The Annals of Applied Probability
1999, Vol. 9, No. 2, 504–528

PRICING CONTINGENT CLAIMS ON STOCKS DRIVEN BY LÉVY PROCESSES[1]

BY TERENCE CHAN

Heriot-Watt University

We consider the problem of pricing contingent claims on a stock whose price process is modelled by a geometric Lévy process, in exact analogy with the ubiquitous geometric Brownian motion model. Because the noise process has jumps of random sizes, such a market is incomplete and there is not a unique equivalent martingale measure. We study several approaches to pricing options which all make use of an equivalent martingale measure that is in different respects "closest" to the underlying canonical measure, the main ones being the Föllmer–Schweizer minimal measure and the martingale measure which has minimum relative entropy with respect to the canonical measure. It is shown that the minimum relative entropy measure is that constructed via the Esscher transform, while the Föllmer–Schweizer measure corresponds to another natural analogue of the classical Black–Scholes measure.

1. Introduction. We consider the problem of pricing contingent claims on a stock whose price at time t, S_t, is modelled by a geometric Lévy process

$$dS_t = \sigma_t S_{t-}\, dY_t + b_t S_{t-}\, dt,$$

where Y is a general Lévy process (satisfying some additional conditions) and not merely a Brownian motion. The classical option pricing theory of Black and Scholes relies on the fact that the payoff of every contingent claim can be duplicated by a portfolio consisting of investments in the underlying stock and in a bond paying a riskless rate of interest; in other words, the risk of buying or writing an option can be completely hedged against. In such complete markets, there is a unique measure which is equivalent to the canonical measure (the "real world" measure) and which makes the discounted price process a martingale. The unique fair price of a contingent claim is then the expectation under this martingale measure of the discounted payoff at maturity, which is essentially the content of the famous Black–Scholes formula.

For the stock prices described above, there are many equivalent measures under which the discounted price process is a martingale, in contrast to the geometric Brownian model. In other words, such a market is incomplete—that is, contingent claims cannot in general be hedged by a suitable portfolio. Because there does not exist a unique equivalent martingale measure, it is not possible simply to use the martingale measure to price a contingent claim in the manner just described. Instead, additional criteria must be used to select

Received April 1997; revised November 1997.
[1]Supported in part by the Carnegie Trust for the Universities of Scotland.
AMS 1991 *subject classifications*. Primary 90A09, 60G35; secondary 60J30, 60J75.
Key words and phrases. Option pricing, incomplete market, equivalent martingale measures.

an appropriate martingale measure from among the uncountably many such measures with which to price a contingent claim. Many different approaches to this problem have been proposed in recent years but there is as yet no definitive way of pricing contingent claims in incomplete markets which is preferable to the other possible methods in all situations. Moreover, compared to the large body of work devoted to finding new approaches to option pricing in incomplete markets, relatively little seems to have been done to compare and to investigate the relationship between the various approaches. Part of the aim of this paper is to go a little way toward redressing the balance. For our particular model, we shall concentrate on various approaches to pricing options which are all based on the idea of using an equivalent martingale measure that is in different respects "closest" to the underlying canonical measure, the main ones being the Föllmer–Schweizer minimal measure and the martingale measure which has minimum relative entropy with respect to the canonical measure.

2. Description of the model. Before describing the model, we first review some preliminary results concerning Lévy processes. For a more detailed treatment, the reader is referred to Protter (1990), Jacod and Shiryaev (1987) and Liptser and Shiryayev (1989).

A Lévy process Y_t is simply a process with stationary and independent increments: in other words, $Y_{s+t} - Y_s$ is independent of $\{Y_u\colon u \le s\}$ and has the same distribution as $Y_t - Y_0$. All Lévy processes are semimartingales and throughout this paper we adopt the convention that all Lévy processes are right continuous with left limits (cadlag).

Since Y has stationary independent increments, its characteristic function must take the form

$$\mathbb{E}[\exp(-i\theta Y_t)] = \exp(-t\psi(\theta))$$

for some function ψ, called the *Lévy exponent* of Y. The Lévy–Khintchine formula says that

$$\psi(\theta) = \frac{c^2}{2}\theta^2 + i\alpha\theta + \int_{\{|x|<1\}} \big(1 - e^{-i\theta x} - i\theta x\big)\,\nu(dx) + \int_{\{|x|\ge 1\}} \big(1 - e^{-i\theta x}\big)\,\nu(dx) \tag{2.1}$$

for α, $c \in \mathbb{R}$ and for some σ-finite measure ν on $\mathbb{R}\backslash\{0\}$ satisfying

$$\int \min(1, x^2)\,\nu(dx) < \infty. \tag{2.2}$$

The measure ν is called the *Lévy measure* of Y.

The Lévy–Khintchine formula (2.1) is intimately connected to the structure of the process Y itself, in particular to the *Lévy decomposition* of Y, which we describe below. From the Lévy–Khintchine formula we can deduce that Y must be a linear combination of a Brownian motion and a quadratic pure jump process X which is independent of the Brownian motion. [A process is said to be quadratic pure jump if the continuous part of its quadratic variation $\langle X\rangle^c \equiv 0$, in which case its quadratic variation becomes simply

$$\langle X\rangle_t = \sum_{0<s\le t} (\Delta X_s)^2,$$

where $\Delta X_s = X_s - X_{s-}$ is the jump size at time s.] It will be convenient to explicitly separate out the Brownian component from the quadratic pure jump component X and we therefore write

$$Y_t = cB_t + X_t, \tag{2.3}$$

where B is a standard Brownian motion on $\mathbb{R}$ and X is quadratic pure jump. We now proceed to describe the Lévy decomposition of X [the full Lévy decomposition of Y is then obtained by combining this with (2.3)].

Let $Q(dt, dx)$ be a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}\backslash\{0\}$ with expectation (or intensity) measure $dt \times \nu$, where ν is the Lévy measure introduced earlier and dt denotes Lebesgue measure. The measure ν (or more precisely $dt \times \nu$) is also sometimes called the compensator of Q. The Lévy decomposition of X says that

$$\begin{aligned} X_t &= \int_{\{|x|<1\}} x\big(Q((0,t],dx) - t\nu(dx)\big) + \int_{\{|x|\ge 1\}} x\, Q\big((0,t],dx\big) \\ &\quad + t\mathbb{E}\bigg[X_1 - \int_{\{|x|\ge 1\}} x\,\nu(dx)\bigg] \\ &= \int_{\{|x|<1\}} x\big(Q((0,t],dx) - t\nu(dx)\big) + \int_{\{|x|\ge 1\}} x\, Q((0,t],dx) + \alpha t, \end{aligned} \tag{2.4}$$

where we have put

$$\alpha = \mathbb{E}\bigg[X_1 - \int_{\{|x|\ge 1\}} x\,\nu(dx)\bigg].$$

The parameter α is called the drift of the Lévy process X.

For the purposes of our model, we require the process Y to satisfy certain additional conditions. The key assumption we require of Y is that

$$\mathbb{E}[\exp(-hY_1)] < \infty \quad \text{for all } h \in (-h_1, h_2), \tag{2.5}$$

where $0 < h_1, h_2 \le \infty$. This implies that Y_t has finite moments of all orders, and in particular, $\mathbb{E}[X_1] < \infty$. In terms of the Lévy measure ν of X, we have

$$\int_{\{|x|\ge 1\}} e^{-hx}\,\nu(dx) < \infty, \tag{2.6a}$$

$$\int_{\{|x|\ge 1\}} x^{\gamma} e^{-hx}\,\nu(dx) < \infty \quad \forall \gamma > 0, \tag{2.6b}$$

$$\int_{\{|x|\ge 1\}} x\,\nu(dx) < \infty \tag{2.6c}$$

for all $h \in (-h_1, h_2)$. [Note that as (2.6a) holds for all h in an open interval, (2.6b) and (2.6c) follow from (2.6a).] With these assumptions in mind, (2.4) can be rewritten as

$$X_t = \int_{\mathbb{R}} x(Q((0,t], dx) - t\nu(dx)) + t\mathbb{E}[X_1] \\ = M_t + at, \tag{2.7}$$

where $M_t = \int_{\mathbb{R}} x(Q((0,t], dx) - t\nu(dx))$ is a martingale and $a = \mathbb{E}[X_1]$. Observe that (2.7) gives the Doob decomposition of X as the sum of a martingale and a previsible process of finite variation. Even though a is not the drift of X in the sense in which the term is usually understood (α is the drift in the technical sense), we shall see later that a plays the role of a drift contribution from the jump component of Y. We refer to a (or more correctly, the process $t \mapsto at$) as the previsible part of X.

In addition, (2.5) implies that instead of the characteristic function, one could consider the Laplace transform of Y_t instead. By a slight abuse of notation, we also use ψ to denote the "Lévy exponent" and write $\mathbb{E}[\exp(-\theta Y_t)] = \exp(-t\psi(\theta))$. Bearing in mind the simplified decomposition (2.7) for processes satisfying (2.5), the Lévy–Khintchine formula (2.1) now becomes

$$\psi(\theta) = -\frac{c^2\theta^2}{2} + a\theta + \int_{\mathbb{R}} (1 - e^{-\theta x} - \theta x)\,\nu(dx). \tag{2.8}$$

A very similar analysis can be carried out for more general semimartingales with jumps and in particular for processes with independent but not necessarily stationary increments. Jacod and Shiryaev (1987) have a full treatment. A random measure $Q(dt, dx)$ is also associated with such a process, but it is not necessarily a Poisson measure. As in the case of Lévy processes, the measure Q describes the mechanism by which jumps of the process occur. The compensator of Q is the unique previsible measure $\nu(dt, dx)$ such that $Q([0,t], \Lambda) - \nu([0,t], \Lambda)$ is a martingale for any Borel set $\Lambda \subset \mathbb{R}\backslash\{0\}$. If the process in question has independent increments, the measure ν is necessarily deterministic, so Q is an inhomogeneous Poisson measure. [For Lévy processes, the stationarity of increments implies that $\nu(dt, dx) = dt\,\nu(dx)$.] The compensator can also be characterized as the unique previsible measure such that

$$\mathbb{E}\left[\int_{[0,t]\times\Lambda} H(s,x)Q(ds,dx)\right] = \mathbb{E}\left[\int_{[0,t]\times\Lambda} H(s,x)\nu(ds,dx)\right] \tag{2.9}$$

for any Borel set Λ and any previsible process H. We also have an analogue of the Lévy–Khintchine formula: $\mathbb{E}[\exp(-\theta X_t)] = \exp\{-\psi_X(t,\theta)\}$, where

$$\psi_X(t,\theta) = a_t\theta + \int_{\mathbb{R}} \big(1 - \exp(-\theta x) - \theta x\big)\,\nu\big([0,t],\ dx\big), \tag{2.10}$$

where $a_t = \mathbb{E}[X_t]$ is the previsible part of X. Together with the quadratic variation of the continuous part of X (which is zero if X is quadratic pure jump as in our case), the compensator measure and previsible part form the three components of the *characteristics* of a semimartingale. The following result is also worth noting: for any measurable function $f(t, x)$,

$$\sum_{0<s\le t} f(s, \Delta X_s) = \int_0^t \int_{\mathbb{R}} f(s, x)\, Q(ds, dx). \tag{2.11}$$

Next, we recall Itô's formula for cadlag semimartingales. If $X^1, X^2, \dots, X^n$ are cadlag semimartingales and f a C^2 function, then

$$\begin{aligned} f(X_t^1 \cdots X_t^n) &- f(X_0^1 \cdots X_0^n) \\ &= \int_0^t f_i(X_{s-}^1 \cdots X_{s-}^n)\, dX_s^i + \tfrac{1}{2}\int_0^t f_{ij}(X_{s-}^1 \cdots X_{s-}^n)\, d[X^i, X^j]_s^c \\ &\quad + \sum_{0<s\le t} \Big[f(X_s^1 \cdots X_s^n) - f(X_{s-}^1 \cdots X_{s-}^n) - f_i(X_{s-}^1 \cdots X_{s-}^n)\, \Delta X_s^i \Big], \end{aligned}$$

where $[X^i, X^j]^c$ is the continuous part of the mutual variation of X^i and X^j, $f_i = \partial f/\partial x_i$, $f_{ij} = \partial^2 f/\partial x_i \partial x_j$ and we have used index summation convention. This will often be abbreviated to

$$\begin{aligned} df(X_t^1, X_t^2 \cdots X_t^n) &\\ &= f_i(X_{t-}^1 \cdots X_{t-}^n)\, dX_t^i + \tfrac{1}{2} f_{ij}(X_{t-}^1 \cdots X_{t-}^n)\, d[X^i, X^j]_t^c \\ &\quad + f(X_t^1 \cdots X_t^n) - f(X_{t-}^1 \cdots X_{t-}^n) - f_i(X_{t-}^1 \cdots X_{t-}^n)\, \Delta X_t^i. \end{aligned}$$

Turning now to a description of the model, on a probability space $(\Omega, \{\mathscr{F}_t\}, \mathbb{P})$, let $Y_t = cB_t + X_t = cB_t + M_t + at$ be a Lévy process of the form described earlier, satisfying the condition (2.5). We assume that the filtration $\{\mathscr{F}_t\}$ is the minimal one generated by Y. The stock price S_t is the solution of the stochastic differential equation

$$\begin{aligned} dS_t &= \sigma_t S_{t-}\, dY_t + b_t S_{t-}\, dt \\ &= \sigma_t S_{t-}(c\, dB_t + dM_t) + (a\sigma_t + b_t) S_{t-}\, dt, \end{aligned} \tag{2.12}$$

where the coefficients σ_t and b_t are deterministic continuous functions. Equation (2.12) has an explicit solution [see Protter (1991)] given by

$$\begin{aligned} S_t &= S_0 \exp\left\{ \int_0^t \sigma_s\, dY_s + \int_0^t \left(b_s - \frac{c^2\sigma_s^2}{2} \right) ds \right\} \\ &\quad \times \prod_{0<s\le t} (1 + \sigma_s \Delta Y_s) \exp(-\sigma_s \Delta Y_s) \\ &= S_0 \exp\left\{ \int_0^t c\sigma_s\, dB_s + \int_0^t \sigma_s\, dM_s + \int_0^t \left(a\sigma_s + b_s - \frac{c^2\sigma_s^2}{2} \right) ds \right\} \\ &\quad \times \prod_{0<s\le t} (1 + \sigma_s \Delta M_s) \exp(-\sigma_s \Delta M_s). \end{aligned} \tag{2.13}$$

From this we see that $\sigma\{S_u \colon u \le t\} = \mathscr{F}_t$ and so a contingent claim Γ_T expiring at time T may be regarded as a nonnegative $\mathscr{F}_T$-measurable random variable.

The Doob decomposition of Y suggests that $b_t + a\sigma_t$ rather than b_t should be regarded as the drift in (2.12). Although in practice, a and b cannot be estimated separately and consequently there is no need to add a drift to X separately from b in (2.12), we have chosen to consider the parameters a and b separately for convenience, because the value of a is often implicit in the specification of a particular process as X and so cannot be chosen independently (e.g., if we specify that X be a Poisson process of rate λ, this forces $a = \lambda$).

In order to ensure that $S_t \geq 0$ for all t almost surely, we need $\sigma_t \Delta M_t \geq -1$ for all t. This in turn implies that the jumps of X must be bounded on at least one side, that is, either bounded from below or bounded from above. Suppose that $\Delta X_t = \Delta M_t \in [-c_1, c_2]$, which is equivalent to saying that the Lévy measure ν is supported on $[-c_1, c_2]$ where c_1, $c_2 \geq 0$ and one (but not both) of c_1, c_2 may be infinite. This implies that at least one of h_1, h_2 in (2.5) must be infinite. In order to ensure that $S_t \geq 0$ we need

$$-\frac{1}{c_2} \leq \sigma_t \leq \frac{1}{c_1} \quad \text{for all } t. \tag{2.14}$$

As far as the Brownian component of Y is concerned, the sign of the volatility σ is inconsequential, but if one were to keep to the usual convention that $\sigma > 0$, then (2.14) shows that the jumps of X should be bounded from below (i.e., $c_1 < \infty$). The conditions (2.5) and (2.14) will of course rule out any processes with "fat-tailed" distributions such as stable processes. However, the allowable Lévy processes here include all the processes considered in Gerber and Shiu (1994): for example the gamma, the inverse Gaussian, the Poisson and the difference of two independent Poisson processes.

The riskless rate of interest is given by a deterministic continuous function r_t and the value P_t of a bond or bank account paying this rate of interest evolves according to the ODE

$$\dot{P}_t = r_t P_t.$$

As with σ and b, we could also allow r to be adapted to $\{\mathscr{F}_t\}$, although this is a less useful generalization in practice. For notational convenience, we denote by $\hat{S}_t$ the discounted stock price defined by

$$\hat{S}_t = \exp\left\{-\int_0^t r_s \, ds\right\} S_t. \tag{2.15}$$

It will be seen in the next section that, in this model, there are many measures, equivalent to the underlying canonical measure $\mathbb{P}$, which makes $\hat{S}_t$ a martingale.

We conclude this section by briefly mentioning some other similar models which have been considered by various authors. Bardhan and Chao (1993) considered a similar model where the noise consists of several Brownian motions and several point processes whose jumps are all of size 1 but whose intensities may not be time-homogeneous and may be random. However, the contingent claims they considered are on more than one stock, where the number of stocks exactly equals the total number of noise terms (Brownian motions and point processes). This, together with the fact that the jump sizes are fixed, ensure that their model is complete. Aase (1988) is essentially an attempt at a more general model than that of Bardhan and Chao, where the point process may have random jump sizes but still a finite number of jumps in any finite time interval. Unfortunately, Aase (1988) claims that the model is also complete even though there are more than one equivalent martingale measure; this is false because it contradicts a well-known theorem of Harrison and Pliska (1981, 1983) to the effect that completeness of the market is equivalent to uniqueness of the equivalent martingale measure. Indeed, Aase (1988) claims that every martingale can be represented as an integral with respect to $\hat{S}_t$, in the form

$$\int_0^t \theta_s \, d\hat{S}_s, \tag{2.16}$$

where θ_t is a previsible process. (The existence of such a representation is equivalent to completeness.) This is false, as the martingale representation theorem [see, e.g., Jacod and Shiryaev (1987)] for the jump processes considered in Aase (1988) (which includes certain classes of Lévy processes) says that every martingale has the representation

$$\int_0^t H(s,x)(\tilde{Q}(ds,dx) - \tilde{\nu}(ds,dx)),$$

where $\tilde{Q}(ds,dx)$ is a random jump measure whose compensator is $\tilde{\nu}$—analogous, respectively to the Poisson and Lévy measures associated with a Lévy process—and where $H(s,x)$ is a previsible Borel function (see the next section for a precise definition). We shall see in the next section that, under any equivalent martingale measure, the jump part of $\hat{S}_t$ has the representation

$$\int_0^t \gamma_s \, d\tilde{M}_s = \int_0^t \int_{\mathbb{R}} \gamma_s x \big(\tilde{Q}(ds,dx) - \tilde{\nu}(ds,dx)\big).$$

Hence, in order that the representation (2.16) holds, we need $H(s,x) = \theta_s \gamma_s x$, which of course is not true in general. Finally, Gerber and Shiu (1994) consider the case where the stock price is modelled by a process of the form $\exp\{\sigma Y_t + bt\}$, where σ and b are constants and Y is a Lévy process satisfying (2.5). This has many similarities with our present model and both are obvious generalizations of the geometric Brownian model. The program carried out in the next section can be equally well carried out for the Gerber–Shiu model, often with only fairly minor modifications. Each model has its own advantages and disadvantages. The main advantage of the Gerber–Shiu model is that the jumps of X can be of any size and do not have to be bounded from one side. The present model based on (2.12) describes the price dynamics in a manner which is intuitively more natural and is also more appealing in other mathematical respects. This is because the starting point of the classical geometric Brownian model is (2.12); that the price S_t also has the form $\exp\{\sigma' Y_t + b' t\}$ is a direct consequence of the stochastic calculus involved, in particular, Itô's formula. For discontinuous Lévy processes, Itô's formula is rather different and so a model which takes as its starting point a differential equation like (2.12) and then takes account of the differences in the underlying stochastic calculus in the subsequent computations is more likely to lead to simpler calculations and more attractive results. This point is illustrated in Section 3.3 in relation to the Esscher transform and minimum relative entropy measure. Gerber and Shiu (1994) deal only with pricing contingent claims by Esscher transforms, without explaining why the Esscher transform is a particularly appropriate martingale measure to use. [However, in their response to the discussions that follow their paper, they give a justification of the Esscher transform in terms of utility; see page 175 of Gerber and Shiu (1994).] We shall show that it is the martingale measure which has minimum relative entropy with respect to the canonical measure.

3. Equivalent martingale measures and pricing formulas. We begin by characterizing all equivalent martingale measures $\mathbb{Q}$ under which the discounted price process $\hat{S}$ defined at (2.15) is a $\{\mathscr{F}_t\}$-martingale. To this end, we first need to characterize all the measures which are absolutely continuous with respect to $\mathbb{P}$.

We continue to use the notation established in the previous section. In particular, $Y_t = cB_t + X_t$ is a Lévy process satisfying (2.5) and X_t is a quadratic pure jump Lévy process with Lévy measure ν supported on a subset of $[-c_1, c_2]$, where at least one of c_1, c_2 is finite. The Doob–Meyer decomposition of X is given by $X_t = M_t + at$, where M is a quadratic pure jump martingale with $M_0 = 0$ and $a = \mathbb{E}[X_1]$. If $Q(dt, dx)$ is the Poisson measure associated with X, let $M(dt, dx) = Q(dt, dx) - dt\,\nu(dx)$ denote the compensated measure. Thus, for example, the martingale part of X can be written as $M_t = \int_0^t \int_{\mathbb{R}} x\, M(ds, dx)$. Further, expectations under the canonical measure $\mathbb{P}$ will be denoted by $\mathbb{E}[\cdot]$ while expectations with respect to any other measure $\mathbb{Q}$ will be denoted by $\mathbb{Q}[\cdot]$.

Let $\mathscr{P}$ denote the previsible σ-algebra on $\Omega \times \mathbb{R}^+$ associated with the filtration $\{\mathscr{F}_t\}$ and let $\tilde{\mathscr{P}} = \mathscr{P} \times \mathscr{B}$, where $\mathscr{B}$ is the Borel σ-algebra on $\mathbb{R}$. A function $H(\omega, t, x)$ which is $\tilde{\mathscr{P}}$-measurable will be called Borel previsible. Thus, suppressing the explicit dependence on ω, a Borel previsible function or process $H(t, x)$ is one such that the process $t \mapsto H(t, x)$ is previsible for fixed x and the function $x \mapsto H(t, x)$ is Borel-measurable for fixed t.

LEMMA 3.1. *Let G_t and $H(t, x)$ be previsible and Borel previsible processes respectively. Suppose that*

$$\mathbb{E}\left[\int_0^t G_s^2\, ds\right] < \infty$$

and $H \geq 0$, $H(t,0) = 1$ *for all* $t \geq 0$. *Let* $h(t,x)$ *be another Borel previsible process such that*

$$\int_{\mathbb{R}} \big[H(t,x) - 1 - h(t,x)\big]\,\nu(dx) < \infty. \tag{3.1}$$

Define a process Z_t *by*

$$\begin{aligned} Z_t = \exp\bigg\{ \int_0^t G_s\,dB_s - \tfrac{1}{2}\int_0^t G_s^2\,ds + \int_0^t \int_{\mathbb{R}} h(s,x)\,M(ds,dx) \\ - \int_{[0,t)\times\mathbb{R}} \big[H(s,x) - 1 - h(s,x)\big]\,\nu(dx)\,ds \bigg\} \\ \times \prod_{0<s\leq t} H(s,\Delta X_s)\exp(-h(s,\Delta X_s)). \end{aligned} \tag{3.2}$$

Then Z *is a nonnegative local martingale with* $Z_0 = 1$ *and* Z *is positive if and only if* $H > 0$.

REMARK. The process h referred to in Lemma 3.1 is, of course, not unique. However, given H, it is essentially unique in the following sense: suppose that $h(t,x)$ and $f(t,x)$ are two Borel previsible processes such that (3.1) holds; then because $\int_{\mathbb{R}}(f(t,x) - h(t,x))\,\nu(dx) < \infty$, it is an easy exercise to check that the process Z is unchanged if h is replaced by f in (3.2): simply write $f = h + (f - h)$. [However, note that it is crucial that $\int_{\mathbb{R}}(f(t,x) - h(t,x))\,\nu(dx) < \infty$: the terms involving h in (3.2) do not cancel precisely because $\int_{\mathbb{R}} h(t,x)\,\nu(dx)$ may diverge.] Thus, once H is fixed, Z does not depend on the choice of the process h satisfying (3.1). Of course, the easiest and most obvious choice of h is $h \equiv H - 1$. However, in the present context, particularly in connection with the Esscher transform discussed below, it is useful to allow more general choices of h. In the case where $x \mapsto H(t,x)$ is twice-differentiable, the natural choice of $h(t,x)$ is

$$h(t,x) = x\frac{\partial H}{\partial x}(t,0) = h_t x \quad \text{say,}$$

for then $H(t,x) \sim 1 + h_t x + O(x^2)$ as $x \to 0$ and because of (2.6c) we simply have to choose H so that

$$\int_{|x|\geq 1} H(t,x)\,\nu(dx) < \infty.$$

We shall henceforth assume that $h(t,x) = h_t x$ is related to $H(t,x)$ in this way.

PROOF OF LEMMA 3.1. It is clear that Z is nonnegative (resp., positive) if and only if $H \geq 0$ (resp., $H > 0$). That Z is a local martingale is a simple consequence of Itô's formula; indeed, noting that $Z_t - Z_{t-} = Z_{t-}(H(t, \Delta X_t) - 1)$, Itô's formula gives

$$\begin{aligned}
Z_t &= 1 + \int_0^t G_s Z_{s-}\, dB_s + \int_0^t \int_{\mathbb{R}} h(s,x) Z_{s-}\, M(ds, dx) \\
&\quad - \int_0^t \int_{\mathbb{R}} Z_{s-}[H(s,x) - 1 - h(s,x)]\, \nu(dx)\, ds \\
&\quad + \sum_{0<s\le t} Z_{s-}\big[H(t, \Delta X_t) - 1 - h(s, \Delta X_s)\big] \\
&= 1 + \int_0^t G_s Z_{s-}\, dB_s + \int_0^t \int_{\mathbb{R}} h(s,x) Z_{s-}\, M(ds, dx) \\
&\quad + \int_0^t \int_{\mathbb{R}} Z_{s-}[H(s,x) - 1 - h(s,x)]\, M(ds, dx) \\
&= 1 + \int_0^t G_s Z_{s-}\, dB_s + \int_0^t \int_{\mathbb{R}} Z_{s-}\big[H(s,x) - 1\big]\, M(ds, dx).
\end{aligned}$$

This last expression is a local martingale. □

The processes G, H and h can be chosen so that $\mathbb{E}[Z_t] = 1$ for all t, in which case Z is a martingale.

The next result is essentially a summary of Theorems 3.24 and 5.19 in Chapter III of Jacod and Shiryaev (1987) as they apply to the present setting.

THEOREM 3.2. *Let $\tilde{\mathbb{P}}$ be a measure which is absolutely continuous with respect to $\mathbb{P}$ on $\mathscr{F}_T$. Then*

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathscr{F}_T} = Z_T,$$

where Z is as in Lemma 3.1, *for some G, H and h for which $\mathbb{E}[Z_T] = 1$. Moreover, under $\tilde{\mathbb{P}}$, the process*

$$\tilde{B}_t = B_t - \int_0^t G_s\, ds \tag{3.3}$$

is a Brownian motion and the process X is a quadratic pure jump process with compensator measure given by $\tilde{\nu}(dt, dx) = dt\, \tilde{\nu}_t(dx)$, where

$$\tilde{\nu}_t(dx) = H(t,x)\, \nu(dx), \tag{3.4}$$

and previsible part given by

$$\tilde{a}_t = \tilde{\mathbb{P}}[X_t] = at + \int_0^t \int_{\mathbb{R}} x(H(s,x) - 1)\, \nu(dx)\, ds. \tag{3.5}$$

REMARK. Jacod and Shiryaev (1987) treat only the case that $h \equiv H - 1$ for the process Z in Lemma 3.1. Also, in their treatment of characteristics of general semimartingales, Jacod and Shiryaev (1987) introduce truncation functions, and the corresponding results in Theorem 3.24 of that book depend in part on the choice of truncation function. In the present situation, assumption (2.6c) renders the introduction of truncation functions unnecessary.

Turning now to the problem of pricing a contingent claim Γ_T, we wish to find an equivalent measure $\mathbb{Q}$ under which the discounted price process $\hat{S}_t$ as defined in (2.15) is a martingale; the price of Γ_T is then $\mathbb{Q}[\exp\{-\int_0^T r_s\,ds\}\Gamma_T]$. By Theorem 3.2, under $\mathbb{Q}$, X has Doob–Meyer decomposition

$$X_t = \tilde{M}_t + at + \int_0^t \int_{\mathbb{R}} x\big(H(s,x)-1\big)\,\nu(dx)\,ds, \tag{3.6}$$

where $\tilde{M}$ is a $\mathbb{Q}$-martingale. In fact,

$$\tilde{M}_t = M_t - \int_0^t \int_{\mathbb{R}} x\big(H(s,x)-1\big)\,\nu(dx)\,ds,$$

where M is the $\mathbb{P}$-martingale in the Doob–Meyer decomposition of X under $\mathbb{P}$. Note that $\Delta\tilde{M}_t = \Delta M_t$. Therefore, writing the discounted share price $\hat{S}_t$ in terms of the $\mathbb{Q}$-martingale $\tilde{M}$ and $\mathbb{Q}$-Brownian motion $\tilde{B}$, we have

$$\begin{aligned}
\hat{S}_t &= S_0 \exp\left\{\int_0^t c\sigma_s\,dB_s + \int_0^t \sigma_s\,dM_s + \int_0^t \left(a\sigma_s + b_s - r_s - \frac{c^2\sigma_s^2}{2}\right) ds\right\} \\
&\quad\times \prod_{0<s\le t} (1+\sigma_s\Delta M_s)\exp(-\sigma_s\Delta M_s) \\
&= S_0 \exp\left\{\int_0^t c\sigma_s\,d\tilde{B}_s + \int_0^t \sigma_s\,d\tilde{M}_s + \int_0^t \left(a\sigma_s + c\sigma_s G_s + b_s - r_s - \frac{c^2\sigma_s^2}{2}\right) ds \right. \\
&\qquad\qquad \left. + \int_0^t \sigma_s \int_{\mathbb{R}} x(H(s,x)-1)\,\nu(dx)\,ds\right\} \\
&\quad\times \prod_{0<s\le t} (1+\sigma_s\Delta\tilde{M}_s)e^{-\sigma_s\Delta\tilde{M}_s}.
\end{aligned}$$

Since

$$\exp\left\{\int_0^t c\sigma_s\,d\tilde{B}_s + \int_0^t \sigma_s\,d\tilde{M}_s - \int_0^t \frac{c^2\sigma_s^2}{2}\,ds\right\} \prod_{0<s\le t}(1+\sigma_s\Delta\tilde{M}_s)\exp(-\sigma_s\Delta\tilde{M}_s)$$

is a $\mathbb{Q}$-martingale, a necessary and sufficient condition for $\hat{S}$ to be a martingale under $\mathbb{Q}$ is the existence of G and H for which the process Z in Lemma 3.1 is a positive martingale and such that

$$c\sigma_s G_s + a\sigma_s + b_s - r_s + \int_{\mathbb{R}} \sigma_s x(H(s,x)-1)\,\nu(dx) = 0 \tag{3.7}$$

for all s, almost surely. Note that h does not appear in (3.7), which is another reflection of the fact that h is essentially unique, given H, in the sense of the remark following Lemma 3.1. It will turn out that G and H are in fact deterministic functions in all the cases considered in the sequel; in this case, (2.5) ensures that Z in Lemma 3.1 is a positive martingale and the key condition for an equivalent martingale measure is then (3.7). Moreover, B and

X are still independent and have independent increments under $\mathbb{Q}$; in this connection, note that $\tilde{\nu}$ is a deterministic measure.

Of course, (3.7) does not specify G and H, and hence the equivalent martingale measure $\mathbb{Q}$, uniquely. Below, we examine various approaches to choosing G and H based on other criteria, additional to (3.7).

3.1. *The Föllmer–Schweizer minimal measure.* Recall that when the noise Y in (2.12) is just a standard Brownian motion, the unique equivalent martingale measure $\mathbb{Q}$ is obtained by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathscr{F}_T} = Z_T, \tag{3.8}$$

where Z satisfies

$$dZ_t = \gamma_t Z_t \, dB_t$$

and the process γ is chosen so as to make $\hat{S}$ a martingale under $\mathbb{Q}$. In the present setting, a natural analogue of this would be to use the martingale measure $\mathbb{Q}$ defined by (3.8), where the Radon–Nikodym derivative Z is now given by

$$dZ_t = \gamma_t Z_{t-}(c\, dB_t + dM_t),$$

or equivalently

$$Z_t = 1 + \int_0^t \gamma_s Z_{s-}(c\, dB_s + dM_s). \tag{3.9}$$

In other words, the Brownian motion in the classical Black–Scholes setting has been replaced by the martingale part of the noise process Y. We saw in the proof of Lemma 3.1 that, in general,

$$Z_t = 1 + \int_0^t G_s Z_{s-}\, dB_s + \int_0^t \int_{\mathbb{R}} Z_{s-}[H(s, x) - 1]\, M(ds, dx).$$

Comparing this last expression with (3.9), we see that we require

$$H(s, x) - 1 = c^{-1} G_s x = h(s, x), \tag{3.10}$$

so that $\gamma_s = c^{-1} G_s$. [When $c = 0$, this just boils down to $G \equiv 0$, $H(s, x) - 1 = \gamma_s x$.] To obtain a martingale measure, we now use the martingale condition (3.7) together with (3.10). Putting

$$v = \int_{\mathbb{R}} x^2\, \nu(dx),$$

it is easily verified that the solution to (3.7) and (3.10) is

$$\begin{aligned} G_s &= \frac{c(r_s - b_s - a\sigma_s)}{\sigma_s(c^2 + v)}, \\ H(s, x) - 1 &= \left(\frac{r_s - b_s - a\sigma_s}{\sigma_s(c^2 + v)}\right) x. \end{aligned} \tag{3.11}$$

In (3.9), we therefore have

$$\gamma_s = \frac{r_s - b_s - a\sigma_s}{\sigma_s(c^2 + v)}. \tag{3.12}$$

Finally, we need some conditions to ensure that $H(s, \Delta X_s) > 0$; otherwise, the measure we have obtained will not be a probability measure but only a signed measure. Since we are assuming throughout this paper that the jump size $\Delta X \in [-c_1, c_2]$, we require the right-hand side of (3.11) of be greater than -1 for all $x \in [-c_1, c_2]$, which is equivalent to the condition that

$$-\frac{1}{c_2} < \left(\frac{r_s - b_s - a\sigma_s}{\sigma_s(c^2 + v)}\right) < \frac{1}{c_1}. \tag{3.13}$$

So far, we have done nothing more than show that one can obtain an equivalent martingale measure by drawing an obvious analogy with the classical Black–Scholes setting. It turns out, however, that the martingale measure given by (3.8), (3.9) and (3.12) is precisely the Föllmer–Schweizer minimal measure introduced in Föllmer and Schweizer (1991), which we shall proceed to show.

The minimal measure is closely connected to a hedging portfolio, which minimizes the risk involved in trying to duplicate a contingent claim Γ_T (provided such a portfolio exists). We briefly sketch the main ideas below, following closely the treatment in Föllmer and Schweizer (1991) but omitting some of the technical assumptions not essential to the exposition.

We adopt the notational convention that for any quantity f_t, the discounted quantity will be denoted by $\hat{f}_t = \exp\{-\int_0^t r_s\,ds\} f_t$. The value V_t of any hedging portfolio can be written as $V_t = \xi_t S_t + \eta_t \exp\{\int_0^t r_s\,ds\}$ and hence the discounted value is

$$\hat{V}_t = \xi_t \hat{S}_t + \eta_t,$$

where ξ and η are, respectively, the number of units of stock and bond. Only strategies for which $V_T = \Gamma_T$ $\mathbb{P}$-a.s. are admissible. Define the cumulative cost at time t by

$$C_t = \hat{V}_t - \int_0^t \xi_s\,d\hat{S}_s$$

and the remaining risk by

$$\mathbb{E}\big[(C_T - C_t)^2 \,\big|\, \mathscr{F}_t\big].$$

(In complete markets, C_t is constant and hence the risk is zero.) The idea is to look for strategies (ξ, η), which minimizes the remaining risk in a local sense: the risk is minimal under all "infinitesimal perturbations" of the strategy at time t. This is equivalent to the following precise technical definition.

DEFINITION 3.1. An admissible strategy (ξ, η) is called *optimal* if the associated cost C is a square-integrable martingale orthogonal to the martingale part (in the Doob decomposition) of $\hat{S}$ under $\mathbb{P}$.

Suppose now that the contingent claim Γ_T admits the following decomposition:

$$\hat{\Gamma}_T = \Gamma_0 + \int_0^T \xi_s \, d\hat{S}_s + L_T \tag{3.14}$$

for some ξ, where L_t is a square-integrable martingale orthogonal to the martingale part of $\hat{S}$ under $\mathbb{P}$. Then letting

$$\hat{V}_t = \Gamma_0 + \int_0^t \xi_s \, d\hat{S}_s + L_t,$$
$$\eta_t = \hat{V}_t - \xi_t \hat{S}_t,$$

we see that (ξ, η) is an optimal admissible strategy. Conversely, an optimal admissible strategy (ξ, η) gives a decomposition of the form (3.14) with $L_t = C_t - C_0$. Thus, the existence of an optimal strategy is equivalent to a decomposition of the form (3.14). Next, we introduce the idea of a minimal martingale measure.

DEFINITION 3.2. An equivalent martingale measure $\mathbb{Q}$ is called *minimal* if any square-integrable $\mathbb{P}$-martingale which is orthogonal to the martingale part of $\hat{S}$ under $\mathbb{P}$ remains a martingale under $\mathbb{Q}$.

Clearly, if an optimal strategy and a minimal equivalent martingale measure $\mathbb{Q}$ exist, we have $\hat{V}_t = \mathbb{Q}[\hat{\Gamma}_T \mid \mathscr{F}_t]$, which motivates taking $V_0 = \mathbb{Q}[\hat{\Gamma}_T]$ as the price of the contingent claim. However, we are not able to say anything about the existence of an optimal strategy in this paper.

It remains to show that the measure given by (3.8), (3.9) and (3.12) is minimal. The argument here is the same as that in Föllmer and Schweizer (1991) for the continuous case. Under $\mathbb{P}$, $\hat{S}$ satisfies

$$\begin{aligned} \hat{S}_t &= \hat{S}_0 + \int_0^t \sigma_s \hat{S}_{s-} (c \, dB_s + dM_s) + \int_0^t (a\sigma_s + b_s - r_s) \hat{S}_{s-} \, ds \\ &= \hat{S}_0 + W_t + A_t, \end{aligned} \tag{3.15}$$

where

$$W_t = \int_0^t \sigma_s \hat{S}_{s-} (c \, dB_s + dM_s)$$

is a $\mathbb{P}$-martingale and

$$A_t = \int_0^t (a\sigma_s + b_s - r_s) \hat{S}_{s-} \, ds$$

is a continuous adapted, and hence previsible, process. Therefore (3.15) gives the Doob decomposition of $\hat{S}$ under $\mathbb{P}$. Consider now a square-integrable $\mathbb{P}$-martingale N which is orthogonal to W, so that $\langle N, W \rangle = 0$. Hence $\langle N, Z \rangle = 0$ for the density Z given by (3.9), which implies that N is a $\mathbb{Q}$-local martingale. Since N and Z are both square-integrable $\mathbb{P}$-martingales, the Cauchy-Schwartz inequality shows that NZ is $L^1(\mathbb{P})$-bounded and so N is in fact a Q-martingale.

3.2. *Pricing by martingale decompositions.* In the continuous case, an important property of the minimal measure is that it gives $\hat{S}$ the law of its martingale part under the Doob decomposition; in fact, the minimal measure can be uniquely characterized by this property [see Föllmer and Schweizer (1991)]. It turns out that when $\hat{S}$ is discontinuous as in the present setting, this is no longer the case. On the other hand, the idea of using the martingale measure under which $\hat{S}$ has the law of its martingale part has a certain intuitive appeal. It is therefore interesting to ask which equivalent martingale measure will give $\hat{S}$ the law of its martingale part in our present setting and to compare it to the minimal measure.

Consider the Doob decomposition of $\hat{S}$ under $\mathbb{P}$ given by (3.15). Let $\mathbb{Q}$ be a martingale measure as described in Theorem 3.2, satisfying the martingale condition (3.7). Under $\mathbb{Q}$, $\hat{S}$ satisfies

$$\hat{S}_t = \hat{S}_0 + \int_0^t \sigma_s \hat{S}_{s-}\,(c\,d\tilde{B}_s + d\tilde{M}_s), \tag{3.16}$$

where $\tilde{M}$ is the $\mathbb{Q}$-martingale in (3.6) and $\tilde{B}$ is the $\mathbb{Q}$-Brownian motion described in Theorem 3.2. Specifically,

$$\tilde{M}_t = \int_0^t \int_{\mathbb{R}} x(\tilde{Q}(ds, dx) - \tilde{\nu}_s(dx)\,ds),$$

where $\tilde{Q}(ds, dx)$ is a nonhomogeneous Poisson measure with compensator measure $\tilde{\nu}_s(dx)$ given by Theorem 3.2. Therefore, comparing (3.16) with the form of W, we see that the only way in which $\hat{S}$ can have the law of W under $\mathbb{Q}$ is to have $\tilde{\nu} \equiv \nu$, which implies that $H \equiv 1$ and $h \equiv 0$ and $\mathbb{Q}$ must be given by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathscr{F}_T} = \exp\left\{\int_0^T G_s\,dB_s - \frac{1}{2}\int_0^T G_s^2\,ds\right\}, \tag{3.17}$$

where, from (3.7),

$$G_s = \frac{r_s - b_s - a\sigma_s}{c\sigma_s}. \tag{3.18}$$

However, for this to make sense, we require $c \neq 0$; in other words, the noise must have a Brownian component. If the driving Lévy process has no Brownian component, there is no martingale measure which would give $\hat{S}$ the law of its martingale part.

The most interesting observation one can make about the measure given by (3.17) and (3.18) is that it corresponds precisely to the classical Black–Scholes formula, if we treat $b + a\sigma$ as an overall drift which (3.17) removes by only changing the drift of the underlying Brownian motion B while leaving the jump part of the noise alone.

In the case of positive semimartingales, there is a multiplicative analogue of the Doob decomposition. If $\hat{S}$ is any positive semimartingale, then it can be written uniquely as $\hat{S}_t = \hat{S}_0 R_t C_t$, where R_t is a (local) martingale and C_t is a previsible process [see Liptser and Shiryayev (1989)]. For the discounted stock price $\hat{S}$, one can easily read off this decomposition from the formula (2.13); we have $\hat{S}_t = \hat{S}_0 R_t C_t$, where

$$R_t = \exp\left\{\int_0^t c\sigma_s\,dB_s + \int_0^t \sigma_s\,dM_s - \frac{1}{2}\int_0^t c^2\sigma_s^2\,ds\right\} \times \prod_{0<s\le t}(1+\sigma_s\Delta M_s)\exp(-\sigma_s\Delta M_s)$$

is a $\mathbb{P}$-martingale and

$$C_t = \exp\left\{\int_0^t (a\sigma_s + b_s - r_s)\,ds\right\}$$

is trivially previsible.

Elliott, Hunter, Kopp and Madan (1995) proposed pricing contingent claims using the martingale measure under which $\hat{S}$ has the law of its martingale part R under this multiplicative distribution. The motivation behind this is that, as we saw in Section 3.1, the minimal measure may in general only be a signed measure. In Elliott, Hunter, Kopp and Madan (1995) it is shown that, under certain regularity conditions, the multiplicative decomposition always results in a proper probability measure. Moreover, this approach has a natural intuitive appeal similar to that based on the Doob decomposition and, for our model, it actually gives the same answer, as we shall now demonstrate.

Let $\mathbb{Q}$ be any measure of the form described in Theorem 3.2 with G and H satisfying (3.7). Then under $\mathbb{Q}$,

$$\hat{S}_t = \hat{S}_0\exp\left\{\int_0^t c\sigma_s\,d\tilde{B}_s + \int_0^t \sigma_s\,d\tilde{M}_s - \tfrac{1}{2}\int_0^t c^2\sigma_s^2\,ds\right\} \times \prod_{0<s\le t}(1+\sigma_s\Delta\tilde{M}_s)\exp(-\sigma_s\Delta\tilde{M}_s).$$

Comparing this with the form of R, we again require $M \equiv \tilde{M}$ and $\mathbb{Q}$ to be the measure given by (3.17) and (3.18) in order for $\hat{S}$ to have the law of R under $\mathbb{Q}$. Also, for this to work there must again be a Brownian component.

3.3. *Pricing by minimum relative entropy and Esscher transform.* Gerber and Shiu (1994) proposed pricing contingent claims by Esscher transforms. Let $\theta \in \mathbb{R}$ be fixed. The Esscher transform of a Lévy process Y [satisfying (2.5)], or equivalently of its underlying canonical measure $\mathbb{P}$, is defined to be the process whose law $\mathbb{Q}_\theta$ is given by

$$\left.\frac{d\mathbb{Q}_\theta}{d\mathbb{P}}\right|_{\mathscr{F}_t} = \exp\{-\theta Y_t + t\psi(\theta)\},$$

where $\psi(\theta) = -\log\mathbb{E}[\exp(-\theta Y_1)]$ is the Lévy exponent of Y given by (2.8). If the stock price process has constant coefficients, which is the case for the model considered in Gerber and Shiu (1994), the value of θ can be chosen so as to make the discounted price process $\hat{S}$ a martingale under $\mathbb{Q}_\theta$. When the stock price process has time-dependent coefficients as in our model, we need to consider generalized Esscher transforms of the form

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_t} = \exp\left\{-\int_0^t \theta_s \, dY_s + \int_0^t \psi(\theta_s)\, ds\right\} \tag{3.19}$$

and choose θ_s to satisfy the martingale condition. Of course, as an alternative to (3.8) and (3.9), (3.19) is another natural analogue of the usual Girsanov transform used for the Black–Scholes formula. In the framework of Theorem 3.2 and Lemma 3.1, if $H(t, x) = \exp(-\theta_t x)$ and $h(t, x) = -\theta_t x$, then in (3.2), the product vanishes and moreover,

$$\int_0^t \int_{\mathbb{R}} h(s, x)\, M(ds, dx) = -\int_0^t \theta_s \, dM_s = -\int_0^t \theta_s \, dX_s + a\int_0^t \theta_s \, ds.$$

Therefore, comparing (2.8), (3.2) and (3.19), we see that the Esscher transform corresponds to the choices $H(t, x) = \exp(-\theta_t x)$, $h(t, x) = -\theta_t x$ and $G \equiv -c\theta$. The martingale condition (3.7) can then be used to specify θ as follows:

$$-c^2\sigma_s\theta_s + a\sigma_s + b_s - r_s + \int_{\mathbb{R}} \sigma_s x\big(\exp(-\theta_s x) - 1\big)\, \nu(dx) = 0. \tag{3.20}$$

[To see that this has a unique solution θ for which $\psi(\theta_s) < \infty$ for all s, define $F(\theta) = \int_{\mathbb{R}} x(e^{-\theta x} - 1)\, \nu(dx) - \theta$ for $\theta \in (-h_1, h_2)$, where h_1 and h_2 are as in (2.5). Then it is easy to check that F is monotonically decreasing and $F(\theta) \to +\infty$ as $\theta \downarrow -h_1$ and $F(\theta) \to -\infty$ as $\theta \uparrow h_2$. Hence equations of the form $F(\theta) = c$ have a unique solution in $(-h_1, h_2)$.]

The Esscher transform is a well-known tool in many actuarial applications [see Esscher (1932)]; Gerber and Shiu were the first to show that it can also be used in option pricing. Gerber and Shiu (1994) provide an interpretation of this approach in terms of maximal expected utility. There is, in fact, another useful interpretation of the Esscher transform; we shall show that it gives rise to the equivalent martingale measure which has minimum relative entropy with respect to $\mathbb{P}$. Intuitively speaking, if one thinks of the measure $\mathbb{P}$ as encapsulating some information about how the market behaves, then pricing options by Esscher transform amounts to choosing the equivalent martingale measure which is closest to $\mathbb{P}$ in terms of its information content.

For a fixed measure P, the relative entropy $I_P(Q)$ of any measure Q with respect to P is defined to be

$$I_P(Q) = \int \log \frac{dQ}{dP}\, dQ = \int \frac{dQ}{dP} \log \frac{dQ}{dP}\, dP.$$

[Note that $I_P(Q) \geq 0$ for any Q. If Q is not absolutely continuous with respect to P, $I_P(Q)$ is infinite.] For an equivalent martingale measure $\mathbb{Q}$ given by Theorem 3.2 and the martingale condition (3.7), the relative entropy in terms of the $\mathbb{Q}$-martingales $\tilde{B}$ and $\tilde{M}$ is therefore

$$\begin{aligned}
I_{\mathbb{P}}(\mathbb{Q}) &= \mathbb{Q}\left[\log \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathscr{F}_T}\right] \\
&= \mathbb{Q}\bigg[\int_0^T G_s\, d\tilde{B}_s + \frac{1}{2}\int_0^T G_s^2\, ds \\
&\qquad + \int_0^T h_s\, d\tilde{M}_s + \int_0^T \int_{\mathbb{R}} h_s x(H(s,x)-1)\,\nu(dx)\,ds \\
&\qquad - \int_0^T \int_{\mathbb{R}} (H(s,x)-1-h_s x)\,\nu(dx)\,ds \\
&\qquad\qquad + \int_0^T \int_{\mathbb{R}} (\log H(s,x)-h_s x)\,\tilde{\nu}_s(dx)\,ds\bigg] \\
&= \mathbb{Q}\left[\frac{1}{2}\int_0^T G_s^2\, ds + \int_0^T \int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx)\,ds\right].
\end{aligned}$$

[In the above calculation, we have written $h(s,x) = h_s x$ and used the results (2.9) and (2.11).]

The problem of finding the equivalent martingale measure of minimum relative entropy can clearly be reduced to that of minimizing

$$\mathbb{Q}\left[\tfrac{1}{2}G_s^2 + \int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx)\right]$$

for fixed s, subject to (3.7). Because the measure $\mathbb{Q}$ varies with each choice of G and H, it is a little less clear that the problem can be reduced further to that of minimizing

$$\tfrac{1}{2}G_s^2 + \int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx) \tag{3.21}$$

for each fixed s and ω, subject to (3.7). But it is clear that this last optimization problem has a deterministic solution in G and H, because all the coefficients σ, b and so on, are assumed to be deterministic. Denote by $\mathbb{Q}^*$ the measure associated with the optimal choice of G and H. The corresponding optimal value I^* of (3.21) is therefore also deterministic and for any other choice of G and H with associated measure $\mathbb{Q}^{G,H}$, we have

$$\tfrac{1}{2}G_s^2 + \int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx) \ge I^*$$

and hence

$$\mathbb{Q}^{G,H}\left[\tfrac{1}{2}G_s^2 + \int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx)\right] \ge I^* = \mathbb{Q}^*[I^*].$$

It might seem at first sight that the natural way to minimize (3.21) subject to (3.7) is to express G in terms of H using (3.7) and then substitute into (3.21). However, the quadratic term in G in (3.21) would result in the square of an integral involving H, which is difficult to handle. It turns out that the most convenient way to solve this optimization problem is first to fix G and choose H to minimize

$$\int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx)$$

subject to (3.7) and then minimize (3.21) (with the optimal H) over G.

Let λ_s be a continuous function and let

$$L(\lambda,H)=\int_{\mathbb{R}} \big[H(s,x)(\log H(s,x)-1)+1\big]\,\nu(dx)+\int_{\mathbb{R}} \lambda_s\sigma_s x(H(s,x)-1)\,\nu(dx).$$

Thus λ is a Lagrange multiplier associated with the constraint (3.7) and L is the associated Lagrangian. Observe that $H\mapsto L(\lambda,H)$ is convex in $H>0$, so to find the optimal H, we require

$$\frac{d}{dt}L(\lambda,H+tF)\bigg|_{t=0}=0$$

for all F, which gives

$$H(s,x)=\exp(-\lambda_s\sigma_s x).$$

The Lagrange multiplier λ can be expressed in terms of G (assumed to be fixed for the moment) via (3.7):

$$\int_{\mathbb{R}} \sigma_s x\big(\exp(-\lambda_s\sigma_s x)-1\big)\,\nu(dx)=r_s-b_s-a\sigma_s-\sigma_s G_s.$$

Since all the optimization is carried out for fixed s, we temporarily drop the explicit dependence on s for the sake of clarity. The previous equation is then simply

$$\int_{\mathbb{R}} \sigma x(e^{-\lambda\sigma x}-1)\,\nu(dx)=r-b-a\sigma-c\sigma G. \tag{3.22}$$

Putting $H\equiv e^{-\lambda\sigma x}$ into (3.21) gives

$$\tfrac{1}{2}G^2+\int_{\mathbb{R}}[1-e^{-\lambda\sigma x}(\lambda\sigma x+1)]\,\nu(dx), \tag{3.23}$$

which we must now minimize over G. We simply differentiate the above with respect to G and solve

$$G+\lambda'(G)\int_{\mathbb{R}} \lambda(\sigma x)^2 e^{-\lambda\sigma x}\,\nu(dx)=0. \tag{3.24}$$

However, differentiating (3.22) shows that

$$\lambda'(G)=c\sigma\bigg(\int_{\mathbb{R}}(\sigma x)^2 e^{-\lambda\sigma x}\,\nu(dx)\bigg)^{-1}$$

and substituting into (3.24) gives $G=-c\sigma\lambda$. [Similarly, we find the second derivative of (3.23) is $1+\lambda'(G)^2\int(\sigma x)^2 e^{-\lambda\sigma x}\,\nu(dx)>0$, so $G=-c\sigma\lambda$ does indeed give the minimum of (3.23).] Thus both G and H are now specified in terms of λ, and restoring the s in (3.22) gives the equation for λ_s,

$$-c^2\sigma_s^2\lambda_s + a\sigma_s + b_s - r_s + \int_{\mathbb{R}} \sigma_s x\big(\exp(-\lambda_s\sigma_s x) - 1\big)\,\nu(dx) = 0. \tag{3.25}$$

Comparing (3.25) and (3.20), we see that this is precisely the measure constructed via the Esscher transform, with $\theta \equiv \sigma\lambda$.

As one of the main motivations behind the study of the Esscher transform presented here is Gerber and Shiu (1994), it is interesting to see if similar results hold for the model of stock price used in that paper, namely,

$$S_t = S_0 \exp\{\sigma Y_t + bt\} \tag{3.26}$$

for constants σ and b. For simplicity, we shall also take $r \equiv 0$, so that $S = \hat{S}$. The Esscher transform of Y is exactly the same as before; in the context of Theorem 3.2 we have $H(t,x) = e^{-\theta x}$, $h(t,x) = \theta x$ and $G \equiv c\theta$ where θ is the (constant) Esscher parameter appearing in (3.19). However, whereas for our model the martingale condition used to specify θ is (3.7), a different martingale condition applies to the Gerber–Shiu model (3.26). Since

$$\begin{aligned} S_t &= S_0 \exp\{c\sigma B_t + \sigma M_t + a\sigma t + bt\} \\ &= S_0 \exp\bigg\{c\sigma\tilde{B}_t + \sigma\tilde{M}_t + a\sigma t + bt + c\sigma\int_0^t G_s\,ds \\ &\qquad\qquad + \sigma\int_0^t\int_{\mathbb{R}} x(H(s,x)-1)\,\nu(dx)\,ds\bigg\}. \end{aligned}$$

Itô's formula gives

$$\begin{aligned} dS_t &= \sigma S_{t-}\,d\tilde{B}_t + \sigma S_{t-}\,d\tilde{M}_t \\ &\quad + S_{t-}\bigg\{c\sigma G_t + a\sigma + b + \frac{c^2\sigma^2}{2} + \sigma\int_{\mathbb{R}} x(H(s,x)-1)\,\nu(dx)\bigg\}\,dt \\ &\quad + S_{t-}\big\{\exp(\sigma\Delta\tilde{M}_t) - 1 - \sigma\Delta\tilde{M}_t\big\} \\ &= \sigma S_{t-}\,d\tilde{B}_t + \sigma S_{t-}\,d\tilde{M}_t + S_{t-} \\ &\quad \times\bigg\{c\sigma G_t + a\sigma + b + \frac{c^2\sigma^2}{2} + \sigma\int_{\mathbb{R}} x(H(s,x)-1)\,\nu(dx)\bigg\}\,dt \\ &\quad + S_{t-}\int_{\mathbb{R}}\big(\exp(\sigma x) - 1 - \sigma x\big)\,\tilde{M}(dt,dx) \\ &\quad + S_{t-}\int_{\mathbb{R}}\big(\exp(\sigma x) - 1 - \sigma x\big)\,\tilde{\nu}(dx)\,dt. \end{aligned}$$

Hence, in order for $S = \hat{S}$ to be a martingale under $\mathbb{Q}$, we require

$$\begin{aligned} &c\sigma G_t + a\sigma + b + \frac{c^2\sigma^2}{2} + \sigma\int_{\mathbb{R}} x(H(t,x)-1)\,\nu(dx) \\ &\qquad + \int_{\mathbb{R}}(e^{\sigma x} - 1 - \sigma x)H(t,x)\,\nu(dx) = 0. \end{aligned} \tag{3.27}$$

In terms of θ, this translates into

$$-c^2\sigma\theta + a\sigma + b + \frac{c^2\sigma^2}{2} + \int_{\mathbb{R}} \big[(e^{\sigma x} - 1)e^{-\theta x} - \sigma x\big]\,\nu(dx) = 0. \tag{3.28}$$

Turning now to the minimum relative entropy measure, we need to minimize (3.21) subject to (3.27). Following exactly the same Lagrangian procedure as before, we find

$$H(x) = \exp\{\lambda(1 - \exp(\sigma x))\},$$
$$G \equiv -c\sigma\lambda$$

(where λ is the Lagrange multiplier). Substituting this into (3.27) gives

$$\begin{aligned} &-c^2\sigma^2\lambda + a\sigma + b + \frac{c^2\sigma^2}{2} \\ &\qquad + \int_{\mathbb{R}} \big[(\exp(\sigma x) - 1)\exp\{\lambda(1 - \exp(\sigma x))\} - \sigma x\big]\,\nu(dx) = 0. \end{aligned} \tag{3.29}$$

Comparing (3.29) with (3.28), we see that it is no longer possible to make the identification $\theta = \sigma\lambda$ and that (3.28) and (3.29), although similar, are in fact different equations. Thus, the Esscher transform for the model (3.26) does not correspond to the measure of minimum relative entropy. However, the two are sufficiently similar to make one further worthwhile observation. If Y only makes very small jumps, its Lévy measure ν is concentrated around 0. Making the approximation

$$1 - e^{\sigma x} \sim -\sigma x \quad \text{for small } x$$

in the integrand of (3.29), we see that the solutions to (3.28) and (3.29) can be approximated to some extent by $\theta \approx \sigma\lambda$. Therefore, the Esscher transform gives a measure which has approximately minimum relative entropy in this sense. This is not surprising since the models (2.12) and (3.26) are identical (but with different σ and b) if Y is Brownian motion, and so they should be close in some sense when the jumps of Y are small.

4. Integro-differential equations for the valuation process. Consider any contingent claim whose payoff depends only on the value at maturity of the underlying security. Thus the payoff at time T can be written as $\Gamma_T = g(S_T)$ for some function g. The typical example is of course a European call with strike price K, for which $g(x) = (x - K)^+$. Let

$$V_t = \mathbb{Q}\bigg[\exp\bigg\{-\int_t^T r_s\,ds\bigg\}g(S_T)\bigg|\mathscr{F}_t\bigg]$$

be the value of the contingent claim at time t.

Recall that in the classical Brownian setting, the valuation process V admits a Feynman–Kac type representation $V_t = u(t, S_t)$, where u is the solution to the Cauchy problem associated with a linear PDE [see, e.g., Karatzas and Shreve (1991)]. The same approach will carry over to the present setting without difficulty, the only difference is that the resulting equation is an integro-differential equation, the integral term being associated with the jumps in the Lévy process. As the essentials of the argument are standard and well known, we shall simply sketch the derivation and leave the details to the reader.

Recall that, under $\mathbb{Q}$, the price of the underlying stock satisfies

$$dS_t = \sigma_t S_{t-}(c\, d\tilde{B}_t + d\tilde{M}_t) + r_t S_{t-}\, dt.$$

Let $\mathscr{G}_t$ be the following linear integro-differential operator:

$$\mathscr{G}_t f(x) = \tfrac{1}{2}c^2\sigma_t^2 x^2 f''(x) + r_t x f'(x) + \int_{\mathbb{R}} \big[f(x+\sigma_t xy) - f(x) - \sigma_t xy f'(x)\big]\,\tilde{\nu}_t(dy).$$

Of course, $\mathscr{G}_t$ is the generator of S_t, as Itô's formula applied to $f(S_t)$ will readily verify. For functions $u(t, x)$, we use the short-hand notation $\dot{u} \equiv \partial u/\partial t$ and $u' \equiv \partial u/\partial x$.

THEOREM 4.1. *Let $u(t, x)$ be the solution to the following Cauchy problem*:

$$\dot{u} + \mathscr{G}_t u - r_t u = 0, \qquad u(T, x) = g(x). \tag{4.1}$$

Then u admits the representation

$$u(t, x) = \mathbb{Q}^{t,x}\left[\exp\left\{-\int_t^T r_s\, ds\right\} g(S_T)\right].$$

PROOF. For fixed t, simply apply Itô's formula to the process $t' \mapsto u(t', S_{t'})\exp\{-\int_t^{t'} r_s\, ds\}$ to show that it is a $\mathbb{Q}$-martingale, and then take its $\mathbb{Q}$-expectation. □

The Markov property combined with the above result shows that $V_t = u(t, S_t)$. Equation (4.1) can then be solved numerically as a practical way to compute the price of the option.

5. Numerical examples. Throughout this section, we take as parameters of our model $\sigma \equiv b \equiv c = 1$ and $r \equiv 0$. We calculate the price of a European call option with strike price $K = 1$ and maturity at time $t = 1$ for various values of the initial share price S_0, using each of the martingale measures discussed in Section 3.

As a first example, we let $X_t = (N_1(t) - N_2(t))/2$ where N_1 and N_2 are independent Poisson processes of rate 1. For this process, the Lévy measure is $\nu = \delta_{-1/2} + \delta_{+1/2}$ and the previsible part is $a = \mathbb{E}[X_1] = 0$. Also (2.14) is satisfied, so share prices S_t are nonnegative almost surely.

For the minimal measure we have, from (3.11),

$$H(x) = 1 - \tfrac{2}{3}x,$$

which satisfies the nonnegativity condition (3.13). Under this measure, N_1 and N_2 are Poisson processes with rates 2/3 and 4/3, respectively, and $\tilde{M}_t = X_t + t/3$ is a $\mathbb{Q}$-martingale.

TABLE 1
Prices based on Poisson process model

S_0	Minimal measure	Black–Scholes measure	Esscher transform measure
0.50	0.135	0.242	0.152
0.75	0.281	0.480	0.304
1.00	0.451	0.749	0.482
1.25	0.636	1.035	0.676
1.50	0.832	1.335	0.880

Under the measure of Section 3.2, the law of X remains unchanged and only the drift of the Brownian part changes. For this reason, we shall henceforth refer to this as the Black–Scholes measure.

For the Esscher transform measure, the solution to (3.20) is $\theta \approx 0.6626$. Under this measure, N_1 and N_2 are Poisson processes with rates $e^{-\theta/2}$ and $e^{\theta/2}$, respectively, and so $\tilde{M}_t = X_t + t\sinh(\theta/2)$ is a $\mathbb{Q}$-martingale.

The processes involved are very simple to simulate; the prices in Table 1 are based on 5000 simulations.

As a further example, we take X to be a Gamma(1,1) process whose law is given by

$$\mathbb{E}\big[\exp(-\lambda X_t)\big] = \left(\frac{1}{1+\lambda}\right)^t.$$

For this process, the Lévy measure is $\nu(dx) = x^{-1}e^{-x}1_{[0,\infty)}(x)\,dx$ and the previsible part is $a = \mathbb{E}[X_1] = 1$. Since the Lévy measure is supported on $[0,\infty)$, there is no sensible choice of parameters which would satisfy (3.13) because in practice one should have $b + a\sigma > r$. We shall therefore not consider the minimal measure for this process. For the Esscher transform measure, the solution to (3.20) is $\theta = \sqrt{2}$. Under this measure, X is still a Gamma process, but now with shape parameter 1 and scale parameter $\theta + 1$, that is,

$$\mathbb{Q}_\theta\big[\exp(-\lambda X_t)\big] = \left(\frac{\theta+1}{\theta+1+\lambda}\right)^t.$$

Although a Gamma process is not too difficult to simulate, the prices in Table 2 are obtained by numerically solving (4.1) rather than by simulating the

TABLE 2
Prices based on the Gamma process model

S_0	Black–Scholes measure	Esscher transform measure
0.50	0.149	0.107
0.75	0.295	0.240
1.00	0.465	0.403
1.25	0.653	0.587
1.50	0.852	0.785

various processes as in the first example. Table 2 gives the values of $u(0, S_0)$, where u is the solution to (4.1).

6. Concluding remarks. We have focussed on only three possible approaches to option pricing under the framework of stock prices driven by Lévy processes. There are many other approaches which have been suggested by various authors. For most of these, explicit pricing formulas of the type presented here should be obtainable. For example, Davis (1994) presents an approach to option pricing based on utility maximization. It is possible to characterize explicitly the associated martingale measure (in the framework of Theorem 3.2) for a large class of utility functions—for example x^γ/γ for $0 < \gamma < 1$ and $\log x$. Unfortunately, none of these utility functions gives rise to any obviously "nice" martingale measures such as the ones discussed here. In this regard, it is worth mentioning that Gerber and Shiu (1994) have noted in their discussion that, for their model, the Esscher price corresponds to the price obtained by the Davis utility approach if one takes as the utility function x^γ/γ, where γ is a function of the Esscher parameter, satisfying an equation similar to (3.20). The main drawback of this idea is that the choice of utility function depends not only on the investor's risk-averseness or other aspects of the investor's preference, but also on the market itself since the Esscher parameter depends on the market parameters. In any case, for the model studied here based on (2.12), the Esscher transform does not admit such an interpretation.

Acknowledgments. This work was carried out while I was visiting the University of Adelaide. I thank the Applied Mathematics Department for its hospitality and John van der Hoek in particular for useful discussions.

REFERENCES

AASE, K. K. (1988). Contingent claims valuation when the security price is a combination of an Itô process and a random point process. *Stochastic Process. Appl.* **28** 185–220.

BARDHAN I. and CHAO, X. (1993). Pricing options on securities with discontinuous returns. *Stochastic Process. Appl.* **48** 123–137.

DAVIS, M. H. A. (1994). A general option pricing formula. Preprint, Imperial College, London.

ELLIOTT, R. J., HUNTER, W. C., KOPP, P. E. and MADAN, D. B. (1995). Pricing via multiplicative price decomposition. *J. Financial Engineering* **4** 247–262.

ESSCHER, F. (1932). On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift* **15** 175–195.

FÖLLMER, H. and SCHWEIZER, M. (1991). Hedging of contingent claims under incomplete information. In *Applied Stochastic Analysis* (M. H. A. Davis and R. J. Elliott, eds.) 389–414. Gordon and Breach, New York.

GERBER, H. U. and SHIU, E. S. W. (1994). Option pricing by Esscher transforms (with discussion). *Trans. Soc. Actuaries* **46** 99–191.

HARRISON, J. M. and PLISKA, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.* **11** 215–260.

HARRISON, J. M. and PLISKA, S. R. (1983). A stochastic calculus model of continuous trading: Complete markets. *Stochastic Process. Appl.* **15** 313–316.

JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.

KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, New York.

LIPTSER, R. SH. and SHIRYAYEV, A. N. (1989). *Theory of Martingales*. Kluwer, Dordrecht.

PROTTER, P. (1990). *Stochastic Integration and Differential Equations*. Springer, New York.

DEPARTMENT OF ACTUARIAL MATHEMATICS
AND STATISTICS
HERIOT-WATT UNIVERSITY
EDINBURGH EH14 4AS
UNITED KINGDOM
E-MAIL: t.chan@ma.hw.ac.uk