ASYMPTOTICS OF IMPLIED VOLATILITY FAR FROM MATURITY

MICHAEL R. TEHRANCHI

ABSTRACT. This note explores the behavior of the implied volatility surface of a European call option far from maturity. Asymptotic formulae are derived with precise control over the error terms. The connection between the asymptotic implied volatility and the cumulant generating function of the logarithm of the underlying stock price is discussed in detail and illustrated by examples.

1. INTRODUCTION

Recall that the implied volatility of a European call option with strike K and time to maturity τ is defined as the unique non-negative solution Σ (if it exists) to the equation

$$C/S = \mathrm{BS}(\log(K/S) - r\tau, \tau\Sigma^2) \tag{1}$$

where C is the current price of the option, S is the price of the underlying stock (assumed to pay no dividends), r is the yield on a zero-coupon bond with the same maturity as the option, and the Black–Scholes call price function $\mathrm{BS} : \mathbb{R} \times [0, \infty) \to [0, 1)$ is defined by

$$\mathrm{BS}(k, v) = \begin{cases} \Phi\big(-\frac{k}{\sqrt{v}} + \frac{\sqrt{v}}{2}\big) - e^k \Phi\big(-\frac{k}{\sqrt{v}} - \frac{\sqrt{v}}{2}\big) & \text{if } v > 0 \\ (1 - e^k)^+ & \text{if } v = 0, \end{cases}$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$ is the standard normal distribution function. Just as the yield of a zero-coupon bond is a dimensionless quantification of the value of the bond, the implied volatility is often used to quote an option's price. Since $v \mapsto \mathrm{BS}(k, v)$ is strictly increasing onto the interval $[(1 - e^k)^+, 1)$, equation (1) has a solution (and hence the implied volatility is well-defined) if and only if $(S - Ke^{-r\tau})^+ \le C < S$.

This note provides a detailed asymptotic analysis of the implied volatility of options very far from maturity. We work within a standard no-arbitrage framework for modelling the prices of a given underlying stock and the options written on it. The motivation of this analysis is to better understand the constraints on the possible shapes of the implied volatility surface imposed by the no-arbitrage condition. Furthermore, the simple asymptotic formula can be used for model calibration, especially in models where the moment-generating function is known explicitly. Many of the techniques used here are familiar from the theory of large deviations, including saddle-point-type approximations. However, the treatment of the borderline and irregular cases seems to be new. These results extend and refine the asymptotics found by Jacquier [8] and Lewis [11].

In section 2, we set up the notation and the standing mathematical assumptions used throughout the paper. In section 3, we obtain an asymptotic formula for the implied volatility of an option very far from maturity, with uniform control on the error. In section 4, we

Date: June 8, 2009.

Key words or phrases. Implied volatility, large deviation, saddle-point approximation.

relate the long implied volatility to asymptotics of the cumulant generating function of the logarithm of the underlying stock price. Finally, in section 5 we compute some explicit examples.

2. The mathematical set-up and notation

We consider a market model where the stock price $(S_t)_{t\geq 0}$ is modelled as a non-negative stochastic process[1] defined on a probability space, adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$. Without loss of generality, we normalize by taking $S_0 = 1$, and to simplify the presentation, we assume that the prevailing risk-free interest rate is identically zero.

We make the following assumption:

Assumption 2.1. There exists a locally equivalent measure $\mathbb{P}$ such that S is a local martingale with $\mathbb{P}(S_t > 0) > 0$ for all $t \geq 0$.

Unless otherwise indicated, all probabilistic statements should be interpreted with respect to $\mathbb{P}$, and the the notation $\mathbb{E}$ will denote expectation with respect to $\mathbb{P}$. It is well-known that this market model admits no arbitrage.

We now introduce a family of European call options to this market. We assume that the price $C(t, T, K)$ of a call with strike K and maturity T is given by

$$C(t, T, K) = S_t - \mathbb{E}(S_T \wedge K|\mathcal{F}_t).$$

In this way, the call price processes $(C(t, T, K))_{t\in[0,T]}$ are local martingales, and the market augmented by these options is free of arbitrage. Furthermore, since

$$(S_t - K)^+ \leq C(t, T, K) < S_t$$

implied volatilities are well-defined.

Remark 2.2. We could have equivalently priced put options by the formula

$$P(t, T, K) = \mathbb{E}[(K - S_T)^+|\mathcal{F}_t]$$

and then priced the call options by put-call parity

$$C(t, T, K) = S_t - K + P(t, T, K).$$

It is important to note that the call prices in this framework are not necessarily martingales. Indeed, notice that since S is only assumed to be a non-negative local martingale, and hence a supermartingale, we generally have the inequality

$$\begin{aligned} \mathbb{E}[C(T, T, K)|\mathcal{F}_t] &= \mathbb{E}[(S_T - K)^+|\mathcal{F}_t] \\ &= \mathbb{E}[S_T - S_T \wedge K|\mathcal{F}_t] \\ &\leq S_t - \mathbb{E}(S_T \wedge K|\mathcal{F}_t) \\ &= C(t, T, K), \end{aligned}$$

with equality only if S is a true martingale. We have chosen to price the calls in this, perhaps unorthodox, way because there are models (for instance, the inverse of a three-dimensional Bessel process) in which the inequality

$$\mathbb{E}[(S_T - K)^+] < (S_0 - K)^+$$

[1] Since most of the following results do not depend on whether the time parameter t is discrete or continuous, no distinction is made unless otherwise indicated.

holds for T large enough. In particular, if S is a strictly local martingale, then the implied volatility of a sufficiently long dated option may not be well-defined if we were to price the calls by expectation. See Cox and Hobson's paper [2] for further discussion of the technicalities that arise when S is a strictly local martingale.

We now come to the object of our study:

Definition 2.3. The (time-0) implied total variance for log-moneyness $k = \log(K/S_0)$ and time to maturity $\tau \geq 0$ is the unique non-negative solution $V(k, \tau)$ of the equation

$$\mathrm{BS}(k, V(k, \tau)) = C(0, \tau, e^k)/S_0 = 1 - \mathbb{E}(S_\tau \wedge e^k).$$

We also define the implied volatility by the formula

$$\Sigma(k, \tau) = \sqrt{\frac{V(k, \tau)}{\tau}},$$

but we will find it more convenient to express most of the results in terms of the implied total variance.

Of course, if S is given by the Black–Scholes model, that is, a geometric Brownian motion of the form $S_t = e^{-\sigma_0^2 t/2 + \sigma W_t}$ for a Wiener process W and a constant $\sigma_0 > 0$, then $\mathbb{E}[(S_\tau - e^k)^+] = \mathrm{BS}(k, \tau\ \sigma_0^2)$ and $\Sigma(k, \tau) = \sigma_0$ for all $k \in \mathbb{R}$ and $\tau > 0$.

We make one further assumption:

Assumption 2.4. $S_t \to 0$ almost surely as $t \uparrow \infty$.

Recall that a non-negative supermartingale must converge almost surely to some non-negative random variable S_∞. The assumption that $S_\infty = 0$ is not motivated by no-arbitrage considerations, but can be justified on the following economic grounds: It is easy to see that Assumption 2.4 is equivalent to the reasonable property (exhibited by many models, including Black–Scholes) that $C(0, T, K) \to S_0$ for all $K > 0$ as $T \uparrow \infty$. Alternatively, Assumption 2.4 holds if and only if $V(k, \tau) \to \infty$ as $\tau \uparrow \infty$. See [13] for details.

Remark 2.5. Notice if $\mathbb{P}(S_\infty > 0) > 0$ then $\sup_{T>0} C(0, T, K) < S_0$. In particular, for each $k \in \mathbb{R}$ the implied total variance $V(k, \tau)$ is bounded by a finite constant, and hence the asymptotic formulae given below do not apply.

3. Asymptotic formula

The main result is the following:

Theorem 3.1. *The implied total variance at long maturities is given asymptotically by*

$$V(k, \tau) = -8 \log \mathbb{E}(S_\tau \wedge e^k) - 4 \log[-\log \mathbb{E}(S_\tau \wedge e^k)] + 4k - 4 \log \pi + \varepsilon(k, \tau) \tag{2}$$

where for all $\kappa > 0$ we have the limit

$$\sup_{-\kappa \leq k \leq \kappa} |\varepsilon(k, \tau)| + \sup_{-\kappa \leq k_1 < k_2 \leq \kappa} \frac{|\varepsilon(k_2, \tau) - \varepsilon(k_1, \tau)|}{k_2 - k_1} \to 0 \text{ as } \tau \uparrow \infty.$$

The proof is contained in a series of lemmas:

Lemma 3.2.

$$\log(1 - \mathrm{BS}(k, v)) = -\frac{v}{8} - \frac{1}{2} \log v + \frac{k}{2} + \frac{1}{2} \log(8/\pi) + \varepsilon_1(k, v)$$

where $\lim_{v \uparrow \infty} \sup_{k \in [-\kappa, \kappa]} |\varepsilon_1(k, v)| = 0$ *for all* $\kappa > 0$.

Proof. We have the calculation

$$\begin{aligned} 1-\mathrm{BS}(k,v) &= \Phi[-d_1(k,v)]+e^k\Phi[d_2(k,v)] \\ &= \phi[d_1(k,v)]\{U[d_1(k,v)]+U[-d_2(k,v)]\} \end{aligned}$$

where as usual $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is the standard normal density,

$$d_1(k,v)=-\frac{k}{\sqrt{v}}+\frac{\sqrt{v}}{2} \text{ and } d_2(k,v)=-\frac{k}{\sqrt{v}}-\frac{\sqrt{v}}{2},$$

and $U(x)=\Phi(-x)/\phi(x)$ is Mills' ratio.

Since $|x\ U(x)-1|\le 1/x^2$ for $x>0$, we have $\sqrt{v}U[d_1(k,v)]\to 2$ and $\sqrt{v}U[-d_2(k,v)]\to 2$ as $v\uparrow\infty$, uniformly for $k\in[-\kappa,\kappa]$. The result now follows. $\square$

Lemma 3.3. *Let the inverse of the Black–Scholes function* $IBS:\mathbb{R}\times[0,1)\to\mathbb{R}_+$ *be defined implicitly by*

$$\mathrm{BS}(k,\mathrm{IBS}(k,c))=c.$$

Then

$$\mathrm{IBS}(k,c)=-8\log(1-c)-4\log[-\log(1-c)]+4k-4\log\pi+\varepsilon_2(k,c)$$

where $\lim_{c\uparrow 1}\sup_{k\in[-\kappa,\kappa]}|\varepsilon_2(k,c)|=0$ *for all* $\kappa>0$.

Proof. We first have the crude estimate

$$\mathrm{BS}(k,N)\le\mathrm{BS}(-N/4,N)=\Phi(\sqrt{N}/4)-e^{-N/4}\Phi(-3\sqrt{N}/4)\le 1-e^{-N}$$

for all $k\ge -N/4$ and N large enough. Therefore, $\mathrm{IBS}(k,c)\ge-\log(1-c)$ whenever $k\ge\log(1-c)/4$ and $1-c$ small enough. In particular, $\inf_{k\in[-\kappa,\kappa]}\mathrm{IBS}(k,c)\to\infty$ as $c\uparrow 1$.

Now, Lemma 3.2 tells us that

$$\mathrm{IBS}(k,c)=-8\log(1-c)-4\log[\mathrm{IBS}(k,c)/8]+4k-4\log\pi+8\varepsilon_1(k,\mathrm{IBS}(k,c)).$$

Dividing both sides by $\mathrm{IBS}(k,c)$ and using the limits $\frac{\log v}{v}\to 0$ and $\varepsilon_1(k,v)\to 0$ uniformly, we have

$$\frac{-8\log(1-c)}{\mathrm{IBS}(k,c)}\to 1$$

uniformly in $k\in[-\kappa,\kappa]$. Since

$$\varepsilon_2(k,c)=8\varepsilon_1(k,\mathrm{IBS}(k,c))+4\log\left(\frac{\mathrm{IBS}(k,c)}{-8\log(1-c)}\right)$$

the proof is complete. $\square$

Proof of Theorem 3.1. Lemma 3.3 yields

$$\varepsilon(k,\tau)=\varepsilon_2(k,1-\mathbb{E}(S_\tau\wedge e^k)).$$

But the inequality $\mathbb{E}(S_\tau\wedge e^k)\le\mathbb{E}(S_\tau\wedge e^\kappa)$, the assumption $S_t\to 0$ almost surely, and the dominated convergence theorem, together imply $\varepsilon(k,\tau)\to 0$ uniformly in $k\in[-\kappa,\kappa]$.

It remains to show $\sup_{-\kappa\le k_1<k_2<\kappa}\frac{|\varepsilon(k_2,\tau)-\varepsilon(k_1,\tau)|}{k_2-k_2}\to 0$. To this end, let $c(k,\tau)=1-\mathbb{E}(S_\tau\wedge e^k)$ and note that for each $\tau>0$, the function $k\to c(k,\tau)$ has both left- and right-derivatives at each point, given by

$$D_-c(k,\tau)=-e^k\mathbb{P}(S_\tau>e^k) \text{ and } D_+c(k,\tau)=-e^k\mathbb{P}(S_\tau\ge e^k)$$

respectively. Therefore, computing the derivatives implicitly from the definition of implied total variance, we have the expression

$$D_-V(k,\tau) = 2\sqrt{V(k,\tau)}\left(\frac{\Phi\{d_2[k,V(k,\tau)]\} - \mathbb{P}(S_\tau \geq e^k)}{\phi\{d_2[k,V(k,\tau)]\}}\right).$$

Of course we have a similar expression for $D_+V(k,\tau)$. On the other hand, differentiating equation (2) yields

$$D_-V(k,\tau) = 4 - 8e^k\frac{\mathbb{P}(S_\tau \geq e^k)}{\mathbb{E}(S_\tau \wedge e^k)}\left(1 + \frac{1}{2\log(\mathbb{E}(S_\tau \wedge e^k))}\right) + D_-\varepsilon(k,\tau).$$

The result follows from noting the following the uniform limits

$$\frac{\sqrt{v}\Phi[d_2(k,v)]}{\phi[d_2(v,k)]} \to 2$$

and

$$\frac{e^k\Phi\{d_2[k,V(k,\tau)]\}}{\mathbb{E}(S_\tau \wedge e^k)} \to \frac{1}{2},$$

and the inequality $\frac{\mathbb{P}(S_\tau \geq e^k)}{\mathbb{E}(S_\tau \wedge e^k)} \leq 1$. □

A first corollary of Theorem 3.1 is the leading order term of the long implied total variance:

Corollary 3.4.

$$\sup_{k\in[-\kappa,\kappa]}\left|\frac{V(k,\tau)}{-8\log\mathbb{E}(S_\tau \wedge 1)} - 1\right| \to 0$$

as $\tau \uparrow \infty$ for all $\kappa > 0$.

Proof. This follows quickly from Theorem 3.1 upon noting the inequality

$$e^{-\kappa} \leq \log\frac{S_\tau \wedge e^k}{S_\tau \wedge 1} \leq e^\kappa$$

for all $k \in [-\kappa,\kappa]$. □

The above corollary shows that the implied volatility surface flattens at long maturities. The next corollary gives an asymptotic formula for the skew:

Corollary 3.5.

$$D_\pm V(k,\tau) = 4\left(\frac{\mathbb{E}[S_\tau \mathbb{1}_{\{S_\tau<e^k\}} - e^k\mathbb{1}_{\{S_\tau>e^k\}}] \pm e^k\mathbb{P}(S_\tau = e^k)}{\mathbb{E}(S_\tau \wedge e^k)}\right) + \varepsilon_\pm(k,\tau)$$

where $\sup_{k\in[-\kappa,\kappa]}|\varepsilon_\pm(k,\tau)| \to 0$ as $\tau \uparrow \infty$ for all $\kappa > 0$. In particular, we have the following bound

$$\limsup_{\tau\uparrow\infty}\max\{|D_-V(k,\tau)|, |D_+V(k,\tau)|\} \leq 4. \qquad (3)$$

Remark 3.6. A consequence of the corollary is that the implied volatility flattens in the precise sense

$$\sup_{-\kappa\leq k_1<k_2<\kappa}|\Sigma(k_1,\tau) - \Sigma(k_2,\tau)| \to 0$$

as $\tau \uparrow \infty$, where $\Sigma(k,\tau) = \sqrt{V(k,\tau)/\tau}$. This is a model independent result, and is not a consequence, for instance, of some notion of ergodicity. This flattening phenomenon has been noticed before: Gatheral [6] has shown that the gradient of the implied volatility, if it exists, decays pointwise like $1/\tau$. See also the paper of Lee [10]. Equation 3 was established in [13], where the constant 4 was shown to be sharp.

4. The cumulant generating function

From Corollary 3.4 and the simple inequality $S \wedge 1 \leq S^p$ which holds for all $0 \leq p \leq 1$ and $S \geq 0$, we obtain the bound

$$\liminf_{\tau \uparrow \infty} \frac{V(k,\tau)}{-8 \inf_{0 \leq p \leq 1} \log \mathbb{E}(S_\tau^p)} \geq 1.$$

In this section, we explore conditions under which the above bound can be strengthened to equality. Indeed, we show how the behavior of the long implied total variance is related to the cumulant generation function of the log stock price.

Definition 4.1. The moment generating function $M_\tau : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$M_\tau(p) = \mathbb{E}(S_\tau^p \mathbb{1}_{\{S_\tau > 0\}})$$

for all $\tau \geq 0$. The cumulant generating function $\Lambda_\tau : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Lambda_\tau(p) = \log M_\tau(p)$$

for all $\tau \geq 0$.

Note that $M_\tau(0) = \mathbb{P}(S_\tau > 0)$ and $M_\tau(1) = \mathbb{E}(S_\tau)$ so by Assumption 2.1 and Hölder's inequality $M_\tau(p)$ takes values in $(0,1]$ for all $p \in [0,1]$. In particular, M_τ is finite-valued on the open interval, and hence can be extended analytically to the vertical strip $\{p + iq : 0 < p < 1, q \in \mathbb{R}\}$ in the complex plane. Furthermore, there exists a neighborhood $\{p + iq : 0 < p < 1, |q| < \varepsilon\}$ on which this extension of M_τ is non-zero and so Λ_τ can be extended to principal branch of the logarithm of M_τ.

Roughly speaking, the asymptotics of the implied total variance depends on the location of the minimum of the convex function Λ_τ. Essentially, three types of behavior are possible, which we call the regular, borderline, and irregular cases below. Note that if the stock price S is strictly positive and a true martingale, then the regular case covers most of the interesting examples. However, there are many popular models in which either S hits zero in finite time with positive probability, or is a strictly local martingale, or both, and the borderline and irregular cases become important.

4.1. The regular case.

In this subsection, we now make the following assumption:

Assumption 4.2. There exists a $0 < p^* < 1$ and a positive increasing function C with $C(\tau) \uparrow \infty$ such that

$$\Lambda_\tau\left(p^* + i\frac{\theta}{C(\tau)}\right) - \Lambda_\tau(p^*) \to -\theta^2/2$$

as $\tau \uparrow \infty$ for all real θ.

Remark 4.3. The above assumption can be verified in practice as follows: Suppose that there exists a strictly convex function $\bar{\Lambda}$ on $(0,1)$ which can be extended to an analytic function on $\{0 < p < 1, |q| < \varepsilon\}$ for some $\varepsilon > 0$ such that

$$\frac{1}{\tau}\Lambda_\tau(p + iq) \to \bar{\Lambda}(p + iq)$$

as $\tau \uparrow \infty$ for all $0 < p < 1$ and $|q| < \varepsilon$. Furthermore, suppose $p^* \in (0,1)$ is the unique minimizer of $\bar{\Lambda}$, so that in particular, we have $\bar{\Lambda}'(p^*) = 0$.

Letting $\bar{\Lambda}''(p^*) = a^2 > 0$, and

$$\frac{\Lambda_\tau(p^* + iq)}{\tau} - \bar{\Lambda}(p^* + iq) = \beta_\tau(q)$$

we have

$$\Lambda_\tau\left(p^* + i\frac{\theta}{a\sqrt{\tau}}\right) - \Lambda_\tau(p^*) = \tau\left[\bar{\Lambda}\left(p^* + i\frac{\theta}{a\sqrt{\tau}}\right) - \bar{\Lambda}(p^*)\right] + \theta\frac{\sqrt{\tau}}{a}\beta_\tau'\left(\frac{\hat{\theta}}{a\sqrt{\tau}}\right)$$

for some $|\hat{\theta}| \le |\theta|$ by the mean-value theorem. Assuming $\sqrt{\tau}\beta_\tau'(\theta/\sqrt{\tau}) \to 0$ uniformly in θ on compacts, Assumption 4.2 is satisfied with $C(\tau) = a\sqrt{\tau}$ since the first bracketed term above converges to $-\theta^2/2$ by Taylor's theorem.

In this section we will let

$$\phi_\tau(\theta) = \frac{1}{\sqrt{2\pi}}\frac{M_\tau\left(p^* + i\frac{\theta}{C(\tau)}\right)}{M_\tau(p^*)}.$$

By assumption, $\phi_\tau(\theta) \to \phi(\theta) = \frac{1}{\sqrt{2\pi}}e^{-\theta^2/2}$, the standard normal density, for each $\theta \in \mathbb{R}$.

Before stating and proving the main result of this section, we need a further assumption. We have assumed that $S_t \to 0$ almost surely, which implies $\Lambda_\tau(p) \to -\infty$ for each $p \in (0,1)$, since the process $(S_t^p)_{t\ge 0}$ is uniformly integrable. In this section we make the following stronger technical assumption:

Assumption 4.4.

$$\frac{C(\tau)}{\Lambda_\tau(p^*)} \to 0$$

as $\tau \uparrow \infty$.

Remark 4.5. In most examples $C(\tau) \sim a\sqrt{\tau}$ and $\Lambda_\tau(p^*) \sim b\tau$ for constants $a, b > 0$, so that Assumption 4.4 is satisfied.

Our first result gives the leading order of long implied total variance with no other assumption on the cumulant generating function. The proof of this result is similar to the proof of the classical Cramér large deviation principle; see Chapter 5.11 of Grimmett and Stirzaker's book [7], for instance. The following result appears (without the uniform control of the error term) in [14] as an application of the Cramér theorem in the case where $\log S$ has independent stationary increments. This result also appears in chaper 6 of Lewis's book [11] for stochatic volatility models under a heuristic asymptotic factorization assumption on the function M_τ.

Theorem 4.6.

$$\sup_{k\in[-\kappa,\kappa]}\left|\frac{V(k,\tau)}{-8\Lambda_\tau(p^*)} - 1\right| \to 0$$

Proof. By the inequality $S_\tau \wedge 1 \le S_\tau^p$ which holds for all $0 < p < 1$, we have the upper bound

$$\log\mathbb{E}(S_\tau \wedge 1) \le \Lambda(p^*).$$

Now for each $\tau \geq 0$ let X_τ be a random variable such that $\mathbb{E}(e^{i\theta X_\tau}) = \frac{M_\tau(p^*+i\theta/C(\tau))}{M_\tau(p^*)}$. Note that by Assumptions 4.2 the distribution of X_τ converges to that of a standard normal random variable. For each $b > 0$ we have

$$\begin{aligned}
\log \mathbb{E}(S_\tau \wedge 1) &= \Lambda_\tau(p^*) + \log \mathbb{E}(e^{-p^* C(\tau) X_\tau} e^{C(\tau)X_\tau} \wedge 1) \\
&\geq \Lambda_\tau(p^*) + \log \mathbb{E}[e^{-p^* C(\tau) X_\tau} (e^{C(\tau)X_\tau} \wedge 1) \mathbb{1}_{\{X_\tau < -b\Lambda_\tau(p^*)/C(\tau)\}}] \\
&\geq \Lambda_\tau(p^*)(1 + p^* b) + \log \mathbb{E}[(e^{C(\tau)X_\tau} \wedge 1) \mathbb{1}_{\{X_\tau < -b\Lambda_\tau(p^*)/C(\tau)\}}]
\end{aligned}$$

Since $(e^{C(\tau)X_\tau} \wedge 1)\mathbb{1}_{\{X_\tau < -b\Lambda_\tau(p^*)/C(\tau)\}}$ is bounded and converges in distribution to a Bernoulli random variable with parameter $1/2$ by Assumpton 4.4, we have

$$\limsup_{\tau \uparrow \infty} \left| \frac{\log \mathbb{E}(S_\tau \wedge 1)}{\Lambda_\tau(p^*)} - 1 \right| \leq b\ p^*$$

Now let $b \downarrow 0$, and apply Corollary 3.4. $\square$

If we supplement this pointwise convergence $\phi_\tau \to \phi$ with some sort of uniform integrability condition, we can get good estimates of the long implied volatility.

Theorem 4.7. *If*

$$\int_{-\infty}^{\infty} \frac{|\phi_\tau(\theta)|}{1 + \theta^2/C(\tau)^2} d\theta \to 1$$

then

$$V(k,\tau) = -8\Lambda_\tau(p^*) + 4k(2p^* - 1) + 8 \log\left(\frac{C(\tau)p^*(1-p^*)}{\sqrt{-\Lambda_\tau(p^*)/2}} \right) + \delta(k,\tau)$$

where $\sup_{k \in [-\kappa,\kappa]} |\delta(k,\tau)| \to 0$ *as* $\tau \uparrow \infty$ *for each* $\kappa > 0$.

Remark 4.8. Based on the theorem above, one might guess that the implied total variance is always approximately affine in the log-moneyness for long maturities. We will see in Section 5 that this conjecture is false: The symmetric binomial model is in the regular case with $p^* = 1/2$, but the long implied total variance is not affine. Of course, for this example, the integrability condition in Theorem 4.7 fails to hold.

We begin with a lemma:

Lemma 4.9. *The following identity is valid for all* $p \in (0,1)$, $k \in \mathbb{R}$, *and* $\tau \geq 0$*:*

$$\mathbb{E}(S_\tau \wedge e^k) = \frac{e^{k(1-p)}}{2\pi} \int_{-\infty}^{\infty} \frac{M_\tau(p+iy)e^{-iky}}{(p+iy)(1-p-iy)} dy.$$

Proof. It is straightforward to show by contour integration the formula

$$\int_{-\infty}^{\infty} \frac{e^{izy}}{(p+iy)(1-p-iy)} dy = 2\pi \begin{cases} e^{-pz} & \text{if } z \geq 0 \\ e^{(1-p)z} & \text{if } z < 0 \end{cases}$$

holds for all real z. Furthermore, since

$$\int_{-\infty}^{\infty} \mathbb{E}\left| \frac{e^{(p+iy)\log S_\tau - iky} \mathbb{1}_{\{S_\tau > 0\}}}{(p+iy)(1-p-iy)} \right| dy \leq M_\tau(p) \int_{-\infty}^{\infty} \frac{dy}{p(1-p) + y^2} < \infty$$

Fubini's theorem implies

$$\begin{aligned}
\frac{e^{k(1-p)}}{2\pi}\int_{-\infty}^{\infty}\frac{M_\tau(p+iy)e^{-iky}}{(p+iy)(1-p-iy)}dy &= \frac{1}{2\pi}\mathbb{E}\left(\int_{-\infty}^{\infty}\frac{e^{p\log S_\tau+k(1-p)+iy(\log S_\tau-k)}\mathbb{1}_{\{S_\tau>0\}}}{(p+iy)(1-p-iy)}dy\right)\\
&= \mathbb{E}(S_\tau\wedge e^k)
\end{aligned}$$

□

Remark 4.10. Notice that for all real z contour integration yields

$$\int_{-\infty}^{\infty}\frac{e^{izy}}{(p+iy)(1-p-iy)}dy = 2\pi\begin{cases}-e^{-pz}(e^z-1)^+ & \text{if } p>1\\ e^{-pz}(1-e^z)^+ & \text{if } p<0.\end{cases}$$

In particular, the above proof cannot be valid if $p^*>1$ or $p^*<0$. We will see shortly that, in fact, genuinely different behavior arises in these cases.

Proof of Theorem 4.7. By the lemma and the change of variables $\theta = yC(\tau)$, we have

$$\mathbb{E}(S_\tau\wedge e^k) = \frac{e^{k(1-p^*)+\Lambda_\tau(p^*)}}{p^*(1-p^*)C(\tau)\sqrt{2\pi}}\int_{-\infty}^{\infty}f_\tau(\theta,k)d\theta$$

where

$$f_\tau(\theta,k) = \frac{\phi_\tau(\theta)e^{-ik\theta/C(\tau)}}{[1+i\theta/(p^*C(\tau))][1-i\theta/((1-p^*)C(\tau))]}.$$

Define auxilliary functions g_τ and h_τ by

$$g_\tau(\theta) = \frac{\phi_\tau(\theta)}{1+\theta^2/C(\tau)^2}$$

and

$$h_\tau(\theta) = \frac{e^{ik\theta/C(\tau)}f_\tau(\theta,k)}{g_\tau(\theta)} = \frac{1+\theta^2/C(\tau)^2}{[1+i\theta/(p^*C(\tau))][1-i\theta/((1-p^*)C(\tau))]}.$$

First note that $g_\tau\to\phi$ and $h_\tau\to 1$ pointwise. We have the computation

$$\begin{aligned}
|f_\tau(\theta,k)-\phi(\theta)| &= |g_\tau(\theta)e^{-ik\theta/C(\tau)}h_\tau(\theta)-\phi(\theta)|\\
&\le |g_\tau(\theta)-\phi(\theta)||h_\tau(\theta)|+\phi(\theta)|e^{-ik\theta/C(\tau)}-1|+\phi(\theta)|h_\tau(\theta)-1|\\
&\le |g_\tau(\theta)-\phi(\theta)|+\phi(\theta)\left(\frac{\kappa|\theta|}{C(\tau)}\wedge 2\right)+\phi(\theta)|h_\tau(\theta)-1|,
\end{aligned}$$

for all $k\in[-\kappa,\kappa]$, where we have used the inequalities $|h_\tau(\theta)|\le 1$ and $|e^{ix}-1|\le|x|\wedge 2$. All three of the terms above vanish pointwise as $\tau\uparrow\infty$, and the second and third are dominated by integrable functions. Furthermore, since

$$\int_{-\infty}^{\infty}|g_\tau(\theta)|d\theta\to 1=\int_{-\infty}^{\infty}\phi(\theta)d\theta$$

by hypothesis, and we have the convergence

$$\int_{-\infty}^{\infty}|g_\tau(\theta)-\phi(\theta)|d\theta\to 0$$

by Scheffé theorem.

If we write $\int_{-\infty}^{\infty} f_\tau(\theta,k)d\theta = 1+\delta_1(k,\tau)$, the above computation shows

$$\sup_{k\in[-\kappa,\kappa]} |\delta_1(k,\tau)| \;\leq\; \int_{-\infty}^{\infty} \sup_{k\in[-\kappa,\kappa]} |f_\tau(\theta,k)-\phi(\theta)|d\theta \to 0$$

by the dominated convergence theorem.

Now, applying Theorem 3.1 we have

$$\begin{aligned}\delta(k,\tau) \;&=\; \varepsilon(k,\tau) - 8\log(1+\delta_1(k,\tau)) \\ &\quad -4\log\left(1+\frac{k(1-p^*)-\log[p^*(1-p^*)C(\tau)\sqrt{2\pi}/(1+\delta_1(k,t))]}{\Lambda_\tau(p^*)}\right)\end{aligned}$$

and the conclusion follows since $\log(C(\tau))/\Lambda_\tau(p^*) \to 0$ by Assumption 4.4. □

Remark 4.11. The asymptotic formula appearing in Theorem 4.7 is essentially Equation (3.8) in Chapter 6 of Lewis's book [11] under a suggestive, if not completely rigorous, assumption of the form $M_\tau(p) \approx e^{-\lambda(p)\tau}u(p)$. Under a similar assumption, Jacquier [8] showed that Lewis's formula as stated cannot be correct as the constant term is missing. The main contribution of this paper is both the explicit sufficient condition on the function Λ_τ under which the formula holds and the uniform control on the error term.

We can now study the skew. If the distribution of S_τ is continuous, the smile is differentiable, and intuitively, $DV(k,\tau) \to 4(2p^*-1)$ by Theorem 4.7. The following theorem makes this intuition precise.

Theorem 4.12. *Suppose the distribution of S_τ is continuous for all $\tau \geq 0$. If*

$$\int_{-\infty}^{\infty} \frac{|\phi_\tau(\theta)|}{1+|\theta|/C(\tau)}d\theta \to 1$$

then

$$DV(k,\tau) = 4(2p^*-1)+\delta'(k,\tau)$$

where $\sup_{k\in[-\kappa,\kappa]}|\delta'(k,\tau)| \to 0$ *as* $\tau\uparrow\infty$ *for each* $\kappa>0$.

Remark 4.13. Note that if the asymptotic formula $DV(k,\tau)\to 4(2p-1)$ holds for some p, then by Corollary 3.5 we must have $0\leq p\leq 1$.

The proof of this result closely follows the proof of Theorem 4.7. Hence, we only outline the argument. Again, we begin with a lemma:

Lemma 4.14. *The following identity is valid for all $p\in(0,1)$, $k\in\mathbb{R}$, and $\tau\geq 0$:*

$$\mathbb{E}[S_\tau \mathbb{1}_{\{S_\tau<e^k\}} - e^k\mathbb{1}_{\{S_\tau>e^k\}}] = (2p-1)\mathbb{E}(S_\tau\wedge e^k) - \frac{e^{k(1-p)}}{\pi i}\lim_{N\to\infty}\int_{-N}^{N}\frac{yM_\tau(p+iy)e^{-iky}}{(p+iy)(1-p-iy)}dy.$$

Proof. Again, contour integration and Fubini's theorem proves the lemma. □

Proof of Theorem 4.12. By the lemma and the proof of Theorem 4.7, we need only show

$$\int_{-\infty}^{\infty}\frac{\phi_\tau(\theta)e^{-ik\theta/C(\tau)}\theta/C(\tau)}{[p^*+i\theta/C(\tau)][(1-p^*)-i\theta/C(\tau)]}d\theta \to 0$$

uniformly in k. The given condition is sufficient for this convergence. □

4.2. **The borderline cases.** In this section we tackle the cases where the saddle-point method outlined above still yields asymptotics. However, since these cases sit on the borderline between the regular and irregular cases, the asymptotic formulae are different. The following assumption will be in force throughout this section.

Assumption 4.15. There exists a $p^* \in \{0,1\}$ and an increasing function C with the following properties: there exists a neighborhood of p^* on which Λ_τ is finite-valued for all $\tau \geq 0$, and hence the mapping $y \mapsto \Lambda_\tau(p^* + yi)$ is well-defined and smooth. The function is such that $C(\tau) \to \infty$ and

$$\Lambda_\tau\left(p^* + i\frac{\theta}{C(\tau)}\right) - \Lambda_\tau(p^*) \to -\theta^2/2$$

for all $\theta \in \mathbb{R}$.

Curiously, for the borderline cases (and with the irregular cases treated below) no extra uniform integrability condition is required to obtain the full asymptotics:

Theorem 4.16. *If* $p^* = 1$ *then*

$$V(k,\tau) = -8\Lambda_\tau(1) - 4\log[-\Lambda_\tau(1)] + 4k - 4\log(\pi/4) + \delta(k,\tau)$$

and if $p^* = 0$ *then*

$$V(k,\tau) = -8\Lambda_\tau(0) - 4\log[-\Lambda_\tau(0)] - 4k - 4\log(\pi/4) + \delta(k,\tau)$$

where in both cases $\sup_{k\in[-\kappa,\kappa]} |\delta(k,\tau)| \to 0$ *for all* $\kappa > 0$.

Proof. Let X_τ be a random variable with characteristic function $\mathbb{E}(e^{i\theta X_\tau}) = \frac{M_\tau\left(p^*+i\frac{\theta}{C(\tau)}\right)}{M_\tau(p^*)}$.

If $p^* = 1$, then

$$\mathbb{E}(S_\tau \wedge e^k) = \mathbb{E}(S_\tau)\mathbb{E}(1 \wedge e^{k-C(\tau)X_\tau})$$

But $1 \wedge e^{k-C(\tau)X_\tau}$ is bounded and converges in distribution to a Bernoulli random variable with mean $1/2$. Furthermore, since

$$\mathbb{E}(1 \wedge e^{-\kappa-C(\tau)X_\tau}) \leq \mathbb{E}(1 \wedge e^{k-C(\tau)X_\tau}) \leq \mathbb{E}(1 \wedge e^{\kappa-C(\tau)X_\tau})$$

we have

$$\frac{\mathbb{E}(S_\tau \wedge e^k)}{\mathbb{E}S_\tau} \to 1/2$$

uniformly for $k \in [-\kappa,\kappa]$, and the conclusion follows from Theorem 3.1.

Similarly, if $p^* = 0$ then

$$\mathbb{E}(S_\tau \wedge e^k) = \mathbb{P}(S_\tau > 0)\mathbb{E}(e^{C(\tau)X_\tau} \wedge e^k)$$

and the conclusion follows as before. □

We can apply Corollary 3.5 to obtain the asymptotic skew, again with no further integrability assumption.

Theorem 4.17. *Assume the distribution of* S_τ *is continuous. If* $p^* = 1$ *then*

$$DV(k,\tau) = 4 + \delta'(k,\tau)$$

and if $p^* = 0$ *then*

$$DV(k,\tau) = -4 + \delta'(k,\tau),$$

where $\sup_{k\in[-\kappa,\kappa]} |\delta'(k,\tau)| \to 0$ *for all* $\kappa > 0$.

Proof. We must estimate the quantity

$$\frac{\mathbb{E}[S_\tau \mathbb{1}_{\{S_\tau < e^k\}} - e^k \mathbb{1}_{\{S_\tau > e^k\}}]}{\mathbb{E}(S_\tau \wedge e^k)} = 1 - 2\frac{e^k \mathbb{P}(S_\tau > e^k)}{\mathbb{E}(S_\tau \wedge e^k)} = 2\frac{\mathbb{E}(S_\tau \mathbb{1}_{\{S_\tau < e^k\}})}{\mathbb{E}(S_\tau \wedge e^k)} - 1$$

As in the proof above, let X_τ have characteristic function $\mathbb{E}(e^{i\theta X_\tau}) = \frac{M_\tau\left(p^* + i\frac{\theta}{C(\tau)}\right)}{M_\tau(p^*)}$.

If $p^* = 1$ then

$$\frac{\mathbb{E}(S_\tau \mathbb{1}_{\{S_\tau < e^k\}})}{\mathbb{E}(S_\tau)} = \mathbb{P}(C(\tau)X_\tau < k) \to 1/2$$

and if $p^* = 0$ then

$$\frac{\mathbb{P}(S_\tau > e^k)}{\mathbb{P}(S_\tau > 0)} = \mathbb{P}(C(\tau)X_\tau > k) \to 1/2.$$

In both cases the convergence is uniform in $k \in [-\kappa, \kappa]$, proving the claim. $\square$

4.3. **The irregular cases.** Now we deal with the irregular cases. As before, we split them into subcases.

Assumption 4.18. There exists a p^* such that either

(1) $p^* > 1$ and $\Lambda_\tau(p^*) - \Lambda_\tau(1) \to -\infty$, or
(2) $p^* < 0$ and $\Lambda_\tau(p^*) - \Lambda_\tau(0) \to -\infty$.

Theorem 4.19. *If $p^* > 1$ then*

$$V(\tau, k) = -8\Lambda_\tau(1) - 4\log[-\Lambda_\tau(1)] + 4k - 4\log\pi + \delta(k, \tau),$$

and if $p^* < 0$ then

$$V(\tau, k) = -8\Lambda_\tau(0) - 4\log[-\Lambda_\tau(0)] - 4k - 4\log\pi + \delta(k, \tau).$$

In both cases, $\sup_{k\in[-\kappa,\kappa]} |\delta(k, \tau)| \to 0$ for all $\kappa > 0$.

Proof. If $p^* > 1$, we have the inequality

$$\begin{aligned}
\mathbb{E}(S_\tau) &\geq \mathbb{E}(S_\tau \wedge e^k) \\
&\geq \mathbb{E}(S_\tau \wedge e^{-\kappa}) \\
&= \mathbb{E}[S_\tau - (S_\tau - e^{-\kappa})^+] \\
&\geq \mathbb{E}(S_\tau) - e^{(p^*-1)\kappa}\mathbb{E}(S_\tau^{p^*})
\end{aligned}$$

for all $k \in [-\kappa, \kappa]$, where we have used the simple inequality $(a - b)^+ \leq a^p b^{1-p}$ which holds for all $a, b > 0$ and $p > 1$. Hence, if $p^* > 1$, we have the bound

$$\left|\frac{\mathbb{E}(S_\tau \wedge e^k)}{\mathbb{E}(S_\tau)} - 1\right| \leq e^{(p^*-1)\kappa + \Lambda_\tau(p^*) - \Lambda_\tau(1)} \to 0.$$

Similarly, if $p^* < 0$ we have have the inequality

$$\begin{aligned}
e^k\mathbb{P}(S_\tau > 0) &\geq \mathbb{E}(S_\tau \wedge e^k) \\
&= \mathbb{E}[e^k \mathbb{1}_{\{S\tau > 0\}} - (e^k - S_\tau)^+ \mathbb{1}_{\{S\tau > 0\}}] \\
&\geq \mathbb{E}[e^k \mathbb{1}_{\{S\tau > 0\}} - (e^\kappa - S_\tau)^+ \mathbb{1}_{\{S\tau > 0\}}] \\
&\geq e^k\mathbb{P}(S_\tau > 0) - e^{(1-p^*)\kappa}\mathbb{E}(S_\tau^{p^*} \mathbb{1}_{\{S_\tau > 0\}})
\end{aligned}$$

for all $k \in [-\kappa, \kappa]$, where we have used the inequality $(b-a)^+ \leq a^p b^{1-p}$ which holds for all $a, b > 0$ and $p < 0$. Thus, if $p^* < 0$ we have the corresponding bound

$$\left| \frac{\mathbb{E}(S_\tau \wedge e^k)}{\mathbb{P}(S_\tau > 0)} - e^k \right| \leq e^{(1-p^*)\kappa + \Lambda_\tau(p^*) - \Lambda_\tau(0)} \to 0.$$

The result now follows from Theorem 3.1. $\square$

In fact, we also have convergence of the skews:

Theorem 4.20. *Assume that S_τ has a continuous distribution for all $\tau \geq 0$. If $p^* > 1$ then*

$$DV(k, \tau) = 4 + \delta'(k, \tau)$$

and if $p^* < 0$ then

$$DV(k, \tau) = -4 + \delta'(k, \tau)$$

where $\sup_{k \in [-\kappa, \kappa]} |\delta'(k, \tau)| \to 0$ for all $\kappa > 0$.

Proof. Note that if $p^* > 1$ then by Chebychev's inequality

$$0 \leq e^k \mathbb{P}(S_\tau > e^k) \leq e^{k(1-p^*)} \mathbb{E}(S_\tau^{p^*}) \leq e^{\kappa(p^*-1) + \Lambda_\tau(p^*)}$$

for all $k \in [-\kappa, \kappa]$. Similarly, if $p^* < 0$ we have the

$$0 \leq \mathbb{E}(S_\tau \mathbb{1}_{\{S_\tau < e^k\}}) \leq e^{\kappa(p^*-1) + \Lambda_\tau(p^*)}.$$

The theorem now follows from Corollary 3.5 and the estimates in the proof of Theorem 4.19. $\square$

5. Examples

In this section we consider some examples to illustrate Theorem 4.7.

5.1. Models with stationary independent increments.

The easiest case to analyse is when $S_t = e^{X_t}$ where X has stationary independent increments such that S is a martingale. If time is continuous, we take X to be a Lévy process but our discussion is also valid for discrete time. When X has stationary independent increments, the cumulant generating function has the nice form

$$\Lambda_\tau(p) = \tau \, \Lambda_1(p)$$

and $\Lambda_1(0) = \Lambda_1(1) = 0$. The function Λ_1 has a unique global minimum at some point $p^* \in (0, 1)$ at which $\Lambda_1'(p^*) = 0$.

Letting $a^2 = \Lambda_1''(p^*) > 0$ we have by Taylor's theorem

$$\tau \left[\Lambda_1 \left(p^* + i \frac{\theta}{a\sqrt{\tau}} \right) - \Lambda_1(p^*) \right] \to -\theta^2/2,$$

and Assumption 4.2 is satified with $C(\tau) = a\sqrt{\tau}$. Clearly Assumption 4.4 is also satisfied. In particular, we have the leading order behavior of the implied volatility is given by

$$\sup_{k \in [-\kappa, \kappa]} \left| \Sigma(k, \tau)^2 + 8 \min_{p \in \mathbb{R}} \log \mathbb{E}(S_1^p) \right| \to 0.$$

Example 5.1 (The Black–Scholes model). This example is simply a reality-check. Let $S_t = e^{-\sigma_0^2 t/2+\sigma_0 W_t}$ for a Brownian motion W. The cumulant generating function is

$$\Lambda_1(p) = \frac{1}{2}\sigma_0^2 p(p-1)$$

which is minimized at $p^* = 1/2$ and we may take $C(\tau) = \sigma_0\sqrt{\tau}$ in Assumption 4.2. Now $\phi_\tau = \phi$ for all $\tau > 0$, and in particular, $\sup_{\tau>0}\phi_\tau$ is integrable. Hence both Theorems 4.7 and 4.12 apply, and we have

$$V(k,\tau) = \sigma_0^2\tau + \delta(\tau,k),$$

and

$$DV(k,\tau) = \delta'(\tau,k).$$

Of course for this example, $\delta(\tau,k) = 0$ identically.

5.1.1. *A sufficient condition.* Let the stock price be given by $S_t = e^{X_t}$ where X has independent stationary increments. As usual, let $M_1(p) = \mathbb{E}(e^{pX_1})$, $\Lambda_1(p) = \log M_1(p)$, and $p^* = \text{argmin}\Lambda_1$. We now exhibit a sufficient condition for Theorem 4.7 to hold:

Theorem 5.2. *Suppose that is such that for some $b > 0$ the inequality*

$$|M_1(p^* + iq)| \le e^{-b\ (q^2\wedge 1)} M_1(p^*)$$

holds for all $q \in \mathbb{R}$. Then the asymptotic implied total variance is given by

$$V(k,\tau) = -8\Lambda_1(p^*)\tau + 4k(2p^*-1) + 4\log\left(\frac{2\Lambda_1''(p^*)[p^*(1-p^*)]^2}{-\Lambda_1(p^*)}\right) + \delta(k,\tau).$$

where $\sup_{k\in[-\kappa,\kappa]}|\delta(k,\tau)| \to 0$ for all $\kappa > 0$.

Proof. Let $a = \sqrt{\Lambda_1(p^*)}$. Since

$$\begin{aligned}\int_{|\theta|>a\sqrt{\tau}} \frac{|\phi_\tau(\theta)|}{1+\theta^2/\tau}d\theta &\le \frac{1}{\sqrt{2\pi}}\int_{|\theta|>a\sqrt{\tau}}\frac{e^{-b\tau}}{1+\theta^2/\tau}d\theta \\ &= \sqrt{\frac{2\tau}{\pi}}e^{-b\tau}\tan^{-1}(1/a) \to 0\end{aligned}$$

and $\mathbb{1}_{\{|\theta|\le a\sqrt{\tau}\}}|\phi_t(\theta)| < \frac{1}{\sqrt{2\pi}}e^{-b\theta^2/a^2}$ is integrable and converges pointwise to $\phi(\theta)$ we have

$$\int_{-\infty}^{\infty}\frac{|\phi_\tau(\theta)|}{1+\theta^2/\tau}d\theta \to 1$$

by the dominated convergence theorem. Hence Theorem 4.7 applies. □

Remark 5.3. For models satisfying the hypothesis of Theorem 5.2, the long implied total variance is approximately affine in both the log-moneyness k and the time to maturity τ:

$$V(k,\tau) \approx A\tau + Bk + C.$$

In chapter 5 of the book by Fouque, Papanicolaou, and Sircar [5], it is observed, in the context of a fast-mean reverting stochastic volatility model, that such affine structure could be exploited for model calibration. Indeed, one need only regress observed values of $V(k,\tau)$ against (k,τ) for small k and large τ to obtain estimates of A, B, and C. For a model with three parameters, such as the variance gamma model studied below, the values of A, B, and C can be inverted to yield the model parameters.

Example 5.4 (Subordinated Brownian motion and the variance gamma process). For an example of a Lévy process which satisfies the sufficient condition given by Theorem 5.2, consider the following construction: Let Y be a subordinator with characteristic function

$$\mathbb{E}(e^{iqY_t}) = e^{iqta+\int_{(0,\infty)} t(e^{iqx}-1)\mu(dx)}$$

where $a \geq 0$ and μ is a measure such that

$$\int_{(0,\infty)} (x \wedge 1)\mu(dx) < \infty.$$

Now let W be an independent Brownian motion and define a new Lévy process X by

$$X_t = \sigma W(Y_t) + \Theta Y_t + mt$$

for real constants σ, Θ such that m defined by

$$m = -a(\Theta + \sigma^2/2) - \int_{(0,\infty)} (e^{(\Theta+\sigma^2/2)x} - 1)\mu(dx)$$

is finite. Then by construction, the process $S_t = e^{X_t}$ defines a martingale.

The cumulant generating function of X_1 is given by

$$\Lambda_1(p) = a\sigma^2 \frac{p(p-1)}{2} + \int_{(0,\infty)} (e^{(p\Theta+p^2\sigma^2/2)x} - e^{(\Theta+\sigma^2/2)x})\mu(dx)$$

and hence

$$\begin{aligned}
\Re\Lambda_1(p+iq) - \Lambda_1(p) &= -a\sigma^2\frac{q^2}{2} + \int_{(0,\infty)} e^{(p\Theta+p^2\sigma^2/2)x}\left(\cos[(\Theta+\sigma^2 p)qx]e^{-q^2\sigma^2 x/2} - 1\right)\mu(dx) \\
&\leq -a\sigma^2\frac{q^2}{2} - q^2 \wedge 1 \int_{(0,\infty)} e^{(p\Theta+p^2\sigma^2/2)x}\left(\frac{\sigma^2 x}{2+\sigma^2 x/2}\right)\mu(dx)
\end{aligned}$$

where we have used the inequalities $\cos\alpha < 1$ and $e^{-\alpha\beta} - 1 \leq -(\alpha \wedge 1)\frac{\beta}{1+\beta}$ for all $\alpha, \beta > 0$. Therefore, Theorem 5.2 applies as claimed.

One realization of the above construction, popularized by Madan, Carr, and Chang in [12], is when $\mu(dx) = \frac{1}{\nu x}e^{-x/\nu}dx$ for some constant $\nu > 0$ and $a = 0$, so that Y is a gamma process, X is a variance gamma process, and

$$\Lambda_1(p) = \left[p\log\left(1-(\Theta+\sigma^2/2)\nu\right) - \log\left(1-(p\Theta+p^2\sigma^2/2)\nu\right)\right]/\nu.$$

In this case, the minimizer $p^* \in (0,1)$ of Λ_1 can be found by solving the equation $\Lambda_1'(p) = 0$ which is equivalent to the quadratic equation

$$\frac{\sigma^2}{2}p^2 + \left(\Theta - \frac{\sigma^2}{\log\left(1-(\Theta+\sigma^2/2)\nu\right)}\right)p - \left(1/\nu + \frac{\Theta}{\log\left(1-(\Theta+\sigma^2/2)\nu\right)}\right) = 0.$$

Note that for the variance gamma model, the call option prices can be expressed as

$$\begin{aligned}
\mathbb{E}[(S_\tau - e^k)^+] &= \mathbb{E}\left[e^{m\tau+(\Theta+\sigma^2/2)Y_\tau}\left(e^{-\sigma^2 Y_\tau+\sigma W(Y_\tau)} - e^{k-m\tau-(\Theta+\sigma^2/2)Y_\tau}\right)^+\right] \\
&= \mathbb{E}\left[e^{m\tau+(\Theta+\sigma^2/2)Y_\tau}\mathrm{BS}(k-m\tau-(\Theta+\sigma^2/2)Y_\tau, \sigma^2 Y_\tau)\right] \\
&= \int_0^\infty e^{m\tau+(\Theta+\sigma^2/2)\nu u}\mathrm{BS}(k-m\tau-(\Theta+\sigma^2/2)\nu u, \sigma^2\nu u)\frac{u^{t/\nu-1}e^{-u}}{\Gamma(t/\nu)}du
\end{aligned}$$

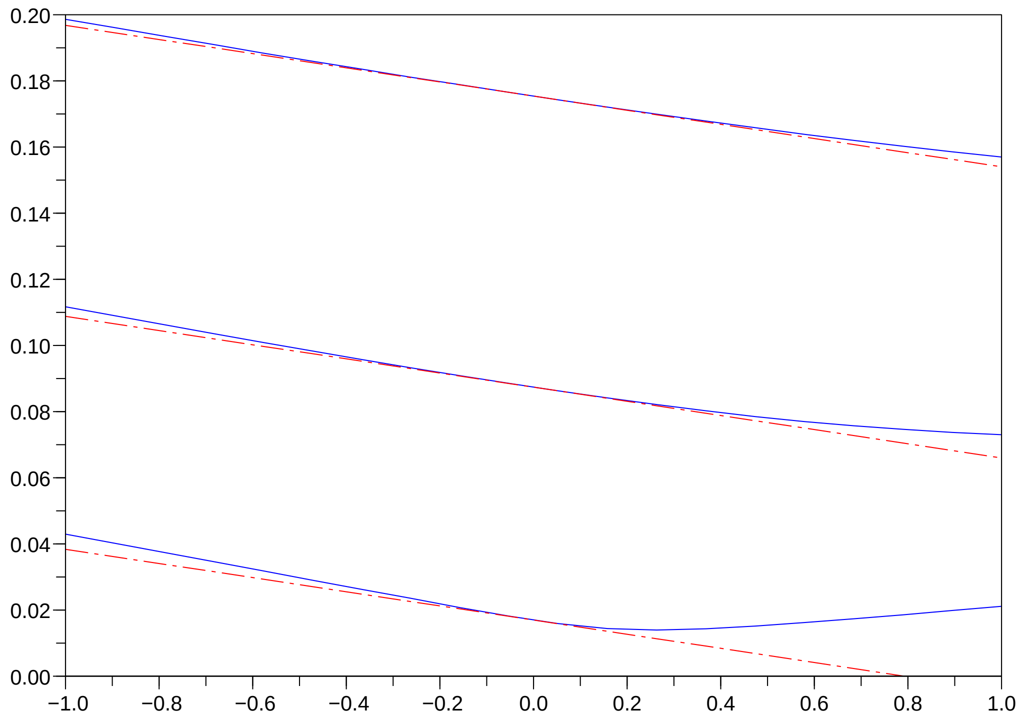

FIGURE 1. A comparison of the results of numerical integration (the blue, solid lines) with the affine approximation (the red, dotted lines) of the variance gamma model with parameters $\sigma = 0.1213$, $\nu = 0.1686$, and $\Theta = -0.1436$ for times to maturity $\tau = 1, 5$ and 10 years. The horizontal axis is log-moneyness $k = \log(K/S_0)$ and the vertical axis is implied variance $V(k, \tau)$.

and hence can be calculated by numerical integration.

In figure 1 we compare the affine approximation as given by Theorem 5.2 with the 'true' smile given by numerical integration with the parameter values $\sigma = 0.1213$, $\nu = 0.1686$, and $\Theta = -0.1436$ as found by Madan, Carr, and Chang [12] to fit S& P 500 European option prices between 1992 and 1994. The fit is surprisingly good even when the time to maturity is five years.

5.1.2. *A counterexample.* We now consider the extremely simple binomial model where we can do the calculations very explicitly. It turns out that the long implied total variance is not approximately affine in log-moneyness k as suggested by Theorem 5.2. Of course, the sufficient condition on the cumulant generating function fails to hold in this example.

Example 5.5 (Binomial model). Suppose $S_{\tau+1} = \xi_{\tau+1} S_\tau$ where ξ is a sequence of independent random variables such that $\mathbb{P}(\xi_\tau = e^b) = \frac{1}{e^b+1} = 1 - \mathbb{P}(\xi_\tau = e^{-b})$ for a constant $b > 0$. In this case, the moment generating function of $\log S_1$ is given by

$$M_1(p) = \frac{e^{bp} + e^{b(1-p)}}{e^b + 1} = \frac{\cosh[b(p-1/2)]}{\cosh(b/2)}.$$

The minimizing exponent is $p^* = 1/2$ by symmetry. A naive application of Theorem 5.2 would predict the following formula:

$$V(k,\tau) = 8\tau \ \log\cosh(b/2) + 4\log\left(\frac{b^2}{8\log\cosh(b/2)}\right) + \delta(k,\tau).$$

However, the above formula is *not correct*! Indeed, note

$$M_1(1/2 + iy) = \cos(by)M_1(1/2)$$

so the sufficient condition in Theorem 5.2 does not hold. In this simple example, we can actually compute the long implied total variance explicitly:

$$V(k,\tau) = 8\tau \ \log\cosh(b/2) - 8\log\cosh g(k,\tau) + 4\log\left(\frac{(\sinh b/2)^2}{2\log\cosh(b/2)}\right) + \delta(k,\tau)$$

where $g(\cdot,\tau)$ is the $2b$-periodic function whose restriction to the interval $(-b,b]$ is given by

$$g(k,\tau) = \begin{cases} |k|/2 & \text{if } \tau \text{ is odd} \\ (b-|k|)/2 & \text{if } \tau \text{ is even} \end{cases}$$

The above asymptotic formula may be regarded as something of a curiosity, as it would be hard to argue that the binomial model provides a good fit to stock price data. However, it serves as a warning that the integrability condition cannot be dropped from the statement of Theorem 4.7.

The above asymptotic formula is a consequence of the following proposition:

Proposition 5.6. *For each $\tau \in \mathbb{N}$, let*

$$F_\tau(y) = \int_{-\infty}^{\infty} \frac{(\cos x)^\tau \sqrt{\tau} e^{ixy}}{a^2 + x^2} dx.$$

Then $F_{2m+1} \to H$ and $F_{2m} \to H(\cdot + 1)$ uniformly on compacts, where H is the 2-periodic function whose restriction to $(-1,1]$ is given by

$$H(y) = \sqrt{2\pi}\frac{\cosh(ay)}{a\sinh a}.$$

Proof. We have by the absolute integrability of the integrand for each fixed $\tau \in \mathbb{N}$ and $y \in \mathbb{R}$ the calculation

$$\begin{aligned} F_\tau(y) &= \sum_{n\in\mathbb{Z}} \int_{(n-1/2)\pi}^{(n+1/2)\pi} \frac{\sqrt{\tau}(\cos x)^\tau e^{ixy}}{a^2 + x^2} dx \\ &= \sum_{n\in\mathbb{Z}} \int_{-\pi/2}^{\pi/2} (-1)^{n\tau} \frac{\sqrt{\tau}(\cos z)^\tau (-1)^{\tau n} e^{izy + iyn\pi}}{a^2 + (z+n\pi)^2} dz \\ &= \int_{-\pi/2}^{\pi/2} \sqrt{\tau}(\cos z)^\tau e^{izy} G(z,y,\tau) dz \end{aligned}$$

where

$$G(z,y,\tau) = \sum_{n\in\mathbb{Z}} \frac{(-1)^{n\tau} e^{iny\pi}}{a^2+(z+n\pi)^2}.$$

For all τ, the series defining $G(\cdot,\cdot,\tau)$ converges absolutely and uniformly on $[-\pi/2,\pi/2]\times[-\kappa,\kappa]$, and hence defines a continuous function.

Make the substitution $z=\theta/\sqrt{\tau}$ in the last integral above. First note that the integrand is uniformly bounded by an integrable function of θ, since $G(\cdot,\cdot,\cdot)$ is bounded on $[-\pi/2,\pi/2]\times[-\kappa,\kappa]\times\mathbb{N}$ and the inequality $\cos z \le 1 - z^2/\pi$ on $[-\pi/2,\pi/2]$ implies

$$|\cos(\theta/\sqrt{\tau})^{\tau}\mathbb{1}_{[-\pi\sqrt{\tau}/2,\pi\sqrt{\tau}/2]}(\theta)| \le e^{-\theta^2/\pi}.$$

Now letting τ take only odd values, we have the pointwise in $\theta\in\mathbb{R}$ uniform in $y\in[-\kappa,\kappa]$ convergence

$$[\cos(\theta/\sqrt{\tau})]^{2m+1} e^{iy\theta/\sqrt{\tau}} G(\theta/\sqrt{\tau},y,\tau)\mathbb{1}_{[-\pi\sqrt{\tau}/2,\pi\sqrt{\tau}/2]}(\theta) \to e^{-\theta^2/2}G(0,y,1)$$

and hence the dominated convergence theorem implies $F_{2m+1}(y)\to\sqrt{2\pi}G(0,y,1)$ uniformly. Similarly, $F_{2m}(y)\to\sqrt{2\pi}G(0,y,0)$ uniformly.

It is now a simple matter to check that the series

$$G(0,y,1) = \sum_{n\in\mathbb{Z}} \frac{(-1)^n e^{iyn\pi}}{a^2+n^2\pi^2}$$

is the Fourier series for the function $\frac{1}{\sqrt{2\pi}}H$. Since H is continuous and of bounded variation, its Fourier series converges everywhere. The case of even τ is similar. $\square$

To show that the long implied volatility in the binomial model is given by the announced formula, we need only work through the proof of Theorem 4.7 and substitute the integral in the above proposition with $a=b/2$ and $y=k/b$ in the appropriate place. Figure 2 compares the true implied total variance for the binomial model with $b=0.15$ with the approximation.

5.2. **Affine models.** In Markovian stochastic volatility models, the moment generating function of the stock price can be found generally by solving a partial differential equation. Affine models, however, have the attractive feature that the moment generating function can be found by solving a coupled family of ordinary differential equations, and in many cases are know in closed form. See Duffie, Filipović, and Schachermayer's paper [3] for a complete account of such models. We include the canonical example of an affine stochastic volatility model, the Heston model, to illustrate the technique. We only sketch the outline of the story. For full details for this example, see the recent preprint of Forde and Jacquier [4].

Example 5.7 (Heston model). Consider the following coupled SDE

$$\begin{aligned} dS_t &= S_t\sqrt{V_t}dW_t^{(S)} \\ dV_t &= \lambda(\theta-V_t)dt + \zeta\sqrt{V_t}dW_t^{(V)} \end{aligned}$$

for correlated Wiener processes $W^{(X)}$ and $W^{(V)}$ with $\langle W^{(X)},W^{(V)}\rangle_t = rt$, and positive constants λ, θ, and ζ.

Let

$$H(t,v,p) = e^{A(t,p)v+B(t,p)}$$

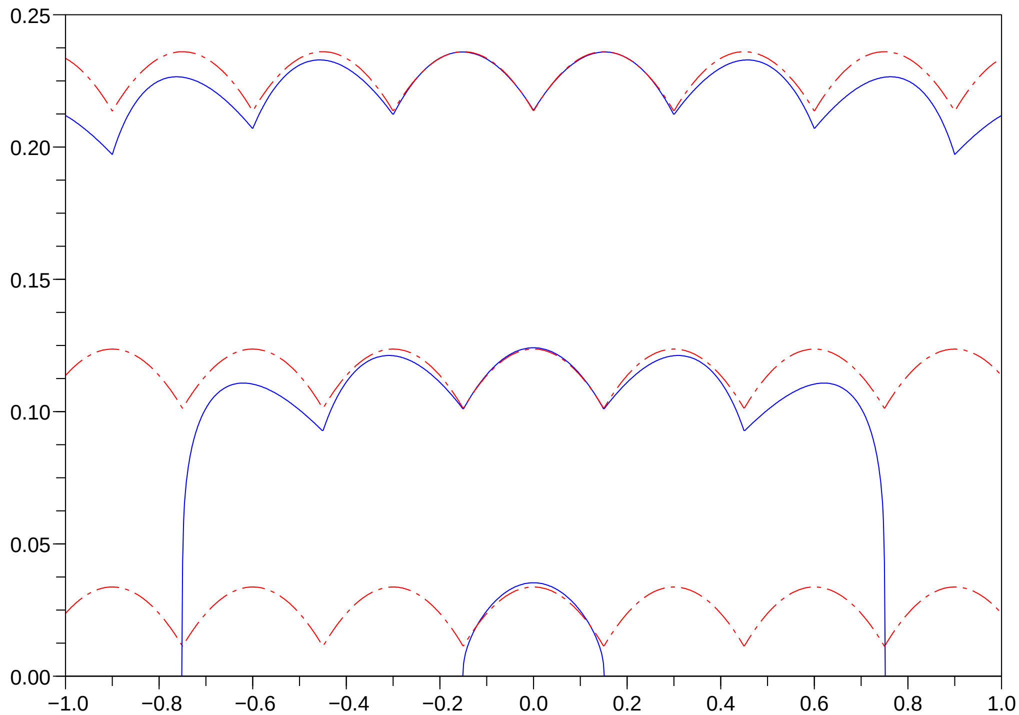

FIGURE 2. A comparison of the true implied variance (the blue, solid line) with the approximation (the red, dotted line) for the binomial model with $b = 0.15$ for times to maturity $\tau = 1, 5$, and 10 years. The horizontal axis is log-moneyness $k = \log(K/S_0)$ and the vertical axis is implied variance $V(k, \tau)$.

where the function A solves the Riccati equation

$$\begin{aligned} \frac{\partial}{\partial t} A(t,p) &= \frac{1}{2} p(p-1) + (pr\zeta - \lambda) A(t,p) + \frac{1}{2}\zeta^2 A(t,p)^2 \\ A(0,p) &= 0 \end{aligned}$$

and B is given by

$$B(t,p) = \lambda\theta \int_0^t A(s,p) ds.$$

It is easy to see by Itô's formula that the process $M_t = S_t^p H(t, V_t, p)$ defines a positive local martingale. We suppose that M is a true martingale for each p in some set $\Gamma \subseteq \mathbb{R}$, so that

$$\Lambda_\tau(p) = A(\tau, p) V_0 + B(\tau, p)$$

for all $p \in \Gamma$.

Following Keller-Ressel's paper [9], we notice that when $(pr\zeta - \lambda)^2 > \zeta^2 p(p-1)$, the Riccati equation has two fixed points, $A^-(p) < A^+(p)$, with $A^-(p)$ stable and $A^+(p)$ unstable.

Hence, when $p \in (0,1)$ or $pr\zeta < \lambda$ the unstable fixed point $A^+(p)$ is positive, so that $A(\tau, p)$ converges to $A^-(p)$, and in particular,

$$\bar{\Lambda}(p) = \lim_{\tau \uparrow \infty} \frac{\Lambda_\tau(p)}{\tau} = \lambda\theta A^-(p).$$

The minimizer p^* of this function $\bar{\Lambda}$ gives information about the long implied volatility, and was found in Chapter 6 of Lewis [11] to be

$$p^* = \frac{1}{2(1-r^2)\zeta}\left[\zeta - 2r\lambda + r\sqrt{\zeta^2 - 4\lambda\zeta r + 4\lambda^2}\right].$$

Note there are parameter values for which p^* defined by the above formula is such that $p^* > 1$ and $p^* r\zeta > \lambda$. (The referee has observed that the parameters $\zeta = 1, r = 3/4$ and $\lambda = 1/4$ have this property.) In these cases, it is not at all clear whether p^* satisfies Assumption 4.2 since $\frac{\Lambda_\tau(p)}{\tau}$ may not converge as $\tau \uparrow \infty$.

5.3. **Irregular cases.** We conclude with examples which illustrate what happens in the irregular case. Remember, the irregular case can only arise when the stock price is either a strictly local martingale, or hits zero in finite time with positive probability.

Example 5.8. The first example is extremely simple and is related to an example appearing in [13] verifying that the bound in Corollary 3.5 is sharp.

Let T be a random time where the distribution of $1/T$ is uniform on $[0,1]$, and let

$$S_t = \mathbb{1}_{\{0 \le t < 1\}} + t\mathbb{1}_{\{1 \le t < T\}}.$$

Then S is a martingale with respect to its natural filtration. Now, it is easy to see $\Lambda_\tau(p) = (p-1)\log t$ for $t \ge 1$. Letting $p^* = -1 < 0$, say, we have

$$\Lambda_\tau(p^*) - \Lambda_\tau(0) = -\log t \to -\infty$$

so Theorem 4.19 applies. The full asymptotics are then

$$V(k,\tau) = 8\log\tau - 4\log\log\tau - 4k - 4\log\pi + \delta(\tau,k).$$

Example 5.9 (CEV models). The CEV models, i.e. the models with constant elasticity of variance, are given by the SDE

$$dS_t = S_t^\alpha dW_t.$$

It is well-known that if $\alpha > 1$ then S is a strictly positive, strict local martingale. To illustrate the phenomenon, we consider the case $\alpha = 2$ corresponding to the inverse of a three-dimensional Bessel process. In this case, one has

$$\mathbb{E}(S_t) = 2\Phi\left(\frac{1}{\sqrt{t}}\right) - 1,$$

so that $\Lambda_\tau(1)/\log\tau \to -1/2$. A routine calculation shows $\Lambda_\tau(2)/\log\tau \to -1$, and hence the long implied total variance can be read off, using Theorem 4.19:

$$V(k,\tau) = 4\log\tau - 4\log\log\tau + 4k + \delta(k,\tau).$$

6. Acknowledgements

A preliminary report of this work was presented at the Conference on Implied Volatility organized by Princeton's Bendheim Center for Finance in Huntington Beach, California, in September 2008. I would like to thank the conference participants and especially Jim Gatheral for useful comments and suggestions.

Much of this paper was written during my visit to Ecole Polytechnique during the winter of 2009. Thanks go to Peter Tankov and Nizar Touzi for interesting discussions during my stay in Paris.

The current version of this paper was presented at the Conference on Small Time Asymptotics, Perturbation Theory and Heat Kernel Methods in Mathematical Finance, in Vienna in February 2009. I would like to thank the participants, with special thanks to Antoine Jacquier for carefully reading an earlier draft of this paper and alerting me his and Forde's preprint [4].

Finally, I would like to thank the referee for a detailed review of the article and for suggesting many improvements to the presentation.

References

[1] Carr, P. and Wu, L. (2003). The finite moment log stable process and option pricing. *Journal of Finance* **58**(2), 753–777.

[2] Cox, A.M.G. and Hobson, D.G. (2005). Local martingales, bubbles and option prices. *Finance and Stochastics* **9**, 477-492.

[3] Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability* **13**(3), 984–1053.

[4] Forde, M. and Jacquier, A. (2009). The large-maturity smile for the Heston model. Preprint. Dublin City University and Imperial College London.

[5] Fouque, J-P., Papanicolaou, G. and Sircar. K.R. (2000). *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge University Press.

[6] Gatheral, J. (1999). The volatility skew: Arbitrage constraints and asymptotic behaviour. Merrill Lynch.

[7] Grimmett, G.R. and Stirzaker, D.R. (2001). *Probability and Random Processes*. Third edition. Oxford University Press, New York.

[8] Jacquier, A. (2007). Asymptotic skew under stochastic volatility. Preprint. Birkbeck College, University of London.

[9] Keller-Ressel, M. (2008). Moment explosions and long-term behavior of affine stochastic volatility models. Preprint. TU Vienna.

[10] Lee, R. (2004). Implied volatility: Statics, dynamics, and probabilistic interpretation. In *Recent Advances in Applied Probability*, R. Baeza-Yates et al, eds. Springer.

[11] Lewis, A. (2000). *Option valuation under stochastic volatility*. Newport Beach: Finance Press.

[12] Madan, D.B., Carr, P.P. and Chang, E.C. (1998). The variance gamma process and option pricing. *European Finance Review* **2**, 79105.

[13] Rogers, L.C.G. and Tehranchi M.R. (2009). Can the implied volatility surface move by parallel shifts? To appear in *Finance and Stochastics*.

[14] Tehranchi, M.R. (2009). Implied volatility – Long maturity behavior. To appear in *Encyclopedia of Quantitative Finance*.

Statistical Laboratory, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK

E-mail address: m.tehranchi@statslab.cam.ac.uk