities. Using the moment formula of [Lee04], this paper identifies such relationships using extreme-strike options.

1.1. Literature Review. The VIX formula (as it has been calculated since 2003) is described in [DDKZ99]. A general description of how volatility derivatives are designed and traded, particularly in the post-2008 crisis markets, is provided in [CL09]. Stochastic and local volatility models are described in [Gat06], and pricing of volatility derivatives based on these models is a straightforward application of standard, partial differential equations methods. An alternative to model-based pricing/hedging are the model-free results found in [CL08, CL10, FG05].

The issues in fitting the Heston model to VIX options are explained in [Gat08]. There has been some success in fitting VIX options to market models of the variance swap term structure (see [CS07, CK13]), and the added explanatory power from the inclusion of jumps has been demonstrated in [MY11]. Some studies have shown that the large-strike implied-volatility skew from VIX options can be fit with a heavy-tailed volatility process (see [BB14, Dri12]), but heavy tails are not necessarily required as shown in [BGK13] using a double-Heston model and in [PS14] using a Markov-chain modulation of the Heston model.

1.2. Main Results Of This Paper. The main result in this paper is the identification of basic connections between VIX options and negative moments in the SPX price. In particular, the existence of negative moments in the SPX's risk-neutral distribution is an indication that the VIX's risk-neutral distribution is not heavy tailed. Conversely, if the risk-neutral distribution of VIX_T^2 is heavy tailed (i.e. it's MGF does not exist on the positive real line), then the price for SPX has no negative moments. The results can be considered model free, as the main assumptions are continuity of the stochastic processes and the martingale property. Part of the main results of this paper is the detailed application of the theory to specific models that are frequently used in SPX and VIX options pricing and an assessment of their relative usefulness, given the observed market behavior of these options.

The rest of the paper is organized as follows: Section 2 introduces the probabilistic framework and describes how to price options on SPX and VIX; Section 3 presents the main results and other ideas relevant to the problem; Section 4 presents various stochastic volatility models and discusses how each relates to this paper's results. Section 5 concludes and Appendices B has some necessary calculations for the examples from Section 4.

2. Probabilistic Framework for Pricing. Let S_t denote the price of the SPX at some time $t \geq 0$. The model considered throughout this paper has an asset whose log returns are given by a stochastic volatility model,

$$d \log S_t = \left(r - \frac{1}{2}\sigma_t^2 \right) dt + \sigma_t dW_t \tag{2.1}$$

where $r \geq 0$ is the risk-free rate of interest, W is a risk-neutral Brownian motion, and σ_t is a volatility process that is non-anticipative of W. This paper will assume that σ_t satisfies strong enough conditions for $\int_0^t \sigma_u dW_u$ to be a true martingale over a finite time horizon.

Condition 2.1 (Finite Second Moment of Stochastic Integral). *For $T < \infty$, the second moment of stochastic integral $\int_0^T \sigma_u dW_u$ is finite,*

$$\mathbb{E}\left(\int_0^T \sigma_u dW_u\right)^2 = \mathbb{E}\int_0^T \sigma_u^2 du < \infty \ .$$

Condition 2.1 implies that $\int_0^t \sigma_u dW_u$ is a true, square integrable martingale on finite time-interval $[0,T]$, but S_t may still be a strict local martingale. A much stronger condition is the *Novikov Condition*, which states that for $T < \infty$ the process S_t is a true martingale if $\mathbb{E}e^{\frac{1}{2}\int_0^T \sigma_u^2 du} < \infty$. The Novikov condition is very strong and often doesn't hold for stochastic volatility models. Other conditions for exponential martingales are presented in [KL12]. This paper will rely on Condition 2.1, will not assume Novikov, and will show the martingale property on a case-by-case basis.

Another important condition is the existence of S_T's negative moments:

Condition 2.2 (Negative Moments). *For $T < \infty$, there exist constant $q > 0$ such that*

$$\mathbb{E}S_T^{-q} < \infty \ .$$

For S_t a super martingale, Condition 2.2 implies existence of the MGF of $\log(S_T)$ in a neighborhood containing zero,

$$\mathbb{E}e^{\xi \log(S_T)} < \infty \qquad \forall \xi \in [-q, 1] \ .$$

Condition 2.2 is used in [Lee04] to obtain small-strike bounds on implied volatility. In particular, the supremum over all $q > 0$ such that Condition 2.2 holds is identified with the asymptotic rate at which implied volatility grows as strike-price goes to zero in the SPX options market; this is part 2 of the moment formula [Lee04] that will be reviewed in Section 3.1.

2.1. Variance Swaps and the VIX Index. Consider European call and put options on S_T for some fixed time $T \in (0,\infty)$ and some fixed strike $K \in [0,\infty)$, both of which are processes

$$\mathrm{C}(t,K,T) \triangleq B_{t,T}\mathbb{E}_t(S_T - K)^+$$
$$\mathrm{P}(t,K,T) \triangleq B_{t,T}\mathbb{E}_t(K - S_T)^+$$

for some $0 \leq t \leq T$, where $B_{t,T} \triangleq e^{-r(T-t)}$ and the expectation operator is defined as $\mathbb{E}_t \triangleq \mathbb{E}\{\ \cdot\ |\mathcal{F}_t\}$ with $\mathcal{F}_t$ denoting filtration generated by $(W_u, \sigma_u)_{u \leq t}$. Throughout the paper, an expectation without a subscript is conditional at time $t = 0$, that is, $\mathbb{E} = \mathbb{E}_0$.

Definition 2.3 (SPX Implied Volatility $\hat{\sigma}$). *Implied volatility for SPX options is denoted with $\hat{\sigma}(t,K,T)$ and is the unique volatility input to the Black-Scholes prices such that*

$$P(t,K,T) = B_{t,T}\left(\Phi\left(-d_-\right)K - \Phi\left(-d_+\right)\mathbb{E}_t S_T\right)$$

where $d_\pm = \frac{\log(\mathbb{E}_t S_T/K)}{\hat{\sigma}(t,K,T)\sqrt{T-t}} \pm \frac{\hat{\sigma}(t,K,T)\sqrt{T-t}}{2}$ and Φ is the standard normal cumulative distribution function. The quantity $\hat{\sigma}$ could be equivalently redefined using call options via the put-call parity.

A variance swap for the time period $[t,T]$ with $t < T$ has a floating leg of $\frac{1}{T-t}\int_t^T \sigma_u^2 du$ (equal to the quadratic variation of $\log S_t$) and a fixed leg that is chosen such that the contract has zero entry cost at time t. This fixed leg is the **variance-swap rate:**

$$\text{variance-swap rate} = \mathbb{E}_t\left\{\frac{1}{T-t}\int_t^T \sigma_u^2 du\right\} \ .$$

When trading in variance swaps, an important instrument is the log contract with time-T payout of $-\frac{2}{T-t}\log(S_T/\mathbb{E}_tS_T)$. As shown in [DDKZ99], the log contract is replicated by a portfolio of European call and put options by taking the expectation of the identity $-\log(S_T/s^*) = \frac{S_T-s^*}{s^*} + \int_0^{s^*}\frac{(K-S_T)^+}{K^2}dK + \int_{s^*}^\infty\frac{(S_T-K)^+}{K^2}dK$ (which holds for any reference point $s^*>0$), and taking $s^* = \mathbb{E}_tS_{t+\tau}$ yields the VIX formula:

$$\mathrm{VIX}_t = \sqrt{\frac{2}{\tau B_{t,t+\tau}}\left(\int_0^{\mathbb{E}_tS_{t+\tau}}\mathrm{P}(t,K,t+\tau)\frac{dK}{K^2} + \int_{\mathbb{E}_tS_{t+\tau}}^\infty \mathrm{C}(t,K,t+\tau)\frac{dK}{K^2}\right)}\ , \tag{2.2}$$

where $\tau = 30$ days. By definition, equation (2.2) is the square root of the log contract's price $\mathrm{VIX}_t = \sqrt{-\frac{2}{\tau}\mathbb{E}_t\log\left(S_{t+\tau}\Big/\mathbb{E}_tS_{t+\tau}\right)}$. By assuming Condition 2.1 for the continuous model in (2.1), the no-arbitrage price of the log-contract is equal to the variance-swap rate, and hence the VIX index is the square-root of the variance-swap rate for the coming 30 days,

$$(\text{Condition 2.1}) \Rightarrow \mathrm{VIX}_t = \sqrt{\text{variance-swap rate}}\ .$$

2.2. VIX Future and VIX Options. Define the future contract on VIX_T at time $t\leq T$ as

$$X_{t,T} \triangleq \mathbb{E}_t\mathrm{VIX}_T = \mathbb{E}_t\sqrt{-\frac{2}{\tau}\mathbb{E}_T\log\left(S_{T+\tau}\Big/\mathbb{E}_TS_{T+\tau}\right)}\ . \tag{2.3}$$

The price $X_{t,T}$ is the central quantity in the VIX market because (unlike the VIX index) it is a trade-able asset. European call and put options on the VIX are the expectation of functions of VIX_T, but should be thought of as options on $X_{T,T}$,

$$\mathrm{C}^{vix}(t,K,T) \triangleq B_{t,T}\mathbb{E}_t(\mathrm{VIX}_T-K)^+ = B_{t,T}\mathbb{E}_t(X_{T,T}-K)^+$$
$$\mathrm{P}^{vix}(T,K,T) \triangleq B_{t,T}\mathbb{E}_t(K-\mathrm{VIX}_T)^+ = B_{t,T}\mathbb{E}_t(K-X_{T,T})^+\ .$$

Considering these options as payoffs on $X_{T,T}$ will make clearer the standards for Δ-hedging in terms of $X_{t,T}$, and also the standard for computing implied volatility. As is the convention for options-trading, VIX options are quoted using implied volatility that is computed by inverting Black-Scholes formula for options on futures, as is also done in [PS14].

Definition 2.4 (VIX Implied Volatility $\hat\nu$). *Implied volatility for VIX options is denoted with $\hat\nu(t,K,T)$ and is the unique volatility input to the Black-Scholes prices such that*

$$C^{vix}(t,K,T) = B_{t,T}\left(X_{t,T}\Phi\left(d_+\right) - K\Phi\left(d_-\right)\right)\ ,$$

where $d_\pm = \frac{\log(X_{t,T}/K)}{\hat\nu(t,K,T)\sqrt{T-t}} \pm \frac{\hat\nu(t,K,T)\sqrt{T-t}}{2}$ and Φ is the standard normal cumulative distribution function. The quantity $\hat\nu$ could be equivalently redefined using VIX put options via put-call parity.

Figures 2.1 and 2.2 show implied volatility for SPX and VIX options for September 9th of 2010, a day that was shortly after the European debt crisis when options were trading with high implied volatility. Notice the right-hand skew of the $\hat\nu$ in Figure 2.2, which corresponds to volatility tail risk and is stylistic feature of VIX options that should be captured by a stochastic volatility model that aims to price VIX options in periods of higher volatility (see [Dri12, BGK13, PS14]).

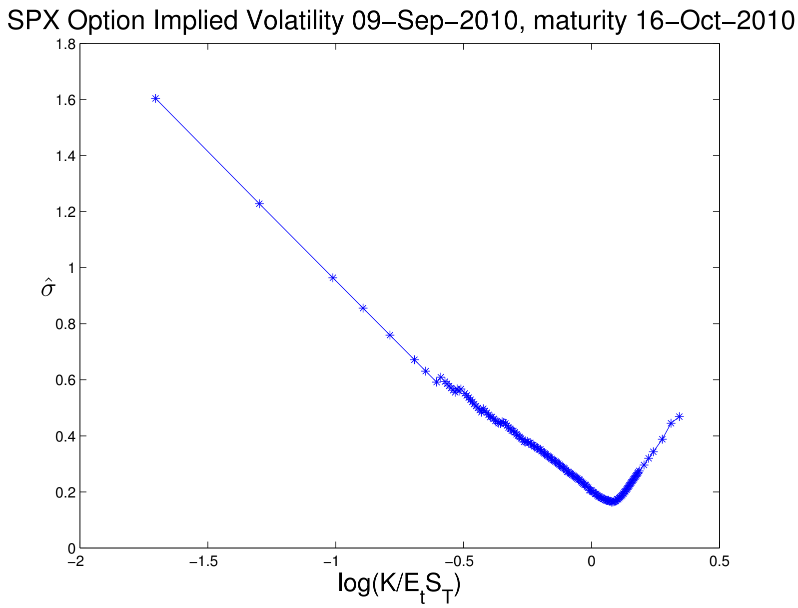

Fig. 2.1: Implied volatility for SPX put options in September 2011. The ex-dividend SPX future price is $\mathbb{E}_t S_T = \$1101.97$ and the risk-free rate is approximately $r = .28\%$. The left-hand skew observed in this figure has been a common sight since the late 1980's. Stochastic volatility models (such as the Heston model) and various Lévy models have successfully fit this skew. However, VIX options introduce another skew that is a derivative of the SPX skew (see Figure 2.2), and this new skew has forced the re-evaluation of standards in stochastic volatility models.

3. Extreme-Strike Asymptotics. Results that are considered model-free usually require some assumptions, such as the VIX being finite almost surely, a condition that is ensured by Condition 2.1. Indeed, Condition 2.1 is the key assumption for the square of the VIX formula in (2.2) to be equal to the variance-swap rate, but it turns out that the negative moments described in Condition 2.2 are sufficient to imply Condition 2.1. The following proposition proves this statement and is the first instance in this paper where a connection is identified between the VIX and negative moments in SPX:

PROPOSITION 3.1. *Suppose $S_t e^{-rt}$ is a martingale and suppose Condition 2.2. Then Condition 2.1 holds.*

Proof. From the martingale property the call option price is bounded $\mathrm{C}(t,K,T) \leq B_{t,T}\mathbb{E}_t S_T = S_t$, and from [Lee04] there is the put-option estimate $\mathrm{P}(t,K,T) \leq B_{t,T}\frac{\mathbb{E}_t S_T^{-q}}{q+1}\left(\frac{q}{q+1}\right)^q K^{1+q}$ for all

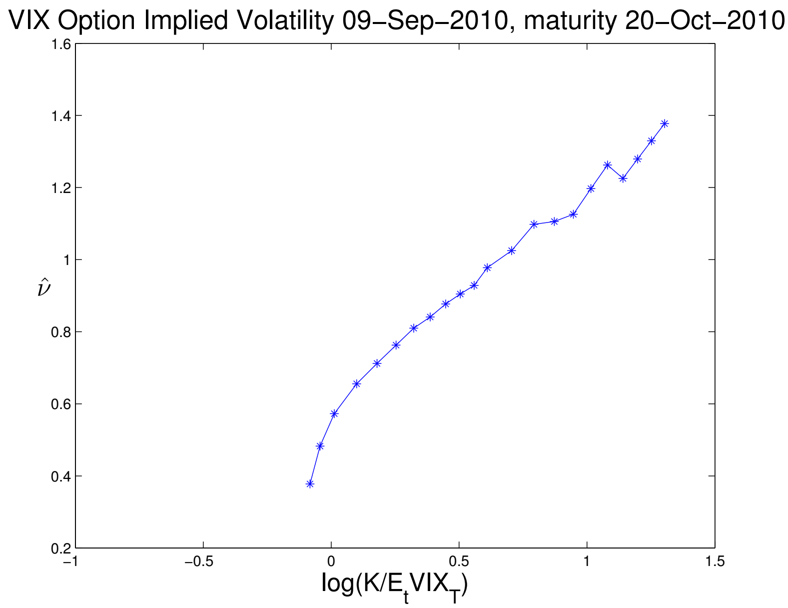

Fig. 2.2: Implied volatility for VIX call options on April 7th, 2011, maturity on June 15th, 2011. The VIX future price $\mathbb{E}_t\mathrm{VIX}_T = 27.1714\%$. Data from this day highlights to commonly-observed right skew from VIX options. This is the stylistic feature that is identified in the literature (see [Dri12, Gat08, PS14]). The right skew affects selection of stochastic volatility models, by suggesting that heavy-tailed volatility models (such as the 3/2 or GARCH models) are better-suited than the light-tailed Heston model.

$K \in [0,\infty)$ and $q > 0$. Then using the square of equation (2.2) and taking $q \in (0,1)$ yields

$$\begin{aligned}\mathbb{E}_t\int_t^T \sigma_u^2 du &\leq \frac{2}{B_{t,T}}\left\{\frac{\mathbb{E}_t S_T^{-q}}{q+1}\left(\frac{q}{q+1}\right)^q\int_0^{\mathbb{E}_t S_T}\frac{K^{1+q}}{K^2}dK + S_t\int_{\mathbb{E}_t S_T}^\infty\frac{1}{K^2}dK\right\}\\ &= \frac{2}{B_{t,T}}\left\{\frac{\mathbb{E}_t S_T^{-q}}{q(q+1)}\left(\frac{q}{q+1}\right)^q(\mathbb{E}_t S_T)^q + \frac{S_t}{\mathbb{E}_t S_T}\right\}\\ &= \frac{2}{B_{t,T}}\left\{\frac{\mathbb{E}_t S_T^{-q}}{q(q+1)}\left(\frac{q}{q+1}\right)^q(S_t)^q + 1\right\} \qquad \forall t \leq T\ .\end{aligned}$$

Hence, $\mathbb{E}\int_0^T \sigma_u^2 du \leq \frac{2}{B_{t,T}}\left\{\frac{\mathbb{E}S_T^{-q}}{q(q+1)}\left(\frac{q}{q+1}\right)^q S_0^q + 1\right\} < \infty$ because $\mathbb{E}S_T^{-q} < \infty$, and Condition 2.1 holds. □

Another instance where finite SPX moments are important is in determining the existence of the MGF of VIX_T^2. The following lemma will be instrumental throughout the rest of the paper:

LEMMA 3.2. *Let S_te^{-rt} be a super martingale on $[0, T+\tau]$ satisfying Condition 2.1. For any*

$\xi \in \mathbb{R}$, the MGF $\mathbb{E}_t e^{\xi\, VIX_T^2}$ satisfies the strict inequality

$$\mathbb{E}_t e^{\xi\, VIX_T^2} < \frac{1}{q}\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t \left(S_T/B_{T,T+\tau}\right)^{\frac{2\xi p}{\tau}} \qquad \forall t \le T \ , \tag{3.1}$$

where $p > 1$ and $q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ (i.e. p and q are conjugate exponents).

Proof. It suffices to prove for any $\xi \ge 0$ because $\mathbb{E}_t e^{-\xi \mathrm{VIX}_T^2} < \mathbb{E}_t e^{\xi \mathrm{VIX}_T^2}$. From Jensen's and Young's inequality,

$$\begin{aligned}
\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} &= \mathbb{E}_t e^{-\frac{2\xi}{\tau}\mathbb{E}_T \log(S_{T+\tau}/\mathbb{E}_T S_{T+\tau})} \\
&= \mathbb{E}_t e^{\mathbb{E}_T \log\left(\left(\frac{S_{T+\tau}}{\mathbb{E}_T S_{T+\tau}}\right)^{-\frac{2\xi}{\tau}}\right)} \\
&< \mathbb{E}_t e^{\log\left(\left(\frac{S_{T+\tau}}{\mathbb{E}_T S_{T+\tau}}\right)^{-\frac{2\xi}{\tau}}\right)} && \text{(Jensen's inequality)} \\
&= \mathbb{E}_t \left(\frac{S_{T+\tau}}{\mathbb{E}_T S_{T+\tau}}\right)^{-\frac{2\xi}{\tau}} \\
&< \frac{1}{q}\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t \left(\frac{1}{\mathbb{E}_T S_{T+\tau}}\right)^{-\frac{2\xi p}{\tau}} && \text{(Young's inequality)} \\
&= \frac{1}{q}\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t \left(\mathbb{E}_T S_{T+\tau}\right)^{\frac{2\xi p}{\tau}} \\
&< \frac{1}{q}\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t \left(S_T/B_{T,T+\tau}\right)^{\frac{2\xi p}{\tau}} \ .
\end{aligned}$$

Jensen's inequality is an equality iff the random variable has zero variance, hence the inequality must be strict. □

Lemma 3.2 is a useful tool when evaluating the market for VIX options, primarily because it shows how existence of a negative moment $\mathbb{E}_t S_{T+\tau}^{-q}$ for some $q > 0$ implies that the VIX-square process is not heavy tailed. Conversely, if the MGF $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$, then $\mathbb{E}_t S_{T+\tau}^{-q} = \infty$ for all $q > 0$. In both cases, for $\frac{2\xi p}{\tau} \le 1$ the (super) martingale property ensures $\mathbb{E}_t \left(S_T/B_{T,T+\tau}\right)^{\frac{2\xi p}{\tau}} < \infty$, so that $\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} < \infty$ implies $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} < \infty$, and $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} = \infty$ implies $\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} = \infty$.

3.1. Moment Formulas. The Moment Formula from [Lee04] consists of parts 1 and part 2 describing the right and left tail, respectively, of the implied volatility smile. Let the price process S_t be a martingale and define $\tilde{p} \triangleq \sup\{p \ge 0 | \mathbb{E}_t S_T^{1+p} < \infty\}$. Part 1 states that

$$\limsup_{K \nearrow \infty} \frac{\hat{\sigma}^2(t,K,T)}{\log(K/\mathbb{E}_t S_T)/(T-t)} = \beta_R \in [0,2] \ , \tag{3.2}$$

where $\tilde{p} = \frac{1}{2\beta_R} + \frac{\beta_R}{8} - \frac{1}{2}$. Equivalently, $\beta_R = 2 - 4\left(\sqrt{\tilde{p}^2 + \tilde{p}} - \tilde{p}\right)$ with $\beta_R = 0$ when $\tilde{p} = \infty$. Next define $\tilde{q} \triangleq \sup\{q \ge 0 | \mathbb{E}_t S_T^{-q} < \infty\}$. Part 2 states that

$$\limsup_{K \searrow 0} \frac{\hat{\sigma}^2(t,K,T)}{\log(K/\mathbb{E}_t S_T)/(T-t)} = \beta_L \in [0,2] \tag{3.3}$$

where $\tilde{q} = \frac{1}{2\beta_L} + \frac{\beta_L}{8} - \frac{1}{2}$. Equivalently, $\beta_L = 2 - 4\left(\sqrt{\tilde{q}^2 + \tilde{q}} - \tilde{q}\right)$ with $\beta_L = 0$ when $\tilde{q} = \infty$. Parts 1 and 2 both take $1/0 \triangleq \infty$. In most "practical" cases the limit supremum in (3.2) and (3.3) can be replaced with a proper limit (see [BF08]). Figure 3.1 shows how the moment formulas apply to the data with β_L and β_R estimated from the most extreme strikes in the options data.

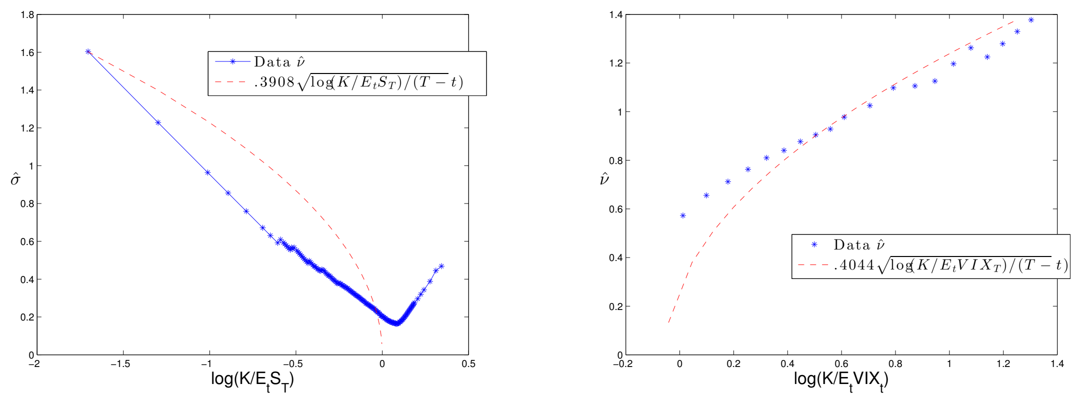

Fig. 3.1: **Left:** From the lowest-strike SPX put option there is the estimate $\beta_L \geq 0.3908$, which is the coefficient of the dashed line that is the extreme-strike asymptotic from the moment formula. **Right:** From the highest-strike VIX option there is the estimate $\beta_R \geq .4044$.

The moment formula can be used to show how moment explosion in the VIX options market affects implied volatility in SPX options. Define $\tilde{p}^{vix} = \sup\left\{p > 0 \middle| \mathbb{E}_t \mathrm{VIX}_T^{1+p} < \infty\right\}$. If $\tilde{p}^{vix} < \infty$ then the MGF $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$ and by Lemma 3.2 it follows that $\mathbb{E}_t S_{T+\tau}^{-q} = \infty$ for all $q > 0$. Hence by part 2 of the moment formula, as stated in equation (3.3), it follows that $\beta_L = 2$ for SPX options with exercise at time $T + \tau$, and the implied volatility limit is at it's maximum, $\limsup_{K \searrow 0} \frac{\hat{\sigma}^2(t, T+\tau, K)}{\log(K/\mathbb{E}_t S_{T+\tau-t})/(T+\tau-t)} = 2$.

Similarly, moments of SPX's distribution can say something about implied volatility of VIX options. Define $\beta_R^{vix} = \limsup_{K \nearrow \infty} \frac{\hat{\nu}^2(t,K,T)}{\log(K/\mathbb{E}_t S_T)/(T-t)}$, and suppose $\tilde{q} = \sup\left\{q > 0 \middle| \mathbb{E}_t S_{T+\tau}^{-q} < \infty\right\} > 0$. Then from Lemma 3.2 it follows that $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} < \infty$ for some $\xi > 0$. Moreover, if VIX_T^2 has finite MGF for positive ξ then $\mathbb{E}_t \mathrm{VIX}_T^n < \infty$ for all $n > 0$, and then by equation (3.2) it follows that $\beta_R^{vix} = 0$, giving the extreme-strike asymptotic $\limsup_{K \nearrow \infty} \frac{\hat{\nu}^2(t,T,K)}{\log(K/X_{t,T})/(T-t)} = 0$. Moreover, finite MGF for VIX_T^2 for positive ξ is equivalent to saying that the VIX-squared does not have a heavy-tailed distribution, and hence $\tilde{q} > 0$ implies VIX_T does not have heavy tails.

More generally, the moment formula and Lemma 3.2 are used to show how implied volatility from VIX options gives a lower bound on the implied volatilities of SPX options.

Proposition 3.3. *Assume Condition 2.1, let the price process $S_t e^{-rt}$ be a true martingale, and let*

$$\tilde{\xi} = \sup\left\{\xi \geq 0 \middle| \mathbb{E}_t e^{\xi \, VIX_T^2} < \infty\right\}.$$

1. *If* $\tilde{\xi} < \infty$ *and* $\mathbb{E}_t S_{T+\tau}^{\frac{4\tilde{\xi}}{\tau}} < \infty$*, then*

$$\tilde{q} = \sup\left\{ q \geq 0 \middle| \mathbb{E}_t S_{T+\tau}^{-q} < \infty \right\} \leq \frac{4\tilde{\xi}}{\tau},$$

and

$$\limsup_{K \searrow 0} \frac{\hat{\sigma}^2(t,K,T+\tau)}{\log(K/\mathbb{E}_t S_{T+\tau-t})/(T+\tau-t)} = \beta_L \geq 2 - 4\left(\sqrt{\left(\frac{4\tilde{\xi}}{\tau}\right)^2 + \frac{4\tilde{\xi}}{\tau}} - \frac{4\tilde{\xi}}{\tau}\right).$$

2. *If* $\tilde{\xi} < \infty$ *and* $\mathbb{E}_t S_{T+\tau}^{\frac{-4\tilde{\xi}}{\tau}} < \infty$*, then* $\frac{4\tilde{\xi}}{\tau} > 1$ *and*

$$\tilde{p} = \sup\left\{ p \geq 0 \middle| \mathbb{E}_t S_{T+\tau}^{1+p} < \infty \right\} \leq \frac{4\tilde{\xi}}{\tau} - 1,$$

and

$$\limsup_{K \nearrow \infty} \frac{\hat{\sigma}^2(t,K,T+\tau)}{\log(K/\mathbb{E}_t S_{T+\tau-t})/(T+\tau-t)} = \beta_R$$
$$\geq 2 - 4\left(\sqrt{\left(\frac{4\tilde{\xi}}{\tau} - 1\right)^2 + \frac{4\tilde{\xi}}{\tau} - 1} - \left(\frac{4\tilde{\xi}}{\tau} - 1\right)\right).$$

Proof. Part 1: Using Lemma 3.2 with $p = q = 2$, it is shown that $\tilde{q} = \sup\{q \geq 0 | \mathbb{E}_t S_{T+\tau}^{-q} < \infty\} \leq \frac{4\tilde{\xi}}{\tau}$. Now define the function

$$f(q) = 2 - 4\left(\sqrt{q^2 + q} - q\right)$$

and notice that $f'(q) < 0$ for all $q > 0$, so that $\beta_L = f(\tilde{q}) \geq f\left(\frac{4\tilde{\xi}}{\tau}\right)$.

Part 2: If $\tilde{\xi} < \infty$ and $\mathbb{E}_t S_{T+\tau}^{\frac{-4\tilde{\xi}}{\tau}} < \infty$, then $\mathbb{E}_t S_{T+\tau}^{\frac{4\tilde{\xi}}{\tau}} = \infty$ and $1 + \tilde{p} \leq \frac{4\tilde{\xi}}{\tau}$, and therefore $\tilde{p} \leq \frac{4\tilde{\xi}}{\tau} - 1$. The remainder of the proof is similar to the argument used in Part 1, expect with an application of equation (3.2). ∎

3.2. Hedging VIX Markets with SPX Options. From a market of VIX options with a continuum of strikes comes a tremendous amount of information about the SPX. In the particular, via the Breeden and Litzenberger formiula one obtains a risk-neutral distribution on the portfolio of calls and puts given in equation (2.2). This section will show that if the MGF of VIX_T^2 can be replicated, then violation of the inequality in (3.1) can result in a static arbitrage. Moreover, in replicating the MGF for VIX_T^2 it is seen that $\mathbb{E}_t e^{\xi \text{VIX}_T^2} < \infty$ for some $\xi > 0$ implies that VIX call options decay quickly as K grows.

Moments on VIX can be replicated with options, $B_{t,T}\mathbb{E}_t \text{VIX}_T^n = \int_0^\infty K^n \frac{\partial^2}{\partial K^2} \text{C}^{vix}(t,K,T)dK = n(n-1)\int_0^\infty K^{n-2}\text{C}^{vix}(t,K,T)dK$ for positive integers n. Furthermore, the MGF of VIX_T^2 can be hedged, and if finite will give a rate at which call options must decay for large strikes.

PROPOSITION 3.4. *Suppose Condition 2.1. For any $\xi \in \mathbb{R}$ there is a perfect hedge of the MGF of VIX_T^2 in terms of VIX options,*

$$B_{t,T}\mathbb{E}_t e^{\xi\, VIX_T^2} = 1 + \int_0^\infty \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} C^{vix}(t,K,T)dK \ .$$

Moreover, if the MGF $\mathbb{E}_t e^{\xi\, VIX_T^2} < \infty$ for some $\xi > 0$, then

$$\lim_{K\to\infty} K^2 e^{\xi K^2} C^{vix}(t,K,T) = 0 \qquad \text{for all } t \leq T \ . \tag{3.4}$$

Proof. It follows from Condition 2.1 that $\mathrm{VIX}_T < \infty$ almost surely, and then elementary calculus is used to check that

$$e^{\xi \mathrm{VIX}_T^2} = 1 + \int_0^\infty \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} \left(\mathrm{VIX}_T - K\right)^+ dK \ .$$

Then, for $t < T$ and any $N < \infty$, the monotone convergence theorem is used to obtain

$$\begin{aligned}
&\int_0^\infty \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} \mathrm{C}^{vix}(t,K,T)dK \\
&= B_{t,T} \lim_{N\to\infty} \mathbb{E}_t \int_0^N \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} \left(\mathrm{VIX}_T - K\right)^+ dK \\
&= B_{t,T}\mathbb{E}_t \lim_{N\to\infty} \int_0^N \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} \left(\mathrm{VIX}_T - K\right)^+ dK \\
&= B_{t,T}\left(\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} - 1\right) \ .
\end{aligned}$$

Moreover, $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} < \infty$ if and only if $\int_0^\infty \left(2\xi + 4\xi^2 K^2\right) e^{\xi K^2} \mathrm{C}^{vix}(t,K,T)dK < \infty$, in which case the integrand must converge zero as K tends toward infinity, hence $\lim_{K\to\infty} K^2 e^{\xi K^2} \mathrm{C}^{vix}(t,K,T) = 0$. ☐

The hedge in Proposition 3.4 is useful because it is part of an argument to show that reversal of the inequality in Lemma 3.2 can result in static arbitrage. Portfolios for hedging $\mathbb{E}_t S_T^p$ for $p > 1$ and $\mathbb{E}_t S_T^{-q}$ for $q > 0$ are given in [Lee04], and for $p = 1/2$ the hedge is obtained from the Laplace transform $\mathbb{E}_t\sqrt{S_T} = \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-\mathbb{E}_t e^{-\lambda S_T}}{\lambda^{3/2}} d\lambda$ (see [CL08]) with

$$\mathbb{E}_t e^{-\lambda S_T} = \frac{1}{B_{t,T}}\left(1 - \lambda \mathbb{E}_t S_T + \lambda^2 \int_0^\infty e^{-\lambda K} \mathrm{C}(t,K,T)dK\right) \ .$$

Assuming $S_t e^{-rt}$ is a super-martingale and assuming Condition 2.2 for $S_{T+\tau}$, if inequality in (3.1) is reversed so that $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} \geq \frac{1}{q}\mathbb{E}_t \left(S_{T+\tau}\right)^{-\frac{2\xi q}{\tau}} + \frac{1}{p}\mathbb{E}_t \left(S_T/B_{T,T+\tau}\right)^{\frac{2\xi p}{\tau}}$, then an arbitrage is obtained by opening a portfolio that is long in SPX and short in VIX options. Letting $\tilde{\xi} = \frac{2\xi}{\tau}$, the strategy is

Buy:

$$\Pi_t^1 = \frac{1}{q}\mathbb{E}_t S_{T+\tau}^{-\tilde\xi q} = \frac{\tilde\xi(\tilde\xi q+1)}{B_{t,T}}\int_0^\infty \frac{\mathrm{P}(t,K,T+\tau)}{K^{\tilde\xi q+2}}dK$$

$$\Pi_t^2 = \frac{1}{p}\mathbb{E}_t S_T^{\tilde\xi p} = \frac{1}{B_{t,T}}\begin{cases} \tilde\xi(\tilde\xi p-1)\int_0^\infty \mathrm{C}(t,K,T)\frac{dK}{K^{2-\tilde\xi p}} & \text{if } \tilde\xi p>1 \\ \frac{1}{p}\mathrm{C}(t,0,T) & \text{if } \tilde\xi p = 1 \\ \frac{1}{2p\sqrt{\pi}}\int_0^\infty\left(S_t-\lambda\int_0^\infty e^{-\lambda K}C(t,K,T)\right)\frac{d\lambda}{\sqrt{\lambda}}dK & \text{if } \tilde\xi p = 1/2 \end{cases}$$

Sell:

$$\Pi_t^3 = \mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} = \frac{1}{B_{t,T}}\int_0^\infty \left(2\xi+4\xi^2K^2\right)e^{\xi K^2}\mathrm{C}^{vix}(t,K,T)dK\ .$$

At time t there is a net cash flow of $\Pi_t^3-\Pi_t^1+\Pi_t^2\geq 0$, but at time T the position can be closed out for a strictly positive net cash flow, as $\Pi_T^1+\Pi_T^2-\Pi_T^3>0$ by Lemma 3.2. Hence, there is an arbitrage.

To summarize, this section has shown a static hedge the MGF for VIX_T^2, which means there is a implementable arbitrage portfolio if the inequality of (3.1) is violated. Moreover, this static hedge is informative because it shows exponential decay for large-K VIX call options if there is finiteness of $\mathbb{E}_t E^{\xi \mathrm{VIX}_T^2}$ for some $\xi>0$. However, this large-K asymptotic cannot be differentiated to get a tail distribution, but the next section will explore how to obtain the VIX's tail distribution.

3.3. Extreme-Value Theory for VIX's Distribution. The moment-formula limits β_R and β_L have a relationship to the rates in the extreme-value distributions of the underlying asset. In particular, the stochastic volatility inspired (SVI) parameterization of the volatility surface leads to an extreme-value distribution that is parameterized by β_R for the right tail and β_L for the left. The SVI parameterization is related to extreme strikes because its construction is consistent with the moment formulas in (3.2) and (3.3), namely that the square of large-strike implied volatility is proportional to log-moneyness divided by the square-root of time-to-maturity (see [Gat06, BGK13]).

Motivation for the section comes from the fact that the limiting behavior for VIX options given in Proposition 3.4 cannot be differentiated in K. In other words, the limit in (3.4) shows a rate of convergence for C^{vix}, but cannot be differentiated to find an asymptotic for the tail distribution $\frac{\partial}{\partial K}C^{vix}$. However, by assuming differentiability in K and assuming that the SVI is not mis-specified, the asymptotic tail distribution is obtained. In particular, the limiting right slope for VIX options $\beta_R^{vix}=\limsup_{K\nearrow\infty}\frac{\hat\nu^2(t,K,T)}{\log(K/\mathbb{E}_tS_T)/(T-t)}$ gives the rate of polynomial decay in the peaks-over-threshold (POT) distribution of VIX_T if $\beta_R^{vix}\in(0,2)$.

The generalized Pareto distribution (GPD) with parameter $\alpha>0$ has cumulative distribution function defined as

$$G_\alpha(y)\triangleq\begin{cases} 1-(1+y/\alpha)^{-\alpha} & \text{for } y\geq 0 \text{ and } \alpha<\infty \\ 1-e^{-y} & \text{for } y\geq 0 \text{ and } \alpha=\infty\ . \end{cases}$$

The Pickands-Balkema-de Haan theorem states that the following are equivalent:

(i) The VIX's distribution function is the in the maximum domain of attraction of H_α defined as $H_\alpha(y) = \exp(G_\alpha(y) - 1)$ (see Appendix C).

(ii) The exists a positive measurable function $a(\cdot)$ such that the POT distribution converges to the GPD in the following way

$$\lim_{x \nearrow \infty} \mathbb{P}_t\left(\frac{\mathrm{VIX}_T - x}{a(x)} \geq y \Big| \mathrm{VIX}_T \geq x\right) = 1 - G_\alpha(y) \ .$$

For details on this theory and more general information about extreme-value theory, the reader is directed to [Deg06, Res87]. Assuming twice-differentiability in K, the Breeden-Litzenberger formula yields the VIX's distribution function, $\mathbb{P}_t(\mathrm{VIX}_T \geq K) = -\frac{1}{B_{t,T}}\frac{\partial}{\partial K}\mathrm{C}^{vix}(t,K,T)$. Then for two large strikes $K_0 < K$ a POT distribution can be written as

$$\mathbb{P}_t\left(\mathrm{VIX}_T \geq K_i \Big| \mathrm{VIX}_T \geq K_0\right) = \frac{\mathbb{P}_t(\mathrm{VIX}_T \geq K_i)}{\mathbb{P}_t(\mathrm{VIX}_T \geq K_0)} = \frac{\int_{K_i}^{\infty}\frac{\partial^2}{\partial K^2}\mathrm{C}^{vix}(t,K,T)dK}{\int_{K_0}^{\infty}\frac{\partial^2}{\partial K^2}\mathrm{C}^{vix}(t,K,T)dK} \ .$$

If the VIX options imply a value $\beta_R^{vix} \in (0,2)$, then it follows that there is $\tilde{p} \in (0,\infty)$ such that $\mathbb{E}_t\mathrm{VIX}_T^{1+\tilde{p}} = \infty$, and the VIX's distribution is heavy tailed. Moreover the stochastic volatility inspired (SVI) parameterization yields a GPD for the tail distribution.

Proposition 3.5. *Suppose the large-strike limit from equation* (3.2) *is applied to VIX call options to obtain*

$$\frac{\hat{\nu}^2(t,K,T)}{\log(K/X_{t,T})/(T-t)} \to \beta_R^{vix} \in (0,2) \qquad \text{as } K \to \infty.$$

Suppose further that the implied volatility surface is fit by the SVI parameterization, where the fit does not admit static arbitrage (i.e. no butterfly or calendar-spread arbitrage). Then the asymptotic behavior for the POT distribution with scaling function $a(x) = \frac{x}{\alpha}$ *is*

$$\lim_{x \nearrow \infty} \mathbb{P}_t\left(\frac{VIX_T - x}{x/\alpha} \geq y \Big| VIX_T \geq x\right) = \left(1 + \frac{y}{\alpha}\right)^{-\alpha}$$

for $y \geq 0$ *and* x *tending toward infinity, where* $\alpha = \frac{1}{2}\left(\sqrt{\frac{1}{\beta_R^{vix}}} + \frac{\sqrt{\beta_R^{vix}}}{2}\right)^2 > 1$ *for all* $\beta_R^{vix} < 2$.

Proof. No butterfly arbitrage in the SVI fit is enough for there to be an explicit formula for the VIX distribution's density function at time T, namely

$$\frac{\partial^2}{\partial K^2}\mathrm{C}^{vix}(t,K,T) = \frac{g(k)}{B_{t,T}\sqrt{2\pi\omega(k)}}\exp\left(-\frac{1}{2}d_-(k)^2\right) \ ,$$

where $k = \log(K/X_{t,T})$, $d_\pm(k) = -\frac{k}{\sqrt{\omega(k)}} \pm \frac{\sqrt{\omega(k)}}{2}$, $\omega(k) = \hat{\nu}^2(t,K,T)(T-t)$, and $g(k) = \left(1 - \frac{k\omega'(k)}{2\omega(k)}\right)^2 - \frac{(\omega'(k))^2}{4}\left(\frac{1}{\omega(k)} + \frac{1}{4}\right) + \frac{\omega''(k)}{2}$. The SVI fit to this slice of the implied-volatility surface is free from butterfly arbitrage if $g(k) \geq 0$ for all $k \in \mathbb{R}$ and $d_+(k) \to -\infty$ as $k \to \infty$ (see [GJ14]).

The moment formula in equation (3.2) says that $\omega(k) \sim k\beta_R^{vix}$ for k large k (i.e. as $K \nearrow \infty$), which applied to the density function yields

$$g(k) \sim \frac{1}{4} - \frac{\beta_R^2}{16} \qquad \text{for } k \text{ tending toward } \infty,$$
$$d_-(k) \sim -\sqrt{\frac{k}{\beta_R^{vix}}} - \frac{\sqrt{k\beta_R^{vix}}}{2} \qquad \text{for } k \text{ tending toward } \infty.$$

Placing these asymptotic approximations into the density yields

$$\frac{\partial^2}{\partial K^2}\mathrm{C}^{vix}(t,K,T)$$
$$\sim \frac{4-(\beta_R^{vix})^2}{16B_{t,T}\sqrt{2\pi k\beta_R^{vix}}}\exp\left(-\frac{k}{2}\left(\sqrt{\frac{1}{\beta_R^{vix}}}+\frac{\sqrt{\beta_R^{vix}}}{2}\right)^2\right) \qquad \text{for } k \text{ tending toward } \infty,$$

and integrating in k (or K with a change of variable) yields the tail distribution

$$-\frac{\partial}{\partial K}\mathrm{C}^{vix}(t,K,T) \sim \frac{\sqrt{2}(4-(\beta_R^{vix})^2)}{16B_{t,T}\sqrt{\beta_R^{vix}}}\int_{\sqrt{k}}^{\infty} e^{-\alpha u^2}du \qquad \text{for } k \text{ tending toward } \infty,$$

where $\alpha = \frac{1}{2}\left(\sqrt{\frac{1}{\beta_R^{vix}}}+\frac{\sqrt{\beta_R^{vix}}}{2}\right)^2 > 1$ for all $\beta_R^{vix} < 2$. Applying this asymptotic approximation of the tail distribution along with L'Hopitâles rule yields the result

$$\frac{\mathbb{P}_t\left(\mathrm{VIX}_T \geq x\left(1+\frac{y}{\alpha}\right)\right)}{\mathbb{P}_t(\mathrm{VIX}_T \geq x)} \sim \frac{\int_{\ell(x,y)}^{\infty} e^{-\alpha u^2}du}{\int_{\ell(x,0)}^{\infty} e^{-\alpha u^2}du} \sim \left(1+\frac{y}{\alpha}\right)^{-\alpha}$$

where $\ell(x,y) = \sqrt{\log(x(1+y/\alpha)/X_{t,T})}$. □

Proposition 3.5 shows that for $\beta_R^{vix} \in (0,2)$ with an SVI fit, the POT distribution with scaling $a(x) = \frac{x}{\alpha}$ converges to a GPD with parameter α. Hence, the Pickands-Balkema-de Haan theorem says this POT distribution is in the maximum domain of attraction of a generalized extreme-value distribution (see Appendix C).

Remark 3.6. *Large-strike asymptotics for SVI are presented in [FGGS11]. In particular, there is a large-k series expansion in powers of $k^{-1/2}$.*

Remark 3.7. *If $\beta_R^{vix} = 2$ then SVI may admit arbitrage as $d_+(k) \to 0$ as $k \to 0$, a limit which can be seen by using the large-k SVI expansion in [FGGS11]. If $\beta_R^{vix} = 0$ then the distribution is not heavy tailed and the extreme-value theory needs to be reworked to find an appropriate scaling function $a(\cdot)$ for the POT distribution.*

4. Examples. This section will give the reader a sense of how the theory from Section 3 applies to some widely-used stochastic volatility models. Models such as the 3/2 and SABR have heavy-tailed volatility, both of which have MGFs that are infinite on the positive real line, but each have different moment formula asymptotics for VIX implied volatility; the SABR model has a volatility process with moments of arbitrary order whereas the 3/2 model has some volatility moment that are infinite. This section also explores other models such as constant elasticity of volatility (CEV) and Heston.

4.1. CEV Model. The following example shows how special care is required in using models where the asset price does not satisfy Condition 2.2. Take $r = 0$, $T = 1$, and the process

$$dS_t = \sqrt{S_t}dW_t = \sigma_t S_t dW_t \qquad t \in [0,1]$$

with $S_0 = 1$ and $\sigma_t = \sigma(S_t) = \frac{1}{\sqrt{S_t}}$. The process S_t is a true martingale because $\int_\epsilon^\infty \frac{1}{s\sigma^2(s)}ds = \infty$ for some $\epsilon > 0$ (see [MM12, Pro13]), but was shown in [Fel51] to have positive probability of hitting zero in finite time, that is, $\mathbb{P}(\min_{t\leq 1} S_t = 0) > 0$ with zero being an absorbing boundary. Therefore $\mathrm{VIX}^2 = \mathbb{E}\int_0^1 \frac{1}{S_u}du = \infty$ and $\mathbb{E}S_t^{-q} = \infty$ for all $t \in (0,1]$ and all $q > 0$.

However, in a manner similar to the real-life variance-swap barriers discussed in [CL09], caps on payoffs based on realized variance can be modeled by using the stopping time

$$\mathcal{T} = \min\left\{t \geq 0 \Big| \int_0^t \frac{1}{S_u}du = M\right\}$$

and then by considering the stopped process $\tilde{S}_t = S_{t\wedge\mathcal{T}}$. The VIX on $\tilde{S}$ is then bounded,

$$\mathrm{VIX} = \sqrt{\mathbb{E}\int_0^{\mathcal{T}\wedge 1}\frac{1}{S_u}du} \leq \sqrt{M}\ ,$$

which is sufficient to have the Novikov condition for the martingales

$$\mathcal{Z}_t^{\pm} = \exp\left\{-\frac{1}{2}\int_0^{\mathcal{T}\wedge t}\frac{1}{S_u}du \pm \int_0^{\mathcal{T}\wedge t}\frac{1}{\sqrt{S_u}}dW_u\right\}\ .$$

Hence, $\tilde{S}_t = \mathcal{Z}_t^+$ and for any $q \in (0,1]$ the negative moments are

$$\begin{aligned}
\mathbb{E}\tilde{S}_t^{-q} &= \mathbb{E}\left\{(\mathcal{Z}_t^+)^{-q}\right\} \\
&= \mathbb{E}\exp\left\{\frac{q}{2}\int_0^{\mathcal{T}\wedge t}\frac{1}{S_u}du - q\int_0^{\mathcal{T}\wedge t}\frac{1}{\sqrt{S_u}}dW_u\right\} \\
&= \mathbb{E}\left\{(\mathcal{Z}_t^-)^q \exp\left\{q\int_0^{\mathcal{T}\wedge t}\frac{1}{S_u}du\right\}\right\} \\
&\leq e^{qM}\mathbb{E}\left\{(\mathcal{Z}_t^-)^q\right\} \leq e^{qM} \\
&< \infty\ ,
\end{aligned}$$

which is sufficient (along with the fact that $\mathbb{E}S_t^2 = \mathbb{E}\left(1 + \int_0^t \sqrt{S}_u dB_u\right)^2 = 1 + t < \infty$) to have Condition 2.2 for $\tilde{S}_t$.

This example is informative because it is a non-trivial case where a barrier (i.e. via a stopping time) is used to ensure $\mathrm{VIX}_t < \infty$ a.s. Without the stopping time there would be positive probability of infinite VIX, which would lead to infinite variance-swap rates.

4.2. SABR Model. This is a non-trivial example of the contraposition of Proposition 3 because Condition 2.1 holds yet Condition 2.2 does not because there are no negative moments on S_t.

For $t \in [0, T+\tau]$, take $r = 0$, $S_0 = 1$ and $Y_0 > 0$, and consider the SABR stochastic volatility model

$$
\begin{aligned}
d\log S_t &= -\frac{1}{2}Y_t^2 dt + Y_t\left(\sqrt{1-\rho^2}dW_t + \rho dB_t\right)\\
dY_t &= \alpha Y_t dB_t
\end{aligned}
$$

where $\alpha > 0$, $W \perp B$, and $-1 \leq \rho \leq -1$. It is well known that S_t is a true martingale if and only if $\rho \leq 0$ (see [Jou04]), and that $\mathbb{E}S_t^{1+p} < \infty$ if and only if $p \leq \frac{\rho^2}{1-\rho^2}$ (see Appendix B.2 or [Jou04]). The process Y_t is log-normal and almost-surely positive, and clearly $\mathbb{E}Y_t^2 < \infty$ implying that $\mathbb{E}\left(\int_0^t Y_u dB_u\right)^2 < \infty$, which implies Condition 2.1 and therefore the log-contract satisfies

$$
-\mathbb{E}\log(S_t) = \frac{1}{2}\mathbb{E}\int_0^t Y_u^2 du < \infty \ .
$$

Moreover, the VIX has all it's moments

$$
\begin{aligned}
\mathbb{E}\mathrm{VIX}_T^{2n} &= \frac{1}{\tau^n}\mathbb{E}\left(\mathbb{E}_T\int_T^{T+\tau} Y_u^2 du\right)^n \leq \frac{1}{\tau^{n-1}}\mathbb{E}\int_T^{T+\tau} Y_u^{2n} du\\
&= \frac{Y_0^{2n}}{\tau^{n-1}}\int_T^{T+\tau} e^{(2n^2-n)\alpha^2 u} du\\
&< \infty \ ,
\end{aligned}
$$

yet has infinite MGF,

$$
\mathbb{E}e^{\xi \mathrm{VIX}_T} = \mathbb{E}e^{\xi\sqrt{\frac{1}{\tau}\mathbb{E}_T\int_T^{T+\tau} Y_u^2 du}} = \mathbb{E}e^{\xi C_{\tau,\alpha} Y_T} = \infty \qquad \forall \xi > 0 \ ,
$$

where $C_{\tau,\alpha} = \sqrt{\frac{1}{\tau Y_T^2}\mathbb{E}_T\int_T^{T+\tau} Y_u^2 du} = \sqrt{\frac{1}{\tau\alpha^2}\left(e^{\alpha^2\tau} - 1\right)} > 0$. It follows that $\mathbb{E}e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$.

In summary, for $\rho < 0$ the price S_t is a martingale, there exists $\tilde{p} = \sup\{p \geq |\mathbb{E}_{S+\tau}^{1+p} < \infty\} > 0$, and $\mathbb{E}e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$. Hence from equation (3.1) it is deduced that $\tilde{q} = \sup\{q \geq 0 | \mathbb{E}S_{T+\tau}^{-q} < \infty\} = 0$. The SABR model is a good model because it fits the SPX options data. However, SABR may have trouble fitting both SPX and VIX options because volatility modeled as geometric Brownian motion will produce a flat smile that does not have the upward skew observed from VIX options.

This example has demonstrated the VIX-SPX relationship identified in Lemma **??**, namely that $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$ implies $\mathbb{E}_t S_{T+\tau}^{-q} = \infty$ for all $q > 0$.

4.3. CEV Volatility Process. This is another non-trivial example of the contraposition of Proposition 3, but the model is slightly better than SABR for pricing both SPX and VIX options because it has implied-volatility function $\hat{\nu}(t, K, T)$ that is increasing in K. For $t \in [0, T+\tau]$, take

r =0, $S_0 = 1$, and consider the stochastic volatility model

$$
\begin{aligned}
d\log S_t &= -\frac{1}{2}Y_t^2 dt + Y_t\left(\sqrt{1-\rho^2}dW_t + \rho dB_t\right)\\
dY_t &= cY_t^2 dB_t
\end{aligned}
$$

where $c > 0$ and $dW_t \cdot dB_t = \rho dt$. The process Y_t is almost-surely positive, a fact that can be deduced from it's transition density (see the transition density in [CH05]).

Also from the transition density it is seen that $\mathbb{E}Y_t^p < \infty$ for $0 \leq p \leq 3$ and $\mathbb{E}Y_t^p = \infty$ for $p > 3$. Since $\mathbb{E}Y_t^2 < \infty$ it follows that $\mathbb{E}\left(\int_0^{T+\tau} Y_u dB_u\right)^2 < \infty$ and Condition 2.1 holds, which means the log-contract satisfies

$$
-\mathbb{E}\log(S_t) = \frac{1}{2}\mathbb{E}\int_0^t Y_u^2 du < \infty \ .
$$

However, the VIX's MGF is infinite on the positive real line. Hence, if $\rho = 0$ or if $\rho < 0$, then S_t is a martingale and applying Part 1 of Proposition 3.3 shows that Condition 2.2 fails because $S_{T+\tau}$ has no negative moments. On the other hand, VIX call options have right-hand extreme strikes with $\beta_R = .4495$, as shown in Figure 4.1.

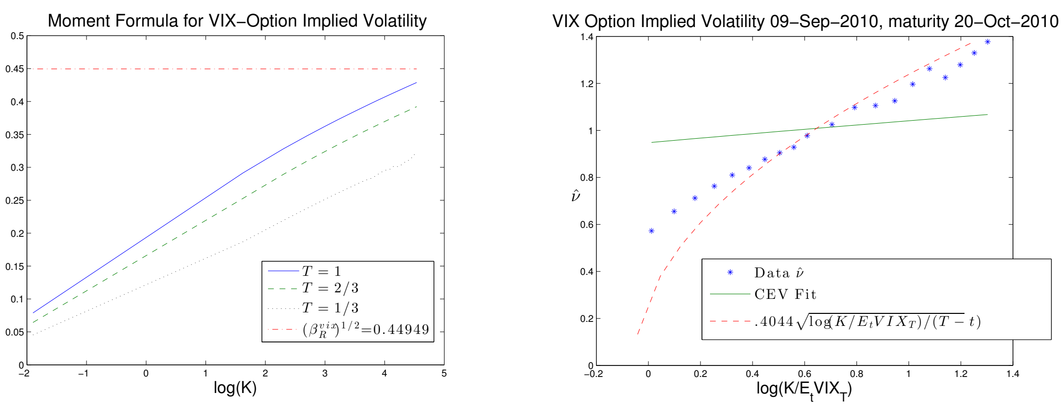

Fig. 4.1: **Left:**Implied volatility for the CEV volatility process in Section 4.3. The plot is of the scaled VIX implied volatility, $\frac{\hat{\nu}(t,K,T)\sqrt{T-t}}{\log(K_{max}/\mathbb{E}_t S_T)} \leq (\beta_R^{vix})^{1/2}$, for $Y_0 = .25$ and $K \leq K_{max} = 93.5$. The lines for $T = 1, 2/3$ and $1/3$ will come closer and closer to $\sqrt{\beta_R^{vix}} = 0.44949$ as K_{max} increases. **Right:** The fit of the the stochastic volatility model proposed in Section 4.3 is fit to the VIX options data from September 9th, 2010. The model captures some of upward skew in VIX implied volatility (certainly more than SABR which produces a flat implied volatility surface), yet does not have enough explanatory power to provide a good fit to the data.

4.4. The Heston Model. The Heston model is interesting because there is some interplay between the quantities $\tilde{q}$ and $\tilde{\xi}$ from Proposition 3.3. Furthermore, one of the model's most interesting features is the role played by time in determining negative moments, namely, for any $q > 0$

there exists $t^* \in (-\infty, T)$ such that $\mathbb{E}_t S_T^{-q} = \infty$ for $t \leq t^*$. However, this section will show that the Heston model's time dependence in tail risk is (for the most part) absent in VIX options.

The Heston model (with $r = 0$) has squared volatility given by a Cox-Ingersol-Ross (CIR) process, so that

$$d\log S_t = -\frac{1}{2}Y_t dt + \sqrt{Y_t}dW_t$$
$$dY_t = \kappa(\overline{Y} - Y_t)dt + \gamma\sqrt{Y_t}dB_t$$

where $dW_t dB_t = \rho dt$ for $\rho \in (-1, 0)$. It is also importnant have the Feller condition, $\gamma^2 \leq 2\overline{Y}\kappa$.

Similar to the computation used in [AP07, FGGS11], the earliest time $t^* < T$ such that $\mathbb{E}_{t^*} S_T^{-q} = \infty$ for some $q > 0$ (see Appendix B.3) is

$$T - t^* = \frac{\gamma^2}{\sqrt{\gamma^2 q(1+q) - (q\gamma\rho + \kappa)^2}} \left(\pi \mathbb{1}_{[q\gamma\rho+\kappa>0]} + \tan^{-1}\left(-\frac{q\gamma\rho + \kappa}{\sqrt{\gamma^2 q(1+q) - (q\gamma\rho+\kappa)^2}} \right) \right) .$$

Asymptotically, $T - t^* \sim \frac{\pi\gamma}{2q\sqrt{1-\rho^2}}$ for $q \gg 1$.

For the VIX, the MGF $\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} < \infty$ for some $\xi > 0$, but there is also a point ξ^* of explosion that can be calculated explicitly. Using the SDE for Y, the VIX is linear in Y_t, $\mathrm{VIX}_t^2 = \frac{1}{\tau}\mathbb{E}_t \int_t^{t+\tau} Y_u du = \overline{Y} + \frac{Y_t - \overline{Y}}{\kappa\tau}(1 - e^{-\tau\kappa}) = a + bY_t$. The explicit transition density for Y_t (as given in [AS99]) is

$$\frac{\partial}{\partial y}\mathbb{P}(Y_T \leq y | Y_t = y_0) = ce^{-u-v}(v/u)^{\alpha/2} I_\alpha(2\sqrt{uv}) \ , \tag{4.1}$$

where $c = 2\kappa/(\gamma^2(1 - e^{-\kappa(T-t)}))$, $u = cy_0 e^{-\kappa(T-t)}$, $v = cy$, $\alpha = 2\overline{Y}\kappa/\gamma^2 - 1 \geq 0$, and I_α is the modified Bessel function of the first kind of order α. The function $e^{\xi b y}$ is integrable against this density for any $\xi b < c$, hence, $\mathbb{E}_t e^{\xi \mathrm{VIX}_T} < \infty$ if and only if $\xi < \xi^*$ where

$$\xi^* = \frac{c}{b} = \frac{2\kappa^2\tau}{\gamma^2(1 - e^{-\kappa(T-t)})(1 - e^{-\tau\kappa})} \ .$$

Note that for large time $\frac{\partial}{\partial y}\mathbb{P}(Y_T \leq y | Y_t = y_0) = ce^{-u-v}(v/u)^{\alpha/2} I_\alpha(2\sqrt{uv}) \approx \frac{y^\alpha e^{-\frac{2\kappa}{\gamma^2}y}}{\left(\frac{\gamma^2}{2\kappa}\right)^{\alpha+1}\Gamma(\alpha+1)}$ (in fact, Y_t converges weakly to a gamma-distributed random variable). Hence for large time the VIX's MGF is

$$\mathbb{E}_t e^{\xi b Y_T} \approx \left(1 - \frac{\gamma^2 \xi b}{2\kappa}\right)^{-(\alpha+1)} \qquad \text{for } T - t \gg 1 \ ,$$

which exists for $\xi < \frac{2\kappa}{\gamma^2 b} = \frac{2\kappa^2\tau}{\gamma^2(1-e^{-\tau\kappa})}$.

The analysis above is interesting because it shows how the Heston model has a time dynamic in SPX tail risk that is not as prevalent in the VIX. To put it another way, there is an increase in tail risk as $\mathbb{E}_t S_T^{-q} = \infty$ for longer time-to-maturity, yet the VIX's MGF exists for a segment of the positive real line for all time. This means that the Heston-model gives VIX prices that do not capture the same long-term risk that is prevalent in the underlying.

Figure 4.2 shows how the Heston model can fit the SPX implied volatility, but has some difficulty in fitting the VIX implied volatility. In particular, the CIR process of the Heston model leads to a downward slope in VIX option implied volatility, which is the stylistic feature pointed out in [Dri12, Gat08, PS14] that suggests the Heston model is mis-specified to the VIX data.

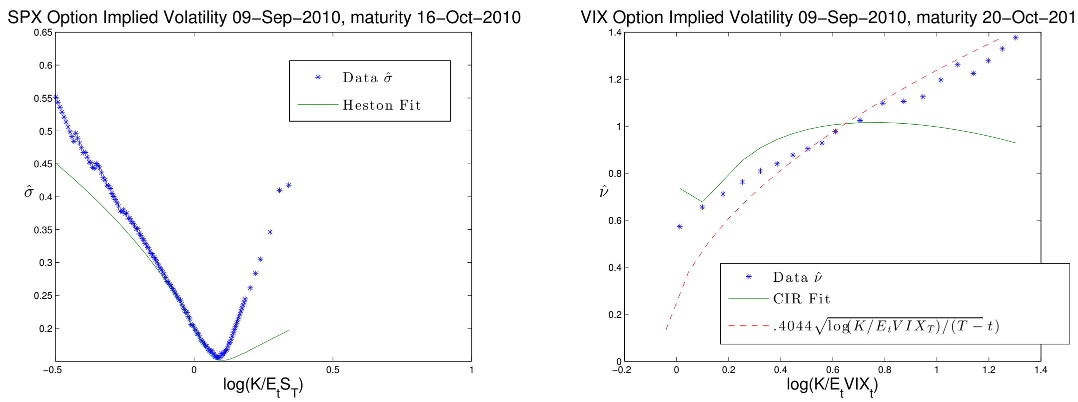

Fig. 4.2: **Left:** A fit of the Heston model to the implied volatility smile from SPX options. **Right:** A fit of the Heston model (or simply a CIR process) to the implied volatility from VIX options. The downward slope in the right-hand skew indicates

The Heston model has $\tilde{\xi} > 0$ and $\beta_L < 2$, yet there is not enough tail behavior for Proposition 3.3 to be an informative estimate. Namely, $\sqrt{\left(\frac{4\tilde{\xi}}{\tau}\right)^2 + \frac{4\tilde{\xi}}{\tau}} - \frac{4\tilde{\xi}}{\tau}$ is a small number for calibrations of Heston model having $\overline{Y}$ that is within the realm of what is (historically) standard in the data. In general, Proposition 3.3 is model free, which means there is an underlying tradeoff between sharpness and specification of the model. Sharper estimates based on Proposition 3.3 may be available for Heston and other stochastic volatility models.

4.5. The 3/2 Model. The 3/2 model is

$$d\log S_t = -\frac{1}{2}Z_t dt + \sqrt{Z_t}dW_t$$
$$dZ_t = Z_t\left(\kappa - (\gamma^2 - \kappa\overline{Y})Z_t\right)dt + \gamma Z_t^{3/2}dB_t \ ,$$

where $2\overline{Y}\kappa \geq \gamma^2$, and $\kappa - \gamma\rho > -\frac{1}{2}\gamma^2$ so that the price process is a martingale (see [Dri12]), and hence Condition 2.1 holds up to time $T+\tau$. This model can be equivalently written as $d\log S_t = -\frac{1}{2Y_t}dt + \frac{1}{\sqrt{Y_t}}dW_t$ where Y is the square-root process from Section 4.4. This is a popular choice for pricing VIX options because the volatility process has heavy tails (see [BB14, Dri12]). However, this model forces low-strike implied volatility of SPX options to be at its maximum in the same manner as the examples in Sections 4.2 and 4.3.

For some $k \geq \alpha+1$ (where $\alpha = \frac{2\overline{Y}\kappa}{\gamma^2} - 1$ as in Section 4.4), convexity arguments lead to non-finite moment,

$$
\begin{aligned}
\mathbb{E}_t(\mathrm{VIX}_T^2)^k &= \mathbb{E}_t\left(\frac{1}{\tau}\int_T^{T+\tau}\mathbb{E}_T\left(\frac{1}{Y_u}\right)du\right)^k \\
&> \mathbb{E}_t\left(\frac{1}{\tau}\int_T^{T+\tau}\left(\frac{1}{\mathbb{E}_T Y_u}\right)du\right)^k \\
&> \mathbb{E}_t\left(\frac{1}{\tau}\int_T^{T+\tau}\left(\frac{1}{Y_T}\right)du\mathbb{1}_{[Y_T>\overline{Y}]}\right)^k \qquad \text{(because } Y_T > \mathbb{E}_T Y_u \text{ if } Y_T > \overline{Y}\text{)} \\
&= \mathbb{E}_t\left(\frac{1}{Y_T}\right)^k \mathbb{1}_{[Y_T>\overline{Y}]} \\
&= \infty \qquad \text{for all } k \geq \alpha+1 \ ,
\end{aligned}
$$

where the transition density given in (4.1) is used to verify that $\mathbb{E}_t\left(\frac{1}{Y_T}\right)^k = \infty$ for $k \geq \alpha + 1$. Therefore the MGF of VIX_T^2 is infinite on the positive real line,

$$
\mathbb{E}_t e^{\xi \mathrm{VIX}_T^2} > \frac{\xi^k}{k!}\mathbb{E}_t(\mathrm{VIX}_T^2)^k = \infty \qquad \text{for all } k \geq \alpha+1 \ ,
$$

and hence S_t has no negative moments. Moreover, a Hölder inequality can be used to show that finiteness for $1 \leq k < \alpha + 1$:

$$
\begin{aligned}
\mathbb{E}_t(\mathrm{VIX}_T^2)^k &= \mathbb{E}_t\left(\frac{1}{\tau}\int_T^{T+\tau}\mathbb{E}_T\left(\frac{1}{Y_u}\right)du\right)^k \\
&< \left(\frac{1}{\tau}\int_T^{T+\tau}\mathbb{E}_t\left(\mathbb{E}_T\left(\frac{1}{Y_u}\right)\right)^k du\right)^{1/k} \\
&< \left(\frac{1}{\tau}\int_T^{T+\tau}\mathbb{E}_t\left(\frac{1}{Y_u^k}\right)^{1/k} du\right)^{1/k} \\
&< \infty \qquad \text{for all } 1 \leq k < \alpha+1 \ .
\end{aligned}
$$

Hence, $\mathbb{E}_t(\mathrm{VIX}_T^2)^k < \infty$ for all $0 \leq k < \alpha + 1$.

In general, the CEV volatility model of Section 4.3, the SABR, and the 3/2 models are candidates for an improved fit to the VIX data because the volatility process that is heavy tailed, whereas the Heston model's CIR process does not capture the right-hand skew in VIX implied volatility. However, the 3/2 model appear to be the best for fitting to VIX-option implied volatility, as the SABR model has flat smile for VIX options, the CEV volatility model has relatively little skew for VIX options (see right-hand plot in Figure 4.1), while the 3/2 model has captured much of the right-hand skew in the data (see Figure 4.3). The moment relationships form the examples of Sections 4.1 to 4.5 are summarized in Table 4.1.

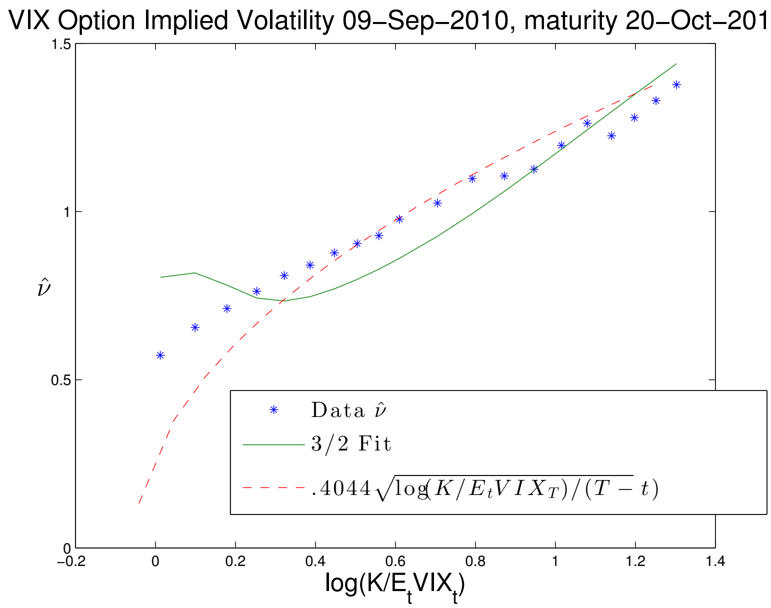

Fig. 4.3: The 3/2 model fit to implied Volatility from VIX options. The heavy-tailed 3/2 process for volatility captures the increasing right-hand skew. The right-hand skew is the stylistic feature that is seen through the VIX data, and it is important to use a stochastic volatility model that fits when pricing VIX options. See [Dri12, Gat08, BGK13, PS14] for more insight on models that fit the right-hand skew.

5. Conclusions. This paper has explored some basic relationships between the markets for SPX and VIX options. The main idea is based on the notion that high-strike VIX call options can be used to hedge tail risk in the SPX. The moment formula was applied to relate the extreme-strike options on the SPX and VIX, and some formulas for comparison were introduced. The primary focus was the relationship between negative moments in SPX and the interval of the positive real line where the MGF of VIX-squared is finite. Negative moments and the MGF were computed for various stochastic volatility models, with the intention of giving the reader a sense of what can be accomplished with different models.

This paper is a step towards a theory for unifying option pricing of VIX and SPX contracts under a single stochastic volatility model. There are studies in the literature that have demonstrated better fits by richer models, but still there must be special taken when using a single model for pricing both SPX and VIX options. The primary contribution of this paper is the model-free understanding of links between these two options markets.

Possible future work will be refined arbitrage bounds for specific models, and more development of the model-free framework with the identification of no-arbitrage bounds that can be enforced with static hedging.

Model	$\mathbb{E}\mathrm{VIX}_T^p = \infty$	$\mathbb{E}e^{\xi \mathrm{VIX}_T^2} = \infty$	$\mathbb{E}S_{T+\tau}^{-q} = \infty$
CEV model, $dS = S^a dW$ with $0 \leq a < 1$	$\forall p > 0$	$\forall \xi > 0$	$\forall q > 0$
SABR with $\rho \leq 0$	$p = \infty$	$\forall \xi > 0$	$\forall q > 0$
CEV volatility, $dS = S\sqrt{Y}dW$ and $dY = cY^2 dB$, with $dWdB = \rho dt$ and $\rho \leq 0$	$\forall p > 3$	$\forall \xi > 0$	$\forall q > 0$
Heston Model	$p = \infty$	$\forall \xi \geq \frac{2\kappa^2\tau}{\gamma^2(1-e^{-\kappa T})(1-e^{-\tau\kappa})}$	$\forall q$ such that $T+\tau \geq T^*(q)$
3/2 model	$\forall p \geq 2(\alpha+1)$	$\forall \xi > 0$	$\forall q > 0$

Table 4.1: A table to summarize the moment relationships from the examples in the Section 4. For the Heston model, the function $T^*(q) = \frac{\gamma^2}{\sqrt{\gamma^2 q(1+q)-(q\gamma\rho+\kappa)^2}}\left(\pi \mathbb{1}_{[q\gamma\rho+\kappa>0]} + \tan^{-1}\left(-\frac{q\gamma\rho+\kappa}{\sqrt{\gamma^2 q(1+q)-(q\gamma\rho+\kappa)^2}}\right)\right)$. For the 3/2 model, $\alpha = \frac{2\overline{Y}\kappa}{\gamma^2} - 1$. For each of these models it was shown that $S_t e^{-rt}$ is a true martingale, so from Lemma 3.2 is follows that $\mathbb{E}S_T^{-q} = \infty$ for all $q > 0$ if $\mathbb{E}e^{\xi \mathrm{VIX}_T^2} = \infty$ for all $\xi > 0$.

Appendix A. The CEV Model. Consider the CEV model with quadratic variance,

$$dS_t = \sigma S_t^2 dW_t \ .$$

This process is a strict local martingale and is discussed in [CH05]. In particular, the process $Y_t = 1/S_t^2$ is among the class of SDEs considered in [Fel51], and has natural boundaries at zero and infinity (i.e. both Y and S have zero probability of touching zero). Furthermore, the transition density of S is

$$\mathbb{P}(S_T \in dz | S_t = s) = \frac{s}{z^3}\frac{dz}{\sqrt{2\pi(T-t)\sigma^2}} \times \left(\exp\left(-\frac{\left(\frac{1}{z}-\frac{1}{s}\right)^2}{2(T-t)\sigma^2}\right) - \exp\left(-\frac{\left(\frac{1}{z}+\frac{1}{s}\right)^2}{2(T-t)\sigma^2}\right)\right) \ .$$

Figure A.1 and Table A.1 show the expectations of some important functions of S_T for $\sigma = 1$.

Appendix B. Moment Calculations.

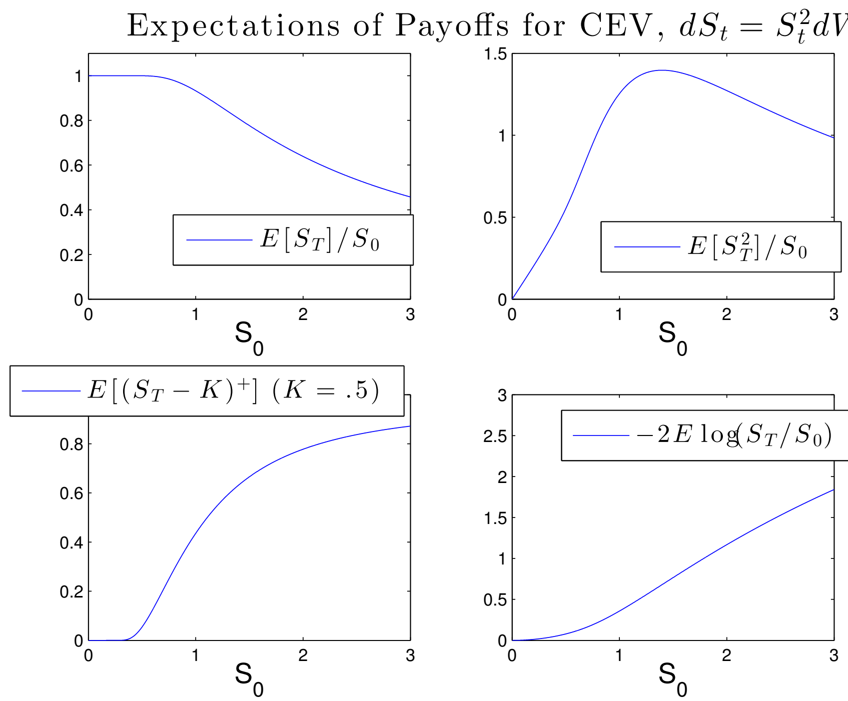

Fig. A.1: Some expectations of functions of the CEV process. **Top Left:** Notice that $\mathbb{E}_0 S_T/S_0 < 1$ for $S_0 > 0$; this is the effect of the local martingale property. **Top Right:** Notice how $\mathbb{E}_0 S_T^2/S_0 = \mathcal{O}(S_0)$; this is useful in Example 4.3 where this CEV model is used as the volatility process. **Bottom Left:** Notice how the call option price in is not convex in S_0. As mentioned in [CH05], this shows that risk-neutral expectations of convex payoffs may not be convex in the initial price if the process is only a local martingale. **Bottom Right:** The log contract, which is increasing but not concave in S_0.

B.1. SABR Model (Section 4.2). Furthermore, for any $p > 0$ the $1 + p$ moment is

$g(S_T)$	$\mathbb{E}\{g(S_T)\|S_0 = s\}$
S_T	$s\left(1 - 2\Phi\left(\frac{-1}{s\sqrt{T}}\right)\right)$,
S_T^2	$\sqrt{\frac{2s^2}{T}}D_+\left(\frac{1}{s\sqrt{2T}}\right)$,
$\log(S_T)$	$\frac{1}{2}\left(2 + \gamma_e + \log(2/T) + \frac{\partial}{\partial a}{}_1F_1\left(0, \frac{3}{2}, \frac{-1}{2Ts^2}\right)\right)$,
$(S_T - K)^+$	$s\left\{\Phi(\kappa-\delta) - \Phi(-\delta) + \Phi(\delta) - \Phi(\delta+\kappa)\right\}$ $-K\left\{\Phi(\kappa+\delta) - \Phi(\delta-\kappa) + \delta^{-1}\left(\Phi'(\kappa+\delta) - \Phi'(\kappa-\delta)\right)\right\}$,
	$\delta = \frac{1}{s\sqrt{T}}$ $\kappa = \frac{1}{K\sqrt{T}}$

Table A.1: The moments for the CEV model $dS = S^2 dW$. The special functions are the normal Gaussian CDF Φ, the Dawson integral $D_+ = e^{-x^2}\int_0^x e^{u^2}du$, the confluent hypergeometric function of the first kind ${}_1F_1 = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 e^{ux}u^{a-1}(1-u)^{b-a-1}du$, and the Euler Gamma $\gamma_e \approx 0.577215665$. For general $\sigma > 0$ the transition density shows that $\sigma \neq 1$ is the scaling time given by $\mathbb{E}^\sigma\{g(S_T)|S_0 = s\} = \mathbb{E}^1\{g(S_{T\sigma^2})|S_0 = s\}$.

$$
\begin{aligned}
\mathbb{E}S_t^{1+p} &= \mathbb{E}e^{(1+p)\log S_t} \\
&= \mathbb{E}\exp\left\{-\frac{1+p}{2}\int_0^t Y_u^2 du + (1+p)\sqrt{1-\rho^2}\int_0^t Y_u dW_u + (1+p)\rho\int_0^t Y_u dB_u\right\} \\
&= \mathbb{E}\exp\left\{\frac{-(1+p)+(1+p)^2(1-\rho^2)}{2}\int_0^t Y_u^2 du + (1+p)\rho\int_0^t Y_u dB_u\right\} \\
&= \mathbb{E}\exp\left\{\frac{-(1+p)+(1+p)^2(1-\rho^2)}{2}\int_0^t Y_0^2 e^{-\alpha^2 u + 2\alpha B_u}du + \frac{(1+p)\rho(Y_t - Y_0)}{\alpha}\right\} \\
&< e^{-\frac{(1+p)\rho Y_0}{\alpha}}\mathbb{E}\left\{\exp\left\{\frac{-(1+p)+(1+p)^2(1-\rho^2)}{2}Y_0^2 e^{-\alpha^2 t}\int_0^t e^{2\alpha B_u}du\right\}\right\} \\
&< \infty\ ,
\end{aligned}
$$

which is the case if (and only if) $\frac{-(1+p)+(1+p)^2(1-\rho^2)}{2} \leq 0$ if and only if $p \leq \frac{\rho^2}{1-\rho^2}$, otherwise if $\frac{-(1+p)+(1+p)^2(1-\rho^2)}{2} > 0$ it blows up:

$$\begin{aligned}
\mathbb{E}\left\{\exp\left\{\xi\int_0^t e^{2\alpha B_u}du\right\}\right\} &> \mathbb{E}\left\{\exp\left\{\xi\int_{t/2}^t e^{2\alpha B_u}du\right\}\right\} \\
&= \mathbb{E}\left\{\exp\left\{\xi e^{2\alpha B_{t/2}}\int_{t/2}^t e^{2\alpha(B_u-B_{t/2})}du\right\}\right\} \\
&> \mathbb{E}\left\{\exp\left\{\xi e^{2\alpha B_{t/2}}\right\}\mathbb{1}_{\left[\int_{t/2}^t(B_u-B_{t/2})du>0\right]}\right\} \\
&= \mathbb{E}\exp\left\{\xi e^{2\alpha B_{t/2}}\right\}\mathbb{E}\left\{\mathbb{1}_{\left[\int_{t/2}^t(B_u-B_{t/2})du>0\right]}\right\} \\
&= \infty\ ,
\end{aligned}$$

because $\mathbb{E}\exp\left\{\xi e^{2\alpha B_{t/2}}\right\} = \infty$ for all $\xi > 0$.

B.2. CEV Volatility Process (Section4.3). To determine whether or not S_t is a true martingale it suffices to consider the expectation

$$\begin{aligned}
\mathbb{E}S_T &= \mathbb{E}\exp\left\{-\frac{1}{2}\int_0^T Y_u^2du + \sqrt{1-\rho^2}\int_0^T Y_udW_u + \rho\int_0^T Y_udB_u\right\} \\
&= \mathbb{E}\exp\left\{-\frac{\rho^2}{2}\int_0^T Y_u^2du + \rho\int_0^T Y_udB_u\right\} \\
&= \mathbb{E}\left\{\left(\frac{Y_T}{Y_0}\right)^{\rho/c}\exp\left\{\frac{\rho(c-\rho)}{2}\int_0^T Y_u^2du\right\}\right\}\ ,
\end{aligned}$$

which is certainly a martingale if $\rho = 0$. For $\rho < 0$, define $\mathcal{Z}_t = \left(\frac{Y_t}{Y_0}\right)^{\rho/c}\exp\left\{\frac{\rho(c-\rho)}{2}\int_0^t Y_u^2du\right\}$ and apply Itô's lemma to get

$$\mathbb{E}S_T = \mathbb{E}\mathcal{Z}_T = 1 + \mathbb{E}\left\{\rho\int_0^T Y_u\mathcal{Z}_udB_u\right\} = 1\ ,$$

where the stochastic integral is a martingale because $\mathbb{E}\int_0^T(Y_u\mathcal{Z}_u)^2du \le \int_0^T(\mathbb{E}Y_u^3)^{2/3}(\mathbb{E}\mathcal{Z}_u^6)^{1/3}du < \infty$ for $\rho < 0$ and any $c > 0$, and hence S_t is a true martingale.

The MGF VIX-squared is infinite on the positive real line, which is seen by writing the second moment of Y_t in terms of the Dawson integral from Table A.1, and then by taking constant $M \gg 1$,

$\epsilon > 0$, and evaluating for $\xi > 0$:

$$
\begin{aligned}
\mathbb{E}e^{\xi \mathrm{VIX}_T^2} &= \mathbb{E}e^{\frac{\xi}{\tau}\mathbb{E}_T \int_T^{T+\tau} Y_u^2 du} \\
&= \mathbb{E}e^{\xi \frac{Y_T}{\tau} \int_T^{T+\tau} \sqrt{\frac{2}{c^2(u-T)}} D_+\left(\frac{1}{Y_T\sqrt{2c^2(u-T)}}\right) du} \\
&= \mathbb{E}e^{\xi \frac{2}{\tau c^2} \int_{\frac{1}{Y_T\sqrt{2c^2\tau}}}^{\infty} \frac{1}{x^2} D_+(x) dx} \\
&> \liminf_{M\to\infty} \mathbb{E}\left\{ \exp\left\{ \xi \frac{2}{\tau c^2} \int_{\frac{1}{Y_T\sqrt{2c^2\tau}}}^{\frac{1}{Y_T\sqrt{2c^2\tau}}+\epsilon} \frac{1}{x^2} D_+(x) dx \right\} \mathbb{1}_{[Y_T \geq M]} \right\} \\
&> \liminf_{M\to\infty} \mathbb{E}\left\{ \exp\left\{ \xi \frac{2}{\tau c^2} \int_{\frac{1}{Y_T\sqrt{2c^2\tau}}}^{\frac{1}{Y_T\sqrt{2c^2\tau}}+\epsilon} \left(\frac{1}{x} - \frac{2x}{3}\right) dx \right\} \mathbb{1}_{[Y_T \geq M]} \right\} \\
&> \liminf_{M\to\infty} \mathbb{E}\Bigg\{ \exp\Bigg\{ \xi \frac{2}{\tau c^2} \int_{\frac{1}{Y_T\sqrt{2c^2\tau}}}^{\frac{1}{Y_T\sqrt{2c^2\tau}}+\epsilon} \\
&\qquad\qquad \times \left(Y_T\sqrt{2c^2\tau} - \frac{4}{3M\sqrt{2c^2\tau}} - \frac{2\epsilon}{3} \right) dx \Bigg\} \mathbb{1}_{[Y_T \geq M]} \Bigg\} \\
&= \liminf_{M\to\infty} e^{-\xi\frac{2\epsilon}{\tau c^2}\left(\frac{4}{3M\sqrt{2c^2\tau}} + \frac{2\epsilon}{3}\right)} \mathbb{E}\left\{ e^{\xi \frac{\epsilon 2\sqrt{2}}{\sqrt{c^2\tau}} Y_T} \mathbb{1}_{[Y_T \geq M]} \right\} = \infty \ ,
\end{aligned}
$$

where we've taken the change of variable $x = \frac{1}{Y_T\sqrt{2c^2(u-T)}}$, and used the Dawson integral's MacLauren series

$$
D_+(x) = \sum_{n=0}^{\infty} \frac{(-1)^n 2^n}{(2n+1)!!} x^{2n+1} > x - \frac{2}{3}x^3 \qquad \text{for } x < 1 \ .
$$

The $\liminf$ is infinite because $\mathbb{E}\left\{e^{\xi Y_T}\mathbb{1}_{[Y_T \geq M]}\right\} > \frac{\xi^4}{4!}\mathbb{E}\left\{Y_T^4 \mathbb{1}_{[Y_T\geq M]}\right\} = \infty$ for all $\xi > 0$ and $M < \infty$.

B.3. Heston Model (Section 4.4). Similar to the computation used in [AP07, FGGS11], the earliest time $t^* < T$ such that $\mathbb{E}_{t^*} S_T^{-q} = \infty$ for some $q > 0$ can be computed from an ansatz for the PDE's solution. Let $X_t = -q \log S_t$, which has Itô differential $dX_t = -q d\log S_t = \frac{qY_t}{2}dt - q\sqrt{Y_t}dW_t$, and the function $u(t,x,y) = \mathbb{E}\{S_T^{-q} | X_t = x, Y_T = y\}$ satisfies the PDE

$$
u_t + \frac{q^2 y}{2} u_{xx} + \frac{qy}{2} u_x - q\gamma\rho y u_{xy} + \frac{\gamma^2 y}{2} u_{yy} + \kappa(\overline{Y} - y) u_y = 0
$$

with terminal condition $u(T,x,y) = e^x$. Now consider the ansatz $u(t,x,y) = e^x e^{a(t)y + b(t)}$ where $a(T) = 0 = b(T)$. There is a Riccati equation for $a(t)$ and an associated ODE for $b(t)$:

$$
\begin{aligned}
\frac{\partial}{\partial t} a(t) &= -\frac{\gamma^2}{2} a^2(t) + (q\gamma\rho + \kappa)\, a(t) - \left(\frac{q^2}{2} + \frac{q}{2}\right) \\
&= -\frac{\gamma^2}{2} \left(a(t) - r_-\right)\left(a(t) - r_+\right) \\
\frac{\partial}{\partial t} b(t) &= -\kappa \overline{Y} a(t) \ ,
\end{aligned}
$$

where $r_{\pm} = \frac{1}{\gamma^2}\left(q\gamma\rho + \kappa \pm \sqrt{(q\gamma\rho+\kappa)^2 - \gamma^2 q(1+q)}\right)$. The solution $a(t) = \infty$ for some $t^* > -\infty$ if there are complex-valued $r_{\pm}$, which happens if $(q\gamma\rho + \kappa)^2 < \gamma^2 q(1+q)$ or (equivalently) if $q \notin [q_-, q_+]$ where

$$q_{\pm} = \frac{2\gamma\rho\kappa - \gamma^2 \pm \sqrt{\left(2\gamma\rho\kappa - \gamma^2\right)^2 + 4\kappa^2\gamma^2(1-\rho^2)}}{2(1-\rho^2)\gamma^2} \tag{B.1}$$

(notice $q_- < 0 < q_+$). To see why $a(t)$ blows up in finite time, rewrite the Riccati equation as a linear 2nd order equation,

$$v''(t) - (q\gamma\rho + \kappa)\, v'(t) + \frac{\gamma^2(q^2+q)}{4} v(t) = 0$$

so that the solution is $a(t) = 2\frac{v'(t)}{\gamma^2 v(t)}$ where for complex roots (e.g. for $q > q^+$) there is general solution

$$v(t) = e^{-\frac{q\gamma\rho+\kappa}{\gamma^2}(T-t)}\left(C_1 \cos\left(d(T-t)\right) + C_2 \sin\left(d(T-t)\right)\right) \ ,$$

with constants $C_1, C_2 \in \mathbb{R}$ and where $d = \frac{\sqrt{\gamma^2 q(1+q) - (q\gamma\rho+\kappa)^2}}{\gamma^2}$. This solution oscillates above and below zero, and hence, will equal zero at some finite time $t^* \in (-\infty, T)$, at which point $a(t^*) = \infty$. At time $t = T$ the terminal condition $a(T) = 0$ requires

$$v'(T) = \frac{q\gamma\rho + \kappa}{\gamma^2} C_1 - dC_2 = 0 \ ,$$

which implies $\frac{C_2}{C_1} = \frac{q\gamma\rho+\kappa}{d\gamma^2}$. Then $\mathbb{E}_t S_T^{-q} < \infty$ if $t \in (t^*, T]$, where $t^* < T$ is such that $v(t^*) = 0$ or

$$\cos\left(d(T-t^*)\right) + \frac{C_2}{C_1}\sin\left(d(T-t^*)\right) = \cos\left(d(T-t^*)\right) + \frac{q\gamma\rho+\kappa}{d\gamma^2}\sin\left(d(T-t^*)\right) = 0 \ ,$$

and solving yields the singularity time

$$T - t^* = \frac{1}{d}\left(\pi \mathbb{1}_{[q\gamma\rho+\kappa>0]} + \tan^{-1}\left(-\frac{q\gamma\rho+\kappa}{\sqrt{\gamma^2 q(1+q) - (q\gamma\rho+\kappa)^2}}\right)\right) \ .$$

Appendix C. Maximum Domain of Attaction. Let $F(y) = \mathbb{P}_t(\mathrm{VIX}_T \leq y)$. Consider samples $(Y_\ell)_{\ell=1}^n$ where $Y_\ell \sim iidF$ for each ℓ, and let $M_n = \max(Y_1, Y_2, \ldots, Y_n)$. For $\alpha > 0$ the generalized extreme-value distribution is

$$H_\alpha(y) \triangleq \begin{cases} \exp\left(-(1+y/\alpha)^{-\alpha}\right) & \alpha < \infty \ , \\ \exp(-\exp(-y)) & \alpha = \infty \ . \end{cases}$$

Distribution function F is said to be in the maximum domain of attraction (MDA) of H_α for $\alpha < \infty$ if and only if

$$a_n M_n \Rightarrow H_\alpha \qquad \text{as } n \to \infty \ ,$$

where $a_n = F^{-1}(1 - 1/n)$ (see [Deg06] or [Res87] page 54-57). This is written as $F \in \mathrm{MDA}(H_\alpha)$.

For $\alpha = \infty$, $F \in mda(H_\alpha)$ if and only if

$$1 - F(y) \sim \exp(-\Psi(y)) \qquad \text{as } y \to \infty$$

for some function $\Psi \in C^2(\mathbb{R}^+)$ with (i) $\Psi(y) \to \infty$ as $y \to \infty$, (ii) $\Psi'(y) > 0$, and (iii) $(1/\Psi'(y))' \to 0$ as $y \to \infty$ (see [Deg06]).

REFERENCES

[AP07] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

[AS99] Y. Aït-Sahalia. Transition densities for interest rate and other nonlinear diffusions. *The Journal of Finance*, 54(4):1361–1395, August 1999.

[BB14] A. Badran and J. Baldeaux. Consistent modeling of VIX and equity derivatives using a 3/2 plus jumps model. *Applied Mathematical Finance*, 21(4):299–312, 2014.

[BF08] S. Benaim and P Friz. Smile asymptotics II: Models with known moment generating functions. *Journal of Applied Probability*, 45(1):16–32, 2008.

[BGK13] C. Bayer, J. Gatheral, and M. Karlsmark. Fast Ninomiya-Victoir calibration of the double-mean-reverting model. *Quantitative Finance*, 13(11):1813–1829, 2013.

[CH05] A. Cox and D. Hobson. Local martingales, bubbles and option prices. *Finance and Stochastics*, 9:477–492, 2005.

[CK13] R. Cont and T. Kokholm. A consistent pricing model for index options and volatility derivatives. *Mathematical Finance*, 23(2):248–274, 2013.

[CL08] P. Carr and R. Lee. Robust replication of volatility derivatives. In *MFA Annual Meeting*, 2008.

[CL09] P. Carr and R. Lee. Volatility derivatives. *Annual Review of Financial Economics*, 1:319–339, 2009.

[CL10] P. Carr and R. Lee. Hedging variance options on continuous semimartingales. *Finance and Stochastics*, 14(2):179–207, April 2010.

[CS07] P. Carr and J. Sun. A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, 10:87–150, May 2007.

[DDKZ99] K. Demeterfi, E. Derman, M. Kamal, and J. Zou. More than you ever wanted to know about volatility swaps. *The Journal of Derivatives*, 6(4):9–32, 1999.

[Deg06] M. Degan. *On Multivariate Generalised Pareto Distributions and High Risk Scenarios*. PhD thesis, ETH Zurich, 2006.

[Dri12] G. Drimus. Options on realized variance by transform methods: a non-affine stochastic volatility model. *Quantitative Finance*, 12(11):1679–1694, 2012.

[Fel51] W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54(1), July 1951.

[FG05] P. Friz and J. Gatheral. Valuation of volatility derivatives as an inverse problem. *Quantitative Finance*, 5(6):531–542, December 2005.

[FGGS11] P. Friz, S. Gerhold, A. Gulisashvili, and S. Sturm. On refined volatility smile expansion in the Heston model. *Quantitative Finance*, 11(8):1151–1164, 2011.

[Gat06] J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley, 2006.

[Gat08] J. Gatheral. Consistent modeling of SPX and VIX options. Presented at The Fifth World Congress of the Bachelier Finance Society in London, 2008. Available at: `http://www.math.nyu.edu/fellows_fin_math/gatheral/Bachelier2008.pdf`.

[GJ14] J. Gatheral and A. Jaquier. Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, 14(1):59–71, 2014.

[Jou04] B. Jourdain. Loss of martingality in asset price models with lognormal stochastic volatility. Technical report, 2004. Preprint CERMICS 2004267. http://cermics.enpc.fr/reports/CERMICS-2004/CERMICS-2004-267. pdf.

[KL12] F. Klebaner and R. Lipster. When a stochastic exponential is a true martingale. Extension of the Beneš method. *Theory Probab. Appl.*, 58(1):38–62, 2012.

[Lee04] R. Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, July 2004.

[MM12] A. Mijatovic and M.Urusov. On the martingale property of certain local martingales. *Probability Theory and Related Fields*, 152:1–30, 2012.

[MY11] D. Madan and M. Yor. The S&P 500 index as a Sato process travelling at the speed of the VIX. *Applied Mathematical Finance*, 18(3):227–244, 2011.

[Pro13] Philip Protter. A mathematical theory of financial bubbles. In *Paris-Princeton Lectures on Mathematical Finance 2013*, volume 2081 of *Lecture Notes in Mathematics*, pages 1–108. Springer International Publishing, 2013.

[PS14] A. Papanicolaou and R. Sircar. A regime-switching Heston model for VIX and S&P 500 implied volatilities. *Quantitative Finance*, 14(10):1811–1827, 2014.

[Res87] S. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer, New York NY, 1987.