THE TERM STRUCTURE OF SIMPLE FORWARD RATES WITH JUMP RISK

Paul Glasserman

Graduate School of Business, Columbia University, New York

S. G. Kou

IEOR Department, Columbia University, New York

This paper characterizes the arbitrage-free dynamics of interest rates, in the presence of both jumps and diffusion, when the term structure is modeled through simple forward rates (i.e., through discretely compounded forward rates evolving continuously in time) or forward swap rates. Whereas instantaneous continuously compounded rates form the basis of most traditional interest rate models, simply compounded rates and their parameters are more directly observable in practice and are the basis of recent research on "market models." We consider very general types of jump processes, modeled through marked point processes, allowing randomness in jump sizes and dependence between jump sizes, jump times, and interest rates. We make explicit how jump and diffusion risk premia enter into the dynamics of simple forward rates. We also formulate reasonably tractable subclasses of models and provide pricing formulas for some derivative securities, including interest rate caps and options on swaps. Through these formulas, we illustrate the effect of jumps on implied volatilities in interest rate derivatives.

Key Words: interest rate models, interest rate derivative securities, jump-diffusion models

1. INTRODUCTION

This paper characterizes the arbitrage-free dynamics of interest rates, in the presence of both jumps and diffusion, when the term structure is modeled through simple forward rates—that is, through discretely compounded forward rates evolving continuously in time—or through forward swap rates. We consider very general types of jump processes (allowing randomness in jump sizes and dependence between jump sizes, jump times, and interest rates) and identify how jump and diffusion risk premia enter into the dynamics of simple forward rates. We also formulate a reasonably tractable subclass of models and provide pricing formulas for some term structure derivatives.

Our investigation builds on several strands of research, in particular on the dynamics of instantaneous continuously compounded rates (as in Heath, Jarrow, and Morton 1992), option pricing with jumps (as in Merton 1976), LIBOR and swap rate market models (including Brace, Gatarek, and Musiela 1997; Jamshidian 1997; Miltersen,

Sandmann, and Sondermann 1997; Musiela and Rutkowski 1997a), and especially the marked point process framework of Björk, Kabanov, and Runggaldier (1997). The motivation for models based on simple forwards (in contrast to the instantaneous rates traditionally treated in continuous-time models) lies in building a model based on observable quantities. Most market rates are indeed based on simple compounding, so instantaneous continuously compounded rates often represent an idealized approximation to market data. This point is relevant whether one tries to infer model parameters from time-series data or from prices of derivative securities because most derivatives contracts are tied to simple rates.

Motivation for including jumps comes from both time-series properties and derivative prices. Specific sources of jumps in interest rates, including economic news and moves by central banks, are put forward in Babbs and Webber (1997), Das (1999b), El-Jahel, Lindberg, and Perraudin (1997), and Johannes (2003). These studies find compelling empirical evidence for jumps. Das (1999b) and Johannes argue that the kurtosis in short-term interest rates is incompatible with a pure-diffusion model. Jumps in interest rates can also be used to try to reproduce the patterns in implied volatilities derived from market prices of interest rate derivatives. The pricing of interest rate derivatives in the presence of jumps is considered in Björk et al. (1997), Burnetas and Ritchken (1997), Das (1999a), Das and Foresi (1996), Duffie and Kan (1996), Duffie, Pan, and Singleton (2000), Jarrow and Madan (1995, 1999) and Shirakawa (1991). The possibility of default (as modeled in Duffie and Singleton 1999 and Jarrow and Turnbull 1995) provides further motivation for including jumps, though we do not consider credit risk here.

Implied volatilities extracted from interest rate caps are putative parameters of simple forward rates, which again motivates adopting simple forwards as the building blocks of a model. (Similarly, implied volatilities extracted from options on interest rate swaps are putative parameters of forward swap rates.) In special cases of the general framework we develop, interest rate caps or swaptions can be priced explicitly, making it possible to investigate what types of patterns in implied volatility can be produced through jumps. The general framework is necessary for the formulation of tractable special cases: it turns out that for caps to be priced using a Poisson-based formula, the actual process of jumps must be substantially more complex than a Poisson process. The additional complexity needed follows from general considerations on precluding arbitrage; in particular, the Poisson property is not in general preserved by the necessary changes of measure.

The rest of this paper is organized as follows. Section 2 develops further motivation and background on modeling simple forward rates and on representing jump processes. Section 3 presents our main results: a general formulation of the arbitrage-free dynamics of simple forwards subject to jumps, and reduction to a tractable subclass. Section 4 presents some pricing formulas and numerical results on implied volatilities. Section 5 undertakes a similar analysis based on swap rates rather than forward rates: we present the arbitrage-free dynamics of the term structure of swap rates with both jumps and diffusion and then provide pricing formulas for options on swaps. All proofs are collected in the Appendixes.

2. MOTIVATION AND BACKGROUND

2.1. Simple Forwards

As in Brace et al. (1997), Jamshidian (1997), and Miltersen et al. (1997), we consider models of the term structure based on simple forward rates with a fixed accrual period δ, expressed as a fraction of a year (e.g., to model 3-month rates we would take $\delta = 1/4$).

With δ fixed, we denote by $L(t, T)$ the forward rate for the interval from T to $T + \delta$ as of time $t \leq T$. Thus, a party entering into a contract at time t to borrow \$1 over the interval $[T, T + \delta]$ will receive \$1 at time T and will return to the lender $\$(1 + \delta L(t, T))$ at time $T + \delta$. Denoting by $B(t, \tau)$ the time-t price of a zero-coupon bond maturing at τ, the forward rate satisfies

$$(2.1) \qquad L(t, T) = \frac{1}{\delta} \left(\frac{B(t, T)}{B(t, T + \delta)} - 1 \right).$$

Conversely, for any $k = 1, 2, \ldots,$

$$(2.2) \qquad B(t, t + k\delta) = \prod_{i=0}^{k-1} \frac{1}{1 + \delta L(t, t + i\delta)}.$$

Simple compounding of this type is characteristic of 3-month or 6-month LIBOR. We will, however, treat the forward rates and associated bonds as default-free, though in practice LIBOR reflect some credit risk.[1]

These simple forward rates should be contrasted with the instantaneous, continuously compounded short rate of classical models and also with the instantaneous forward rates modeled in the framework of Heath et al. (1992). The instantaneous forwards $f(t, T)$ of the Heath-Jarrow-Morton framework satisfy

$$(2.3) \qquad L(t, T) = \frac{1}{\delta} \left(\exp \left\{ \int_T^{T+\delta} f(t, s)\, ds \right\} - 1 \right),$$

but this relation cannot in general be inverted, so the distinction is not simply one of choice of variables. Arbitrage-free models based on simple forwards have been advanced by Brace et al. (1997), Jamshidian (1997), and Miltersen et al. (1997), and this work has given rise to a rapidly expanding related literature. Among other attractive features, these models are based on quantities that are more directly observable in the market than are the instantaneous rates of much of the earlier literature.

Working with simple forward rates often facilitates calibration to derivatives prices, in particular caps and floors. The information about the underlying forward rates in the market prices of caps and floors is commonly summarized through an implied volatility derived from the (so-called) Black (1976) formula (see our equation (4.1) in Section 4). These implied volatilities are frequently used as inputs to models for pricing other derivatives. In more detail, a caplet for the period $[T, T + \delta]$ struck at K pays $\delta(L(T, T) - K)^+$ at $T + \delta$. The Black formula may be viewed as evaluating the discounted expected payoff

$$(2.4) \qquad B(0, T + \delta) E[\delta(L(T, T) - K)^+],$$

under the assumption that $L(T, T)$ is lognormally distributed with mean $L(0, T)$ and $\log L(T, T)$ having variance $\sigma_T^2 T$. The implied volatility is the value of σ_T that equates (2.4) to the market price.

A simple way to introduce dynamics that yield a lognormal distribution for $L(t, T)$ specifies

$$(2.5) \qquad \frac{dL(t, T)}{L(t, T)} = \sigma_T \, dW_t,$$

[1] Miltersen et al. note that, through results of Duffie and Singleton (1999), their model can be used to represent defaultable interest rates under appropriate assumptions and with some redefinition of terms. A similar reinterpretation should be possible in our setting as well. See also Duffie et al. (2000).

with W_t a standard Brownian motion. The absence of a drift in this specification further implies that the conditional expectation of $L(T, T)$ at time t is $L(t, T)$, as is implicit in the Black formula. It turns out, however, that a model specifying (2.5) for all T fails to be arbitrage free. More precisely, there is no probability measure under which forward rates for all maturities simultaneously evolve according to (2.5) in an arbitrage-free model.

Despite this apparent inconsistency, Brace et al. (1997), Jamshidian (1997), Miltersen et al. (1997), and Musiela, and Rutkowski (1997a) were nevertheless able to construct arbitrage-free models of the term structure in which cap prices indeed conform to the Black formula. The models are, in effect, kept arbitrage free through inclusion of an appropriate stochastic drift in (2.5) for each maturity T. The forward rates are thus not simultaneously lognormal, but each becomes lognormal under a maturity-specific change of measure. Each such change of measure is associated with a change of numeraire which further serves to justify discounting by a zero-coupon bond. These ideas are discussed in greater detail in Section 3 and Appendix B.

On one hand, these models provide a theoretical basis for the market convention of quoting or interpreting cap prices through the Black formula; on the other hand, they also make evident an incompatibility between market prices and the models intended to explain them. For in these models the same implied volatility should apply to all caps and floors of a given maturity, regardless of strike price, whereas volatilities implied by market prices vary systematically with strike. This *volatility skew* is particularly pronounced in the Japanese market, but is also present in the US dollar market.

There are various means by which one might try to incorporate an implied volatility skew. These include adding a stochastic volatility, changing from a lognormal to constant elasticity of variance (CEV) form of volatility (as in Andersen and Andreasen 2000), or allowing for jumps. Empirical evidence in equity markets (Bakshi, Cao, and Chen 1997; Bates 2000; Das and Sundaram 1999) suggests that both jumps and stochastic volatility play an important role in the implied volatility skew observed there. It is therefore natural to investigate how jumps can be incorporated in a model of simple forwards.

A naive extension of the naive "Black model" in (2.5) specifies

$$(2.6) \quad dL(t, T) = -\lambda m L(t, T)\, dt + \sigma_T L(t, T)\, dW_t + L(t-, T) d\left(\sum_{i=1}^{N_t}(Y_i - 1)\right),$$

where N_t is a Poisson process with arrival rate λ and the Y_i are i.i.d. lognormal random variables with mean $1 + m$. (By writing $L(t-, T)$ we specify the value of $L(\cdot, T)$ just before a possible jump at t.) This is a jump-diffusion of the type considered by Merton (1976) as a model of a stock price, with the drift modified to make $L(t, T)$ a martingale. The marginal distributions of $L(t, T)$ under (2.6) are Poisson mixtures of lognormal distributions. "Pricing" a caplet according to (2.4) therefore results in a "Merton-Black formula,"

$$(2.7) \qquad\qquad \sum_{k=0}^{\infty} e^{-\lambda T}\frac{(\lambda T)^k}{k!}\mathrm{BC}_k,$$

where each BC_k is an evaluation of the Black formula but with arguments depending on k; this will be made explicit in Corollary 4.1. This pricing formula is nearly as tractable as the Black formula. Moreover, if caplets are priced according to (2.7), their Black-implied volatilities will vary with strike. Indeed, by varying the parameters of (2.6) it is possible to reproduce a variety of patterns in implied volatilities as functions of strike; see Section 4.1.

This simple example serves to motivate the questions we investigate: Is (2.6) consistent with an arbitrage-free model of the term structure? Can the naive pricing formula (2.7) be reconciled with a genuine pricing model? More generally, when is a jump-diffusion model of the term structure of simple forward rates arbitrage-free? We provide answers to these questions (and their analogs for forward swap rates) in subsequent sections.

2.2. Modeling Jumps

Addressing these questions requires an investigation of the dynamics of the term structure with respect to different choices of numeraire asset and under the associated probability measures. This in turn requires consideration of jump processes more general than the compound Poisson process appearing in (2.6). The marked point process framework developed by Björk et al. (1997) provides a convenient framework.

A marked point process (MPP) is characterized by a sequence $\{(\tau_n, X_n), n = 1, 2, \ldots\}$. The τ_n take values in $(0, \infty)$ and satisfy $\tau_1 < \tau_2 < \cdots < \tau_n < \tau_{n+1} < \cdots$, $\sup_n \tau_n = \infty$; interpret these as the times of *potential* jumps. The *marks* X_n may in general take values in an abstract space; we will use them to determine the sizes of the jumps at the points τ_n, though they are not themselves the jump sizes. Forward rates of different maturities may respond to the marks with jumps of different magnitudes. For our purposes, it will suffice to consider marks taking values in $[0, \infty)$.

To construct a jump process, first let N_t be the number of points in $[0, t]$: $N_t = \sup\{n \geq 0 : \tau_n \leq t\}$. Let h be a real-valued function of the marks (and possibly also of the points) and consider the jump process $J(t) = \sum_{n=1}^{N_t} h(X_n, \tau_n)$. The function h transforms the abstract mark X_n into a jump magnitude. In (2.6), it takes the form $h(x, \tau) \equiv h(x) = x - 1$.

We construct our models on a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geq 0\}, \mathsf{P})$ on which are defined a multidimensional Brownian motion W and r marked point process $\{(\tau_n^{(i)}, X_n^{(i)}), n = 1, 2, \ldots\}, i = 1, 2, \ldots, r$, not necessarily independent of each other or the Brownian motion. With each forward rate we associate jump-size functions $H_i, i = 1, \ldots, r$, and define

$$(2.8) \qquad J(t) = \sum_{i=1}^{r} \sum_{n=1}^{N_t^{(i)}} H_i\big(X_n^{(i)}, \tau_n^{(i)}\big),$$

with $N_t^{(i)}$ the counting process associated with the ith marked point process. The dynamics of a forward rate $L(t, T)$ take the form

$$(2.9) \qquad dL(t, T) = \alpha(t)L(t, T)\, dt + \gamma(t)L(t, T)\, dW(t) + L(t-, T)\, dJ(t),$$

for adapted processes α and γ satisfying regularity conditions. The r marked point processes in (2.8) can be dependent on each other but we require that the jump times $\tau_n^{(i)}, n \geq 1, 1 \leq i \leq r$, be distinct (this is needed for a generalization of the Girsanov theorem).

We assume that each marked point process $\{(\tau_n, X_n)\}$ has an *intensity* $\lambda(dx, t)$. Intuitively, $\lambda(dx, t)$ is the arrival rate of points with marks in dx. More precisely, the intensity has the property that, for all suitably integrable h,

$$(2.10) \qquad \sum_{n=1}^{N_t} h(X_n, \tau_n) - \int_0^t \int_0^\infty h(x, s)\lambda(dx, s)\, dt_s$$

is a martingale in t, the inner integral being over the mark space. (The key assumption here is that the arrival rate is absolutely continuous in time; otherwise, in addition to the dt term in (2.9) we would need a $\lambda(dx, dt)$ term.) For a marked point process in which the points follow a Poisson process and the marks are i.i.d. random variables (as would be the case in the Merton model (2.6)) the intensity takes the form

$$\lambda(dx, t) = \lambda \cdot f(x)\,dx,$$

where λ is the Poisson arrival rate and f is the common density of the marks.

A valuable feature of modeling jumps through marked point processes and their intensities arises in considering term structure dynamics under different probability measures. In a pure-diffusion setting, changing probability measures typically corresponds to adding a drift to a Brownian motion. In a model with jumps, changing probability measures can involve changing the jump intensity as well. We will see that even if we want the jumps in a forward rate to follow a Poisson process under one measure, we have no choice but to suppose that they follow a more general marked point process under other measures.

3. MODEL CONSTRUCTION

We now proceed to investigate conditions under which a term structure model of the general form (2.9) is consistent with the absence of arbitrage. The main task lies in identifying the appropriate form of the risk premium determining the drift in (2.9), once the other parameters have been specified.

A prerequisite to this investigation is a precise notion of what it means for a model to be arbitrage free in the presence of jumps, which further entails defining a class of admissible trading strategies. Our objective is the construction of models formulated purely in terms of simple forwards and their parameters; one could in principle specify a class of admissible trading strategies and develop the associated theory in this setting. Rather than make such a digression here, we choose a more efficient and only slightly less general route: We define a model of the term structure of simple forwards $L(t, T)$ to be arbitrage free if it can be embedded in an arbitrage-free model of instantaneous forwards $f(t, T)$ via (2.3). The necessary theory for instantaneous forwards has been developed by Björk et al. (1997), so we may invoke their results. We stress, however, that the models we construct are purely models of simple forwards and make no reference to hypothetical instantaneous forwards. Indeed, the instantaneous forwards appear nowhere in the rest of this section. Jamshidian (1999) has recently developed a model of simple forwards driven by very general discontinuous semimartingales; his framework does not use underlying instantaneous rates but rather works with simple rates throughout.

To simplify both the analysis and notation, we formulate our results in a discrete-tenor setting in which the maturity T is restricted to a finite set of dates $0 = T_0 < T_1 < \cdots < T_M < T_{M+1}$. (In Appendix A we prove an intermediate result that does hold simultaneously for all maturities T and from which we prove Theorem 3.1.) We will further assume that the intervals $T_{i+1} - T_i$ are equally spaced with a common spacing of δ (e.g., a quarter year or a half year). Let $L_n(t) = L(t, T_n)$ so that L_n is the forward rate for the accrual period $[T_n, T_{n+1}]$. Similarly, let $B_n(t) = B(t, T_n)$ denote the price of a zero coupon bond maturing at T_n. Let $\eta(t) = \inf\{k \geq 0 : T_k \geq t\}$ so that $\eta(t)$ is the index of the next maturity as of time t.

The results of this section are proved in Appendixes A and B under regularity conditions. Ideally, all conditions would be made explicit in the statements of the results. As it

does not seem possible to do this succinctly, we leave a discussion of the implicit technical conditions to Appendix A.

3.1. General Case

We are now ready to formulate our first main result. As in Section 2.2, our building blocks are a d-dimensional Brownian motion $W(t)$ and r marked point processes $\{(\tau_j^{(i)}, X_j^{(i)}), j = 1, 2, \ldots\}, i = 1, \ldots, r$, with intensities $\lambda^{(i)}$.

THEOREM 3.1.   *For each $n = 1, \ldots, M$ let $\gamma_n(\cdot)$ be a bounded, adapted, R^d-valued process and let $H_{ni}, i = 1, \ldots, r$, be deterministic functions from $[0, \infty)$ to $[-1, \infty)$. The model*

$$(3.1) \quad \frac{dL_n(t)}{L_n(t-)} = \alpha_n(t)\,dt + \gamma_n(t)\,dW(t) + dJ_n(t), \qquad 0 \le t \le T_n, \quad n = 1, \ldots, M,$$

with

$$(3.2) \qquad\qquad J_n(t) = \sum_{i=1}^{r} \sum_{j=1}^{N_t^{(i)}} H_{ni}\big(X_j^{(i)}\big)$$

is arbitrage-free if

$$(3.3) \qquad \alpha_n(t) = \gamma_n(t)\varphi_0(t) + \gamma_n(t) \sum_{k=\eta(t)}^{n} \frac{\delta\gamma_k(t)^\top L_k(t)}{1 + \delta L_k(t)}$$

$$(3.4) \qquad - \int_0^\infty \sum_{i=1}^{r} H_{ni}(x) \prod_{k=\eta(t)}^{n} \frac{1 + \delta L_k(t-)}{1 + \delta L_k(t-)(1 + H_{ki}(x))} \varphi_i(x, t)\lambda^{(i)}(dx, t)$$

for some R^d-valued process φ_0 and strictly positive scalar processes $\varphi_i(x, \cdot), x \in (0, \infty), i = 1, \ldots, r$, satisfying $\int_0^t \|\varphi_0(s)\|^2\,ds < \infty$, $\int_0^t \int_0^\infty \varphi_i(s, x)\lambda^{(i)}(dx, s)\,ds < \infty, i = 1, \ldots, r$.

REMARKS.

(i) We have implicitly taken $W(t)$ to be a column vector, $\gamma_k(t)$ a row vector, $\gamma_k(t)^\top$ its transpose, and φ_0 a row vector.

(ii) The key feature of this result is that, although it is a continuous-time model, all quantities in (3.1)–(3.4) refer exclusively to properties of the simple forwards $L_k(t)$, in particular to their volatilities $\gamma_k(t)$ and jump-size functions H_{ki}.

(iii) Each component of φ_0 should be interpreted as the market price of risk associated with the corresponding component of the Brownian motion. Similarly, each φ_i may be interpreted as a market price of jump risk associated with the ith marked point process (MPP): $\varphi_i(x, t)$ is the risk premium associated with the arrival of mark x at time t from the ith MPP. An economy of risk-neutral agents corresponds to $\varphi_0 \equiv 0$ and $\varphi_i \equiv 1, i = 1, \ldots, r$. This suggests that if (3.3)–(3.4) holds under the physical measure P, then under a risk-neutral measure $W(t)$ should have drift $-\varphi_0(t)$ and the ith marked point process should have intensity $\varphi_i(x, t)\lambda^{(i)}(dx, t)$. We will see in Appendix A that this is indeed the case.

(iv) The terms in (3.3) are essentially the drift identified by Brace et al. (1997), Jamshidian (1997), and Miltersen et al. (1997), though they worked directly under the risk-neutral measure and so implicitly have $\varphi_0 \equiv 0$. The distinguishing feature of this result is therefore (3.4), the jump-process contribution to the drift. There is in fact a strong

analogy between the diffusion and jump contributions. Focusing on the risk-neutral case, calculations based on (2.2) show that the term

$$\gamma_n(t) \sum_{k=\eta(t)}^{n} \frac{\delta \gamma_k(t)^\top L_k(t)}{1 + \delta L_k(t)}$$

is the product of the forward rate volatility $\gamma_n(t)$ and the volatility of the bond price $B_{n+1}(t)$. Similarly, each term

$$H_{ni}(x) \prod_{k=\eta(t)}^{n} \frac{1 + \delta L_k(t-)}{1 + \delta L_k(t-)(1 + H_{ki}(x))}$$

is the product of a forward rate "jump volatility" H_{ni} and a "jump volatility" of $B_{n+1}(t)$, where by "jump volatility" we mean percentage change at a point of the underlying MPP.

3.2. Forward Measure

For pricing derivatives tied to just one forward rate (such as a caplet), it is often convenient to choose as numeraire a zero coupon bond maturing at the end of the accrual period associated with the forward rate. Thus, to price a claim contingent on L_n we take as numeraire the bond B_{n+1}. Observe from (2.1) that

$$\delta L_n(t) = \frac{B_n(t) - B_{n+1}(t)}{B_{n+1}(t)}$$

is the ratio of a portfolio of assets to $B_{n+1}(t)$, so that under the measure associated with B_{n+1} as numeraire, $L_n(t)$ is a martingale. This in fact is why this particular choice of numeraire is convenient. The measure associated with this numeraire is usually called the *forward measure* or *terminal measure* for maturity T_{n+1}; see Musiela and Rutkowski (1997b) for background.

We now fix the ultimate maturity T_{M+1} and formulate the dynamics of $L_1(t), \ldots, L_M(t)$ under the T_{M+1} forward measure P_{M+1}. Through a generalization of the usual Girsanov theorem (as in Bjork et al. 1997), changing measures corresponds to a change of drift for the underlying Brownian motion and a change of intensity for the MPPs. For emphasis, we let W_{M+1} denote a standard Brownian motion under P_{M+1} and we let $\lambda_{M+1}^{(i)}$ denote the intensity of the ith MPP under P_{M+1}.

THEOREM 3.2. *Under the forward measure P_{M+1}, if for each $n = 1, \ldots, M$ we have*

$$(3.5) \qquad \frac{dL_n(t)}{L_n(t-)} = \alpha_n(t)\, dt + \gamma_n(t)\, dW_{M+1}(t) + dJ_n(t),$$

with $J_n(t)$ as in (3.2) and

$$(3.6) \qquad \alpha_n(t) = - \sum_{k=n+1}^{M} \frac{\delta \gamma_n(t) \gamma_k(t)^\top L_k(t)}{1 + \delta L_k(t)}$$

$$- \int_0^\infty \sum_{i=1}^{r} H_{ni}(x) \prod_{k=n+1}^{M} \frac{1 + \delta L_k(t-)(1 + H_{ki}(x))}{1 + \delta L_k(t-)} \lambda_{M+1}^{(i)}(dx, t),$$

then the model is arbitrage-free. In particular, (3.6) implies

$$(3.7) \qquad \frac{dL_M(t)}{L_M(t-)} = \gamma_M(t)\, dW_{M+1}(t) + dJ_M(t) - \int_0^\infty \sum_{i=1}^r H_{Mi}(x)\lambda_{M+1}^{(i)}(dx, t).$$

We will see that only the comparatively simple case of (3.7) is needed to determine the prices of caps and floors. (The integral in (3.7) compensates the jump process J_M to make L_M a martingale.) The full dynamics in (3.6) would be needed to determine the prices of more complex derivatives.

3.3. A Tractable Subclass of Models

We now illustrate the use of this framework by specializing to a class of tractable models through choice of volatilities, jump-size functions, and intensities. The same strategy could be applied to develop other instances of the framework. We choose the volatilities $\gamma_n(t)$ to be deterministic; this would make $L_M(t)$ lognormal under P_{M+1} in the absence of jumps. If we can arrange to have

$$(3.8) \qquad \lambda_{M+1}^{(i)}(dx, t) = \lambda^{(i)} \cdot g_i(x)\, dx,$$

for some constant $\lambda^{(i)}$ and density g_i, then, under P_{M+1}, each MPP is a compound Poisson process with arrival rate $\lambda^{(i)}$ and i.i.d. jumps distributed according to g_i. Moreover, (3.7) becomes

$$\frac{dL_M(t)}{L_M(t-)} = \gamma_M(t)\, dW_{M+1}(t) + dJ_M(t) - \sum_{i=1}^r \lambda^{(i)} \nu_i\, dt$$

with $\nu_i = \int_0^\infty H_{Mi}(x) g_i(x)\, dx$. This is slight generalization of the Merton jump-diffusion in (2.6) and can be reduced to exactly the form in (2.6) by, for example, taking $H_{M1}(x) = x - 1$, $H_{Mi}(x) \equiv 0$, $i \neq 1$, and taking g_1 to be lognormal. Thus, through (3.8), the naive model of a single forward rate in (2.6) can be reconciled with an arbitrage-free model of the full term structure provided the dynamics in (2.6) are understood to apply under the $M + 1$ forward measure.

From a modeling perspective (3.8) has the unappealing feature that it ties the specification of the jump dynamics to a particular maturity. Moreover, the foregoing discussion does not address the question of whether, in a single model, multiple forward rates can admit the dynamics in (2.6) under their respective forward measures. These considerations lead us to specify the dynamics of all forward rates simultaneously under one measure (we choose the risk-neutral measure Q, as given in Lemma A.1, which intuitively corresponds to the choice of $\varphi_i = 1$, $i = 1, \ldots, r$, and $\varphi_0 = 0$) while keeping in mind the target intensities under each forward measure.

With each L_n we associate a subset I_n of $\{1, \ldots, r\}$ and specify

$$(3.9) \qquad H_{ni}(x) = \begin{cases} x - 1, & i \in I_n \\ 0, & \text{otherwise.} \end{cases}$$

Thus, I_n is the subset of MPPs to which L_n is sensitive. Denote by $\lambda_{\mathsf{Q}}^{(i)}$ the risk-neutral intensity of the ith MPP.

PROPOSITION 3.1.   *If the risk-neutral intensities* $\lambda_{\mathsf{Q}}^{(i)}(dx, t) = \lambda_{\mathsf{Q}}^{(i)}(x, t)\, dx$ *satisfy*

$$(3.10) \qquad \sum_{i \in I_n} \lambda_{\mathsf{Q}}^{(i)}(x, t) = \prod_{k=\eta(t)}^n \frac{1 + \delta x L_k(t-)}{1 + \delta L_k(t-)} \hat{\lambda}_n f_n(x)$$

for some constant $\hat{\lambda}_n$ and some probability density f_n on $(0, \infty)$, then L_n satisfies

(3.11)

$$\frac{dL_n(t)}{L_n(t-)} = -\hat{\lambda}_n m_n \, dt + \gamma_n(t) \, dW_{n+1}(t) + d\left(\sum_{j=1}^{N_t}(Y_j - 1)\right), \qquad n = 1, 2, \dots, M,$$

under its associated forward measure P_{n+1}*, with* W_{n+1} *a standard Brownian motion,* N_t *a Poisson process with rate* $\hat{\lambda}_n$*, the* Y_j *independent and distributed according to* f_n*, and* m_n *the mean of* $(Y_j - 1)$*.*

This result shows that, so long as (3.10) can be enforced, it is possible to construct a model of the term structure of simple forward rates in which multiple forward rates becomes Merton jump-diffusions under their respective forward measures. To show that condition (3.10) is not vacuous, we give an explicit example.

Suppose $r \geq M$ and take $I_n = \{n, n+1, \dots, M\}$ so that (3.10) becomes

$$\sum_{i=n}^{M} \lambda_{\mathsf{Q}}^{(i)}(x, t) = \prod_{k=\eta(t)}^{n} \frac{1 + \delta x L_k(t-)}{1 + \delta L_k(t-)} \hat{\lambda}_n f_n(x).$$

To "solve" for the individual intensities subtract to get

$$\lambda_{\mathsf{Q}}^{(n)}(x, t) = \prod_{k=\eta(t)}^{n} \frac{1 + \delta x L_k(t-)}{1 + \delta L_k(t-)} \left\{ \hat{\lambda}_n f_n(x) - \hat{\lambda}_{n+1} f_{n+1}(x) \frac{1 + \delta x L_{n+1}(t-)}{1 + \delta L_{n+1}(t-)} \right\}.$$

We require nonnegativity of $\lambda_{\mathsf{Q}}^{(n)}$. Since $L_{n+1} \geq 0$,

$$\frac{1 + \delta x L_{n+1}}{1 + \delta L_{n+1}} \leq \max(1, x),$$

so nonnegativity is assured provided that for all $x \geq 0$

(3.12)

$$\frac{f_n(x)}{f_{n+1}(x)} \geq \frac{\hat{\lambda}_{n+1}}{\hat{\lambda}_n} \max(1, x).$$

In the lognormal case, with f_n having the density of $\exp(N(a_n, s_n^2))$ and $y = \log(x)$,

$$\log\left(\frac{f_n(x)}{f_{n+1}(x)}\right) = \log\left(\frac{s_{n+1}}{s_n}\right) - \frac{1}{2}y^2\left(\frac{1}{s_n^2} - \frac{1}{s_{n+1}^2}\right) + y\left(\frac{a_n}{s_n} - \frac{a_{n+1}}{s_{n+1}}\right) - \frac{1}{2}\left(\frac{a_n^2}{s_n^2} - \frac{a_{n+1}^2}{s_{n+1}^2}\right).$$

If we have $s_n > s_{n+1}$, then it can be shown that (3.12) is satisfied if and only if

(3.13)

$$4b_n^2\left\{d_n - \log\left(\frac{\hat{\lambda}_{n+1}}{\hat{\lambda}_n}\right)\right\} \geq \max\left(c_n^2, (c_n - 1)^2\right),$$

where

$$b_n = \frac{1}{2}\left(\frac{1}{s_{n+1}^2} - \frac{1}{s_n^2}\right), \qquad c_n = \frac{a_n}{s_n} - \frac{a_{n+1}}{s_{n+1}}, \qquad d_n = -\frac{1}{2}\left(\frac{a_n^2}{s_n^2} - \frac{a_{n+1}^2}{s_{n+1}^2}\right) + \log\left(\frac{s_{n+1}}{s_n}\right).$$

These conditions generally require that forward rates for more distant maturities jump less frequently and have smaller jumps than forward rates for closer maturities. We stress, however, that this is just one example satisfying (3.10) and other constructions are possible.

This example serves to illustrate another point about our general construction. In order to have all M forward rates evolve according to processes of the form (3.11) under their

respective forward measures, we need M marked point processes. Given $r < M$ marked point processes we can construct a model in which r of the forward rates evolve according to (3.11) under their respective forward measures.

In addition to the normal distribution, it is interesting to consider other distributions for the logarithm of the jump sizes. One attractive candidate is the (asymmetric) double exponential (Laplace) distribution. Compared to a normal distribution with the same mean and standard deviation, the double exponential distribution has a higher peak and heavier tails—it is *leptokurtic*. This property of the jump distribution leads, in turn, to leptokurtosis in the underlying returns.

For the logarithm of the jump sizes to have an asymmetric double exponential distribution, we take the density of the jump sizes to be

$$f_n(x) = \frac{1}{x} p_n \eta_{1,n} e^{-\eta_{1,n} \log(x)} 1_{\{x \geq 1\}} + \frac{1}{x} q_n \eta_{2,n} e^{\eta_{2,n} \log(x)} 1_{\{0 < x < 1\}}, \qquad \eta_{1,n} > 1, \eta_{2,n} > 0,$$

where p_n and $q_n = 1 - p_n$ are probabilities of positive and negative jumps, and the positive and negative jump means are $1/\eta_{1,n}$ and $1/\eta_{2,n}$, respectively. Then

$$\log\left(\frac{f_n(x)}{f_{n+1}(x)}\right) = \log\left(\frac{p_n \eta_{1,n} x^{-\eta_{1,n}}}{p_{n+1} \eta_{1,n+1} x^{-\eta_{1,n+1}}}\right) 1_{\{x \geq 1\}} + \log\left(\frac{q_n \eta_{2,n} x^{-\eta_{2,n}}}{q_{n+1} \eta_{2,n+1} x^{-\eta_{2,n+1}}}\right) 1_{\{0 < x < 1\}}.$$

In this case, (3.12) is satisfied if and only if

$$\frac{p_n \eta_{1,n}}{p_{n+1} \eta_{1,n+1}} \geq \frac{\hat{\lambda}_{n+1}}{\hat{\lambda}_n}, \quad \frac{q_n \eta_{2,n}}{q_{n+1} \eta_{2,n+1}} \geq \frac{\hat{\lambda}_{n+1}}{\hat{\lambda}_n}, \qquad \eta_{1,n+1} - \eta_{1,n} \geq 1, \eta_{2,n+1} - \eta_{2,n} \geq 0.$$

Closed-form prices for caplets can then be obtained from explicit formulas in Kou (2002) for options on futures contracts under the double exponential jump model. In our setting, these formulas provide a tractable alternative to the Merton-like formulas of the normal jump model.

4. PRICING INTEREST RATE DERIVATIVES

In this section, we derive tractable expressions for the prices of some interest rate derivatives in the models of the previous section. The formulas apply in the setting of Proposition 3.1 with deterministic γ_n—that is, when the MPP intensities are chosen so that the relevant forward rates are Merton jump-diffusions under their respective forward measures.

4.1. Pricing Caps and Floors

An interest rate caplet for the period $[T_n, T_{n+1}]$ and struck at K is a derivative security paying $\delta(L_n(T_n) - K)^+$ at time T_{n+1}. We can explicitly find the time-t price $C_n(t)$ of the nth caplet if (3.10) holds with f_n a lognormal density (or, e.g., log-Laplace). This condition, and the requirement that all $\gamma_k, k = 1, \ldots, M$, be deterministic will be in force throughout this section. With Y_n having distribution f_n, let $1 + m_n$ be the mean of Y_n and let s_n^2 be the variance of $\log(Y_n)$.

Pricing under the risk-neutral measure Q leads to

$$C_n(t) = \delta E_Q\left[e^{-\int_t^{T_{n+1}} r(s)\, ds}(L_n(T_n) - K)^+ \,\big|\, \mathcal{F}_t\right],$$

$t \leq T_{n+1}$, with $r(s)$ the short rate as before. It is more convenient in the present context to price under the forward measure for T_{n+1}. Noting that $C_n(t)/B_{n+1}(t)$ is a martingale under this measure (since it is the ratio of an asset price to the numeraire price) leads to

$$C_n(t) = \delta B_{n+1}(t) \mathsf{E}_{T_{n+1}} \left[\frac{C_n(T_{n+1})}{B_n(T_{n+1})} \,\Big|\, \mathcal{F}_t \right] = \delta B_{n+1}(t) \mathsf{E}_{T_{n+1}} \left[(L_n(T_n) - K)^+ \,|\, \mathcal{F}_t \right].$$

Absent jumps, $L_n(T_n)$ would be lognormally distributed under P_{n+1}, and this price would be given by the Black (1976) formula. In our context, the price is given by a "forward" version of Merton (1976) jump-diffusion option pricing formula. This variant stands in the same relation to Merton's formula as the Black formula does to the Black-Scholes formula. It can be derived by conditioning on the number of jumps of the Poisson process in $[0, T_n]$.

Let

(4.1)

$$BC(F, T, K, \sigma^2, b) = b \left[F\Phi\left(\frac{\log(F/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \right) - K\Phi\left(\frac{\log(F/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \right) \right]$$

denote the Black formula with initial price F, maturity T, strike K, volatility parameter σ, and discount factor b. We now have the following corollary.

COROLLARY 4.1.   *If (3.10) holds and all γ_k are deterministic then the time-t price of the nth caplet, $t < T_n$, is given by*

$$(4.2) \quad C_n(t) = \delta \sum_{j=0}^{\infty} e^{-\hat{\lambda}_n(T_n - t)} \frac{(\hat{\lambda}_n(T_n - t))^j}{j!} BC\big(L_n^{(j)}(t), T_n - t, K, v_j(t)^2, B_{n+1}(t)\big),$$

where $L_n^{(j)}(t) = L_n(t) \cdot e^{-\hat{\lambda}_n m_n(T_n - t)} \cdot (1 + m_n)^j$ *and* $v_j(t)^2 = (\rho^2(t) + j s_n^2)/(T_n - t)$, $\rho^2(t) = \int_t^{T_n} \|\gamma_n(u)\|^2 \, du$. *For $T_n \leq t \leq T_{n+1}$, $C_n(t) = \delta B_{n+1}(t)(L_n(T_n) - K)^+$.*

By summing the prices of individual caplets one can price a cap, which is simply a portfolio of caplets with consecutive maturities. A floor can be priced by summing the prices of single-period floors. A single-period floor paying $(K - L_n(T_n))^+$ at T_{n+1} has a time-t price, $t < T_n$, of

$$(4.3) \qquad \delta \sum_{j=0}^{\infty} e^{-\hat{\lambda}_n(T_n - t)} \frac{(\hat{\lambda}_n(T_n - t))^j}{j!} BP\big(L_n^{(j)}(t), T_n - t, K, v_j(t)^2, B_{n+1}(t)\big),$$

where

$$BP(F, T, K, \sigma^2, b) = b \left[K\Phi\left(-\frac{\log(F/K) - \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \right) - F\Phi\left(-\frac{\log(F/K) + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \right) \right];$$

and of course for $T_n \leq t \leq T_{n+1}$, the price is $\delta B_{n+1}(t)(K - L_n(T_n))^+$.

Although they involve infinite series, the expressions in (4.2) and (4.3) can be evaluated numerically very quickly to a high degree of accuracy through truncation. The ability to value these basic instruments quickly is important in fitting model parameters to market data. Indeed, caps and floors are among the most liquid of all interest rate derivatives and are therefore a natural source from which to extract the information in market prices.
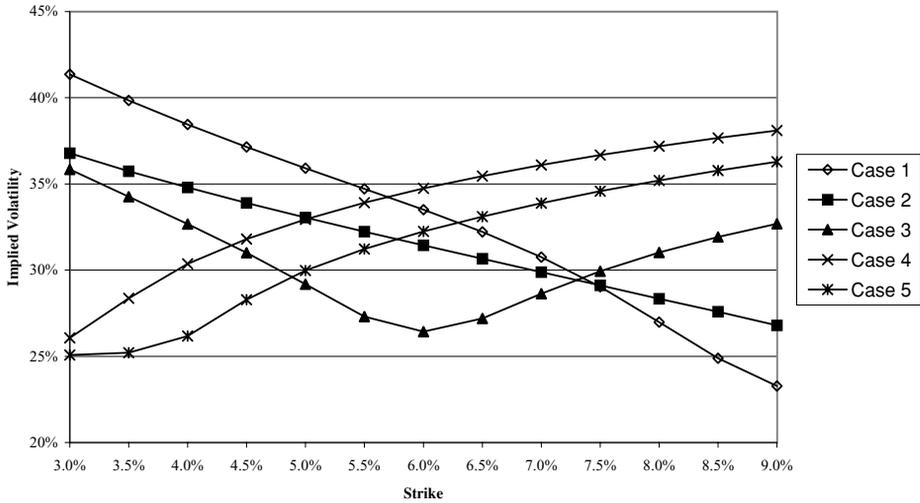
FIGURE 4.1. Implied volatility curves with $T = 2$, $\delta = 1/2$, all forward rates at 6%, $\gamma = .05$, and the following parameters. Case 1: $\hat{\lambda} = 0.75$, $m = -0.25$, $s = 0.30$. Case 2: $\hat{\lambda} = 1.5$, $m = -0.20$, $s = 0.15$. Case 3: $\hat{\lambda} = 0.5$, $m = 0$, $s = 0.45$. Case 4: $\hat{\lambda} = 1.5$, $m = 0.20$, $s = 0.20$. Case 5: $\hat{\lambda} = 1$, $m = 0.20$, $s = 0.25$.

4.2. Numerical Illustrations

To illustrate the flexibility of this model in producing a variety of patterns of implied volatilities, we consider two sets of numerical examples: one showing flexibility through hypothetical parameters and one based on a cross section of implied volatilities in the Japanese yen market on a fixed date.

Our first set of examples illustrates the variety of "skews" and "smiles" that can be captured through the model. For this, we choose a set of model parameters, calculate caplet prices at a range of strikes using Corollary 4.1, and then find the corresponding implied volatilities based on the Black (1976) formula. More precisely, these implied volatilities are the values of σ_T that equate the Black formula price to the price computed using (4.2) with all other parameters held fixed.

We use a maturity of $T = 2$ years, an accrual period of $\delta = \frac{1}{2}$ year, a flat term structure with all forward rates at 6%, and a constant γ of 5%. We denote by m and s the parameters of the lognormal jumps and by $\hat{\lambda}$ the jump rate. Figure 4.1 shows that by varying these parameters we obtain upward sloping, downward sloping, and U-shaped implied volatilities. Although it is difficult to disentangle the effects of the parameters, the shape of the implied volatility curve appears to be determined primarily by m, with $m > 0$ tending to produce an upward slope, $m < 0$ tending to produce a downward slope, and $m = 0$ producing a U-shaped curve. See Das and Sundaram (1999) for an investigation of implied volatility shapes in jump-diffusion and stochastic volatility models.

Our second illustration attempts to reproduce a specific set of implied volatilities. Andersen and Andreasen (2000) graph implied volatilities for 2-year and 9-year caplets in the Japanese LIBOR market as of late May 1998. Their implied volatilities based on midmarket prices are reproduced in Figure 4.2. Corresponding 6-month forward rates for the same period are 1.181% for the 2-year maturity and 2.913% for the 9-year maturity.[2]

---

[2] We thank Leif Andersen of General Re Financial Products for providing these values and also the midmarket implied volatilities in the figure.
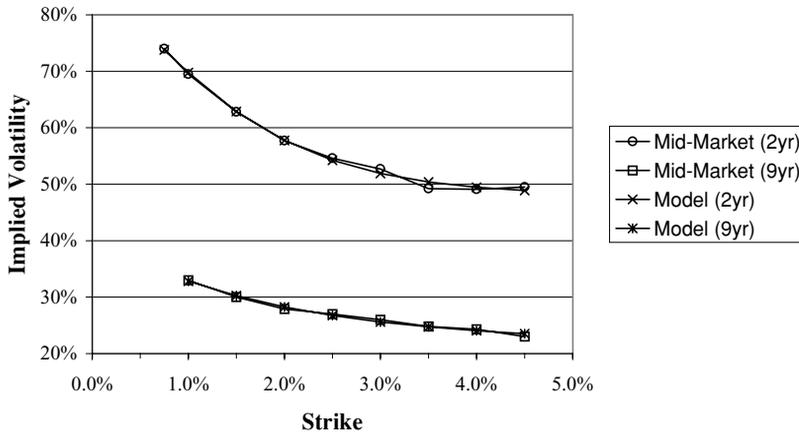
FIGURE 4.2. Midmarket and model-implied volatilities for Japanese LIBOR caplets in May 1998. The parameters used for the model-implied volatilities are $\hat{\lambda}_4 = 0.63$, $\gamma_4 = 0.28$, $s_4 = 0.65$, $m_4 = -0.5$ and $\hat{\lambda}_{18} = 0.16$, $\gamma_{18} = 0.10$, $s_{18} = 0.60$, $m_{18} = -0.3$.

Figure 4.2 also shows implied volatility curves derived from (4.2) using the same forward rates and the parameters in the caption. (With 6-month accrual intervals, the 2-year and 9-year caplets correspond to $n = 4$ and $n = 18$ in the notation of (4.2).) The figure suggests the possibility of a very close fit even to a very sharp market skew. Because the parameters used in (4.2) apply under a martingale measure rather than the objective measure, we can interpret the parameters in Figure 4.2 as suggesting that the market attaches a large risk premium to the possibility of a sharp downward movement in Japanese interest rates.

4.3. Pricing Bond Options

As noted in, for example, Miltersen et al. (1997), from the prices of caps and floors it is possible to derive prices of puts and calls on zero coupon bonds, provided the maturity of the bond is one period later in the tenor structure than the expiration of the option. For example, to price a put on B_{n+1} expiring at T_n and struck at $K < 1$ is to evaluate

$$\text{Put}_{n+1}(t) = B_{n+1}(t)\mathsf{E}_{T_{n+1}}\left[\frac{(K - B_{n+1}(T_n))^+}{B_{n+1}(T_n)}\,\Bigg|\,\mathcal{F}_t\right],$$

again using the fact that B_{n+1} is the numeraire associated with P_{n+1}. From (2.1) we find that $B_{n+1}(T_n) = 1/(1 + \delta L_n(T_n))$. Thus, simple algebra shows that

$$\frac{(K - B_{n+1}(T_n))^+}{B_{n+1}(T_n)} = \delta K\left(L_n(T_n) - \frac{1 - K}{\delta K}\right)^+,$$

so a put struck at K can be valued as a portfolio of K caplets struck at $(1 - K)/(\delta K)$. Through some algebraic simplification, this results in the following pricing formula (taking $t = 0$ for simplicity):

$$\text{Put}_{n+1} = K\delta B_{n+1}(0)\sum_{j=0}^{\infty}\frac{e^{-\hat{\lambda}_n T_n}(\hat{\lambda}_n T_n)^j}{j!}\left\{L_n^{(j)}\Phi\big(-d_-^{(j)}\big) - \frac{1 - K}{K\delta}\Phi\big(-d_+^{(j)}\big)\right\},$$

with

$$d_{\pm}^{(j)} = \frac{\hat{\lambda}_n m_n T_n - j \log(1 + m_n)}{v_j \sqrt{T_n}} + \frac{1}{v_j \sqrt{T_n}} \left\{ \log\left(\frac{(1-K)B_{n+1}(0)}{K\{B_n(0) - B_{n+1}(0)\}} \right) \pm \frac{1}{2} v_j^2 T_n \right\}.$$

In summary, we have the following result.

COROLLARY 4.2.   *The price of a put expiring at T_n on a zero coupon bond maturing at T_{n+1}, with strike price K, is given by*

$$\sum_{j=0}^{\infty} \pi_n^{(j)} \left\{ K(B_n(0) - B_{n+1}(0)) \cdot e^{-\hat{\lambda}_n m_n T_n} (1 + m_n)^j \Phi\left(-d_-^{(j)}\right) - (1-K)B_{n+1}(0)\Phi\left(-d_+^{(j)}\right) \right\}$$

and the price of the corresponding call is given by

$$\sum_{j=0}^{\infty} \pi_n^{(j)} \left\{ (1-K)B_{n+1}(0)\Phi\left(d_+^{(j)}\right) - K(B_n(0) - B_{n+1}(0)) \cdot e^{-\hat{\lambda}_n m_n T_n} (1 + m_n)^j \Phi\left(d_-^{(j)}\right) \right\},$$

where $\pi_n^{(j)} = e^{-\hat{\lambda}_n T_n} (\hat{\lambda}_n T_n)^j / (j!)$.

Pricing more complex derivatives is likely to entail simulation; see Glasserman and Merener (2003) for accurate simulation methods for the models developed here. Nevertheless, the availability of tractable expressions for simple instruments is valuable in choosing model parameters (to be used in a simulation) consistent with market prices of the simple instruments.

A computational procedure for pricing bond options in a Heath-Jarrow-Morton model with jumps is developed in Das (1999a). Bond options are often tractable in affine jump-diffusion models; see Das and Foresi (1996), and Duffie et al. (2000).

5. SWAP RATES AND SWAPTIONS

A framework similar to that in Sections 3 and 4 can be developed based on swap rates rather than forward rates. This formulation leads to tractable expression for swaptions—that is, options on swaps.

Consider an interest rate swap over $[T_n, T_{M+1}]$ in which two parties exchange fixed rate and floating rate payments at dates $T_{n+1}, \ldots, T_{M+1}$. The floating rate payment at a date T_j is $\delta L_{j-1}(T_{j-1})$ times a notional principal; at a fixed rate of K, the fixed rate payment at each date is the product of the notional principal and δK. Under a *forward* swap, the parties commit at some time $t < T_n$ to enter into a swap over $[T_n, T_{M+1}]$.

At a date $t \le T_n$, the *forward* swap rate $S_n(t)$ is the fixed rate that gives the forward swap over $[T_n, T_{M+1}]$ a value of 0. It is standard that forward swap rates can be expressed in terms of zero coupon bond prices as

(5.1) $$S_n(t) = \frac{B_n(t) - B_{M+1}(t)}{\delta \sum_{j=n+1}^{M+1} B_j(t)};$$

see, for example, Musiela and Rutkowski (1997b, p. 389). From a set of forward swap rates $(S_1, \ldots, S_M)$ (each S_i applying to $[T_i, T_{M+1}]$) one can determine the forward rates $(L_1, \ldots, L_M)$ and vice versa, so the two sets of variables offer alternative ways of describing the term structure.

Jamshidian (1997) characterizes the arbitrage-free dynamics of forward swap rates in a pure diffusion setting; we give a corresponding result for a model with jumps as well as diffusion. We state the result under the forward measure P_{M+1}. Through a change of measure a similar result could be formulated under the risk-neutral measure or under the physical measure. In this notation, Jamshidian's characterization becomes

$$\frac{dS_n(t)}{S_n(t)} = \alpha_n^o(t) + \theta_n(t)\,dW_{M+1}(t)$$

with W_{M+1} a standard Brownian motion under P_{M+1} and

$$(5.2) \qquad \alpha_n^o = -\sum_{l=n+1}^{M+1} \frac{\delta s_{n,l}\theta_n\theta_l^\top S_l}{(1+\delta S_l)s_{n,n+1}}, \qquad s_{n,l} = \delta \sum_{k=l}^{M+1}\prod_{j=n+1}^{k}(1+\delta S_j).$$

As in previous sections we extend the framework to include r MPPs; let $\lambda_{M+1}^{(i)}$, $i = 1, \ldots, r$, denote the intensity of the ith MPP under P_{M+1}.

THEOREM 5.1.   *For each $n = 1, \ldots, M$ let $\theta_n(\cdot)$ be a bounded, adapted, R^d-valued process and let $G_{ni}, i = 1, \ldots, r$, be deterministic functions from $[0, \infty)$ to $[-1, \infty)$. The model*

$$(5.3)$$

$$\frac{dS_n(t)}{S_n(t-)} = \alpha_n(t)\,dt + \theta_n(t)\,dW_{M+1}(t) + dJ_n(t), \qquad 0 \le t \le T_n, \quad n = 1, \ldots, M,$$

with $J_n(t) = \sum_{i=1}^{r}\sum_{j=1}^{N_t^{(i)}} G_{ni}(X_j^{(i)})$ is arbitrage free if

$$(5.4)$$

$$\alpha_n(t) = \alpha_n^o(t) - \int_0^\infty \sum_{i=1}^{r} G_{ni}(x)\frac{\sum_{k=n+1}^{M+1}\prod_{j=n+1}^{k}(1+\delta S_j(t-)[1+G_{ji}(x)])}{\sum_{k=n+1}^{M+1}\prod_{j=n+1}^{k}(1+\delta S_j(t-))}\lambda_{M+1}^{(i)}(dx, t).$$

The proof of this result is similar to that of Theorem 3.2. We omit the details but in Appendix C we give some insight into the form of the result.

We now formulate a subclass of models leading to tractable expressions for swaptions. Consider an option expiring at $T \le T_n$ to enter into a swap over $[T_n, T_{M+1}]$ paying a fixed rate K and receiving floating rate payments (a payer swaption). At expiry, the value of the option is

$$(5.5) \qquad \delta \sum_{j=n+1}^{M+1} B_j(T)(S_n(T) - K)^+,$$

so a swaption may be viewed as an option on a swap rate; see, for example, Musiela and Rutkowski (1997b, p. 397). Tractable expressions for swaptions follow from appropriate choices for the parameters in the dynamics of S_n. Let

$$(5.6) \qquad \lambda_{n,M}^{(i)}(dx, t) = \frac{\sum_{k=n+1}^{M+1}\prod_{j=n+1}^{k}(1+\delta S_j(t-)(1+G_{ji}(x)))}{\sum_{k=n+1}^{M+1}\prod_{j=n+1}^{k}(1+\delta S_j(t-))}\lambda_{M+1}^{(i)}(dx, t).$$

We now have the following proposition.

PROPOSITION 5.1.   *Suppose the θ_n are deterministic. If there are constants $\hat{\lambda}_n$ and lognormal densities f_n for which*

(5.7)

$$\sum_{i=1}^{r} \lambda_{n,M}^{(i)}(dx, t) = \hat{\lambda}_n f_n(x)\, dx, \quad and \quad \sum_{i=1}^{r} G_{ni}(x)\lambda_{n,M}^{(i)}(dx, t) = (x - 1)\hat{\lambda}_n f_n(x)\, dx,$$

then the time-t value of a payer swaption expiring at $T > t$ for a swap over $[T_n, T_{M+1}]$ is

(5.8)

$$C_n(t) = \sum_{j=0}^{\infty} e^{-\hat{\lambda}_n(T-t)} \frac{(\hat{\lambda}_n(T - t))^j}{j!} BC\left(S_n^{(j)}(t), T - t, K, v_j(t)^2, \delta \sum_{k=n+1}^{M+1} B_k(t) \right),$$

where $S_n^{(j)}(t) = S_n(t) \cdot e^{-\hat{\lambda}_n m_n(T-t)} \cdot (1 + m_n)^j$, $v_j(t)^2 = (\rho^2(t) + js_n^2)/(T - t)$, $\rho^2(t) = \int_t^T \|\theta_n(u)\|^2\, du$, and f_n is the lognormal density of $\exp(N(\log(1 + m_n) - \frac{1}{2}s_n^2, s_n^2))$.

This result provides a "Merton-Black formula" for the price of a payer swaption. The case of a *receiver* swaption, with $(K - S_n)^+$ in (5.5) in place of $(S_n - K)^+$, works analogously; see (4.3). The market convention is to quote implied volatilities for swaptions based on a version of Black's formula (i.e., assuming that the underlying forward swap rate is a lognormal). Through (5.8), the range of patterns illustrated in Figure 4.1 is thus available in fitting parameters to implied volatilities in the market prices of swaptions. It should be noted, however, that Proposition 5.1 and Corollary 4.1 cannot hold simultaneously; for example, θ_n and γ_n cannot both be deterministic in a consistent model of forward rates and swap rates, as discussed in Jamshidian (1997).

To verify that the hypothesis (5.7) is not vacuous, we show that it can be satisfied by modifying the example that follows Proposition 3.1. We choose G_{ni} according to the right side of (3.9) with $I_n = \{n, n + 1, \ldots, M\}$; (5.7) is satisfied if

$$\sum_{i=n}^{M} \lambda_{M+1}^{(i)}(x, t) R_n(x, t) = \hat{\lambda}_n f_n(x) \quad \text{with} \quad R_n(x, t) = \frac{\sum_{k=n+1}^{M+1} \prod_{j=n+1}^{k}(1 + \delta S_j(t-)x)}{\sum_{k=n+1}^{M+1} \prod_{j=n+1}^{k}(1 + \delta S_j(t-))}.$$

By taking differences (as in the steps leading to (3.12)) we find that this is feasible if $R_{n+1}(x, t)\hat{\lambda}_n f_n(x) \geq R_n(x, t)\hat{\lambda}_{n+1} f_{n+1}(x)$ for all $n < M$. An induction argument establishes that

$$\frac{R_n(x, t)}{R_{n+1}(x, t)} \leq \begin{cases} (1 + M - n)/(M - n), & x < 1, \\ x, & x \geq 1. \end{cases}$$

A sufficient condition for feasibility is thus

$$\frac{\hat{\lambda}_n f_n(x)}{\hat{\lambda}_{n+1} f_{n+1}(x)} \geq \max\left(x, \frac{1 + M - n}{M - n} \right);$$

this restricts the parameters of the densities f_n as in the steps following (3.12).

APPENDIX A: DERIVATION OF ARBITRAGE RESTRICTIONS

Here we prove Theorem 3.1 by showing that the class of models in the theorem can be embedded in an arbitrage-free model of instantaneous forward rates. To do this, we first

need some background on models of instantaneous forward with jumps, as developed in Björk et al. (1997); see also Jarrow and Madan (1995, 1999).

Let $f(t, T)$ denote the instantaneous forward rate for maturity T as of time $t \leq T$; that is,

$$(A.1) \qquad B(t, T) = \exp\left(-\int_t^T f(t, u)\, du\right).$$

Consider a model of the dynamics of the forward curve of the form

(A.2)

$$df(t, T) = \alpha_f(t, T)\, dt + \sigma(t, T)\, dW(t) + d\left(\sum_{i=1}^r \sum_{n=1}^{N_t^{(i)}} h_i(t, X_n^{(i)}, T)\right), \quad 0 \leq t \leq T \leq T^*,$$

with, as before, T^* arbitrary, $W(t)$ a d-dimensional Brownian motion, and $\{(\tau_n^{(i)}, X_n^{(i)})\}$, $i = 1, \ldots, r$, marked point processes.

The following result is an adaptation of part of Theorem 3.13 in Björk et al. (1997) to our setting. Björk et al. treat in detail the questions of admissible trading strategies, absence of arbitrage, and existence of risk-neutral measures in models of instantaneous forward rates. It is in the sense of their Definition 3.6 that we use the term "arbitrage free" in the following proposition.

PROPOSITION A.2.   *The model in (A.2) is arbitrage free if*

$$(A.3) \quad \alpha_f(t, T) = \sigma(t, T)[\sigma^*(t, T)^\top + \varphi_0(t)]$$

$$- \int_0^\infty \sum_{i=1}^r h_i(t, x, T) \exp\left(-\int_t^T h_i(t, x, s)\, ds\right) \varphi_i(x, t)\lambda^{(i)}(dx, t),$$

with

$$(A.4) \qquad \sigma^*(t, T) = \int_t^T \sigma(t, u)\, du$$

for some predictable d-dimensional process φ_0 and strictly positive measurable φ_i satisfying

(A.5)

$$\int_0^t \|\varphi_0(s)\|^2\, ds < \infty, \qquad \int_0^t \int_0^\infty \varphi_i(s, x)\lambda^{(i)}(dx, s)\, ds < \infty, \qquad i = 1, \ldots, r, \text{ P-}a.s.$$

Proof.   Construct a single MPP with mark space $E = (0, \infty) \times \{1, \ldots, r\}$ by assigning to the new MPP a point with mark $(X_n^{(i)}, i)$ at time $\tau_n^{(i)}$, $i = 1, \ldots, r, n = 1, 2, \ldots$ Replace each $h_i(x, t, T)$ with $h((x, i), t, T)$ to combine the r functions h_i into a single function on the enlarged mark space. Replace the symbol $\int_0^\infty \sum_{i=1}^r$ with $\int_E$. These substitutions reduce the case of r MPPs to the case of a single MPP with an augmented mark space and allow us to apply Theorem 3.13 of Björk et al. (1997). The conditions on α_f in the proposition are easily seen to imply the condition in equation (21) of Björk et al. from which the result follows.   □

Theorem 3.13 of Björk et al. (1997) would allow a slightly more general formulation of the drift condition in this proposition and a statement of necessity of this condition as well as sufficiency. The formulation above makes the extension of the condition of Heath et al. (1992) more transparent and is adequate for our purposes.

Inspection of (A.3) suggests that under a risk-neutral measure, W should have a drift of $-\varphi_0$ and the ith MPP should have an intensity of $\varphi_i(x, t)\lambda^{(i)}(dx, t)$. We now make this precise.

LEMMA A.1. *Suppose (A.5) holds and define a process R_t by*

$$\log R_t = -\int_0^t \varphi_0(s)\,dW(s) - \frac{1}{2}\int_0^t \|\varphi_0(s)\|^2\,ds$$

$$+ \sum_{i=1}^r \sum_{n=1}^{N_t^{(i)}} \log \varphi_i\big(X_n^{(i)}, \tau_n^{(i)}\big) + \int_0^t \int_0^\infty \sum_{i=1}^r (1 - \varphi_i(x, s))\lambda^{(i)}(dx, ds),$$

$R_0 = 1$, and suppose $\mathsf{E}_\mathsf{P}[R_t] = 1$ *for all t. Then there exists a probability measure Q with $d\mathsf{Q} = R_t\,d\mathsf{P}$ on $\mathcal{F}_t$ such that, under Q, $W(t) + \varphi_0(t)$ is a standard Brownian motion and $\{(\tau_n^{(i)}, X_n^{(i)})\}$ has intensity $\lambda_\mathsf{Q}^{(i)}(dx, t) = \varphi_i(x, t)\lambda^{(i)}(dx, t)$.*

Proof. As in the proof of Proposition A.2, construct a combined MPP from the individual MPPs by enlarging the mark space to $(0, \infty) \times \{1, \ldots, r\}$. From Theorem 3.12 of Björk et al. (1997) we find that the combined MPP has a Q-intensity given by $\lambda_\mathsf{Q}(A \times I, t) = \int_A \sum_{i \in I} \varphi_i(x, t)\lambda^{(i)}(dx, t)$, for any measurable $A \subseteq (0, \infty)$ and $I \subseteq \{1, \ldots, r\}$. The ith MPP is the subsequence of the combined MPP with marks in $(0, \infty) \times \{i\}$ and thus has Q-intensity

$$\lambda_\mathsf{Q}^{(i)}(A, t) = \lambda_\mathsf{Q}(A \times \{i\}, t) = \int_A \varphi_i(x, t)\lambda^{(i)}(dx, t);$$

that is, $\lambda^{(i)}(dx, t) = \varphi_i(x, t)\lambda^{(i)}(dx, t)$. The change of drift in W follows from Theorem 3.12 of Björk et al. (1997). $\square$

Armed with these results, we can determine the dynamics of simple forward rates implied by the arbitrage-free dynamics of instantaneous forward rates. To streamline some proofs we introduce a piece of notation. A marked point process $\{(\tau_n, X_n)\}$ can be described by a random measure μ on the product of the time axis and the mark space: the measure μ simply assigns unit mass to each point (τ_n, X_n). This makes it possible to write, for example,

$$\sum_{n=1}^{N_t} h(X_n, \tau_n) = \int_0^t \int_0^\infty h(x, t)\mu(dx, dt),$$

with N_t the counting process for the MPP. In the following we use $\mu^{(i)}$ to denote the random measure assigning unit mass to each $(\tau_n^{(i)}, X_n^{(i)})$, $n = 1, 2, \ldots$.

THEOREM A.2. *If $f(t, T)$ satisfies (A.2) with α_f as in (A.3), then $L(t, T)$ satisfies*

(A.6)

$$\frac{\delta}{1 + \delta L(t-, T)}\,dL(t, T) = [\sigma^*(t, T+\delta) - \sigma^*(t, T)] \cdot \{dW(t) + \varphi_0(t)\,dt + \sigma^*(t, T+\delta)^\top dt\}$$

$$+ \sum_{i=1}^r \int_0^\infty e^{-\int_t^T h_i(t,x,s)\,ds}\left(e^{-\int_T^{T+\delta} h_i(t,x,s)\,ds} - 1\right)\varphi_i(x, t)\lambda^{(i)}(dx, t)\,dt$$

$$+ \sum_{i=1}^r \int_0^\infty \left(e^{\int_T^{T+\delta} h_i(t,x,u)\,du} - 1\right)\mu^{(i)}(dx, dt),$$

with $t \le T \le T^*$ and σ^* as in (A.4). The risk-neutral dynamics are given by

(A.7)

$$\frac{\delta}{1 + \delta L(t-, T)} dL(t, T) = [\sigma^*(t, T + \delta) - \sigma^*(t, T)] \cdot \left\{ dW_Q(t) + \sigma^*(t, T + \delta)^\top dt \right\}$$

$$+ \sum_{i=1}^{r} \int_0^\infty e^{-\int_t^T h_i(t,x,s)\, ds} \left(e^{-\int_T^{T+\delta} h_i(t,x,s)\, ds} - 1 \right) \lambda_Q^{(i)}(dx, t)\, dt$$

$$+ \sum_{i=1}^{r} \int_0^\infty \left(e^{\int_T^{T+\delta} h_i(t,x,u)\, du} - 1 \right) \mu^{(i)}(dx, dt),$$

where W_Q is standard Brownian motion under Q and $\lambda_Q^{(i)}$ is the intensity of the ith MPP under Q.

Proof. We prove the risk-neutral case (A.7); this suffices because Lemma A.1 indicates how to move between the two measures. Since we consider only the risk-neutral measure, we write W (rather than W_Q) for a Brownian motion under Q. For notational simplicity, we will let

$$\tilde{Z}(t, T) = \int_T^{T+\delta} f(t, s)\, ds, \qquad \tilde{h}_i(t, x, y) = \exp\left(-\int_t^y h_i(t, x, u)\, du \right).$$

Recall that

(A.8) $$L(t, T) = \frac{1}{\delta} \left(\exp\left\{ \int_T^{T+\delta} f(t, u)\, du \right\} - 1 \right) = \frac{1}{\delta} (\exp\{\tilde{Z}(t, T)\} - 1).$$

Proposition A.1 gives the following dynamics for the forward rate:

(A.9) $$df(t, T) = \frac{\partial}{\partial T} \left(\frac{1}{2} \|\sigma^*(t, T)\|^2 \right) dt + \sum_{i=1}^{r} \int_0^\infty \frac{\partial \tilde{h}_i(t, x, T)}{\partial T} \lambda_Q^{(i)}(dx, t)\, dt$$

$$+ \sigma(t, T)\, dW(t) + \sum_{i=1}^{r} \int_0^\infty h_i(t, x, T) \mu^{(i)}(dx, dt).$$

By interchanging differentiation and integration and interchanging the order of integration (see Assumption 2.1 of Björk et al. 1997 and the surrounding discussion), we obtain from (A.9) that

$$d\tilde{Z}(t, T) = \int_T^{T+\delta} df(t, u)\, du$$

$$= \left\{ \int_T^{T+\delta} \sigma(t, u)\, du \right\} dW(t) + \sum_{i=1}^{r} \int_0^\infty \left\{ \int_T^{T+\delta} h_i(t, x, u)\, du \right\} \mu^{(i)}(dx, dt)$$

$$+ \left\{ \int_T^{T+\delta} \frac{\partial}{\partial u} \left(\frac{1}{2} \|\sigma^*(t, u)\|^2 \right) du \right\} dt$$

$$+ \sum_{i=1}^{r} \int_0^\infty \left\{ \int_T^{T+\delta} \frac{\partial \tilde{h}_i(t, x, u)}{\partial u} du \right\} \lambda_Q^{(i)}(dx, t)\, dt.$$

Since

$$\int_T^{T+\delta} \frac{\partial}{\partial u}\left(\frac{1}{2}\|\sigma^*(t,u)\|^2\right) du = \frac{1}{2}\|\sigma^*(t,T+\delta)\|^2 - \frac{1}{2}\|\sigma^*(t,T)\|^2,$$

$$\int_T^{T+\delta} \frac{\partial \tilde{h}_i(t,x,u)}{\partial u} du = \tilde{h}_i(t,x,T+\delta) - \tilde{h}_i(t,x,T),$$

$$\int_T^{T+\delta} \sigma(t,u) du = \sigma^*(t,T+\delta) - \sigma^*(t,T),$$

we further have

(A.10)

$$d\tilde{Z}(t,T) = \left\{\frac{1}{2}\|\sigma^*(t,T+\delta)\|^2 - \frac{1}{2}\|\sigma^*(t,T)\|^2\right\} dt + (\sigma^*(t,T+\delta) - \sigma^*(t,T))\, dW(t)$$

$$+ \sum_{i=1}^r \int_0^\infty \{\tilde{h}_i(t,x,T+\delta) - \tilde{h}_i(t,x,T)\}\lambda_Q^{(i)}(dx,t)\, dt$$

$$+ \sum_{i=1}^r \int_0^\infty \left\{\int_T^{T+\delta} h_i(t,x,u)\, du\right\} \mu^{(i)}(dx,dt).$$

Therefore, applying Itô's formula for jump processes (p. 140 of Elliott 1982) to $L(t,T)$ in (A.8) yields

$$dL(t,T) = \frac{1}{\delta}\exp\{\tilde{Z}(t,T)\}\{d\tilde{Z}(t,T)\}^c + \frac{1}{2\delta}\exp\{\tilde{Z}(t,T)\}\{d\tilde{Z}(t,T)\}^c\{d\tilde{Z}(t,T)\}^c$$

$$+ \sum_{i=1}^r \frac{1}{\delta}\int_0^\infty \left(\exp\left\{\tilde{Z}(t-,T) + \int_T^{T+\delta} h_i(t,x,u)\, du\right\}\right.$$

$$\left. - \exp\{\tilde{Z}(t-,T)\}\right)\mu^{(i)}(dx,dt),$$

where $\{\cdot\}^c$ denotes the continuous part of the process in braces. Combining this with equations (A.10) and (A.8) leads to

$$dL(t,T) = \frac{1}{\delta}(1+\delta L(t,T)) \cdot \left(\left\{\frac{1}{2}\|\sigma^*(t,T+\delta)\|^2 - \frac{1}{2}\|\sigma^*(t,T)\|^2\right\} dt\right.$$

$$+ [\sigma^*(t,T+\delta) - \sigma^*(t,T)]\, dW(t)$$

$$+ \frac{1}{2}\|\sigma^*(t,T+\delta) - \sigma^*(t,T)\|^2\, dt$$

$$\left. + \sum_{i=1}^r \int_0^\infty \{\tilde{h}_i(t,x,T+\delta) - \tilde{h}_i(t,x,T)\}\lambda_Q^{(i)}(dx,t)\, dt\right)$$

$$+ \frac{1}{\delta}(1+\delta L(t-,T))\sum_{i=1}^r \int_0^\infty \left[\exp\left(\int_T^{T+\delta} h_i(t,x,u)\, du\right) - 1\right]\mu^{(i)}(dx,dt).$$

Noting that, for any row vectors a, b, $\|a\|^2 - \|b\|^2 + \|a-b\|^2 = 2(a-b)a^\top$, we get (A.7). □

Proof of Theorem 3.1.   To establish this result as a consequence of Theorem A.2, we need to show that σ and the h_i in (A.2) can be chosen so that the dynamics in Theorem A.2 (with $T \in \{T_1, \ldots, T_M\}$) simplify to those in Theorem 3.1. This will show that it is indeed possible to embed the simple forward rates in Theorem 3.1 in an arbitrage-free model of instantaneous forwards.

For simplicity, we focus on the risk-neutral setting. Given γ_n and H_{ni}, $i = 1, \ldots, r$, we reduce (A.7) to the dynamics in Theorem 3.1 (with $\varphi_0 \equiv 0$ and $\varphi_i \equiv 1$) by finding σ and h_i, $i = 1, \ldots, r$, satisfying

$$(A.11) \qquad \sigma^*(t, T_{n+1}) - \sigma^*(t, T_n) = \frac{\delta \gamma_n(t) L_n(t)}{1 + \delta L_n(t)}$$

and, for each $i = 1, \ldots, r$,

$$(A.12) \qquad \int_{T_n}^{T_{n+1}} h_i(t, x, s)\, ds = \log\left(\frac{1 + \delta L_n(t-)(1 + H_{ni}(x))}{1 + \delta L_n(t-)}\right),$$

for all $n = 1, \ldots, M$. We can always enforce these conditions; we could, for example, choose each $\sigma(t, \cdot)$ and $h_i(t, x, \cdot)$ constant over intervals (T_k, T_{k+1}), but it is also possible to choose these to be smooth functions of maturity.

Using (A.11), we can rewrite (A.7) evaluated at $T = T_n$ as

$(A.13)$

$$dL_n(t) = \gamma_n(t) L_n(t)\, dW(t) + \left(\gamma_n(t) L_n(t) \sigma^*(t, T_n)^\top + \frac{\delta \|\gamma_n(t)\|^2 L_n^2(t)}{1 + \delta L_n(t)}\right) dt$$

$$+ \frac{1}{\delta}(1 + \delta L_n(t)) \sum_{i=1}^{r} \int_0^\infty e^{-\int_t^{T_n} h_i(t,x,s)\, ds} \left(e^{-\int_{T_n}^{T_{n+1}} h_i(t,x,s)\, ds} - 1\right) \lambda_Q^{(i)}(dx, t)\, dt$$

$$+ \frac{1}{\delta}(1 + \delta L_n(t-)) \sum_{i=1}^{r} \int_0^\infty \left(e^{\int_{T_n}^{T_{n+1}} h_i(t,x,u)\, du} - 1\right) \mu^{(i)}(dx, dt).$$

If we choose $\sigma(t, u) = 0$ for $t \le u < t + \delta$, repeated use of (A.11) shows that

$$\gamma_n(t) L_n(t) \sigma^*(t, T_n)^\top + \frac{\delta \|\gamma_n(t)\|^2 L_n^2(t)}{1 + \delta L_n(t)} = \sum_{k=\eta(t)}^{n} \frac{\delta \gamma_n(t) \gamma_k(t)^\top L_k(t) L_n(t)}{1 + \delta L_k(t)}.$$

Thus, the first dt term in (A.13) matches (3.3). The dW terms in (A.13) and (3.1) also match. For the jump terms, notice that (A.12) is equivalent to

$$e^{\int_{T_n}^{T_{n+1}} h_i(t,x,s)\, ds} = \frac{\delta H_{ni}(x) L_n(t-)}{1 + \delta L_n(t-)} + 1 = \frac{1 + \delta L_n(t-)(1 + H_{ni}(x))}{1 + \delta L_n(t-)},$$

that is, to

$$(A.14) \qquad \frac{1}{\delta}(1 + \delta L_n(t-)) \left(e^{\int_{T_n}^{T_{n+1}} h_i(t,x,s)\, ds} - 1\right) = H_{ni}(x) L_n(t-).$$

If we choose $h_i(t, x, s) = 0$, $t \le s < t + \delta$, it follows that

$$\frac{1}{\delta}(1 + \delta L_n(t)) \sum_{i=1}^{r} \int_0^\infty e^{-\int_t^{T_n} h_i(t,x,s)\, ds} \left(e^{-\int_{T_n}^{T_{n+1}} h_i(t,x,s)\, ds} - 1\right) \lambda_Q^{(i)}(dx, t)\, dt$$

$$= \frac{1}{\delta}(1 + \delta L_n(t)) \sum_{i=1}^{r} \int_0^\infty e^{-\int_t^{T_{n+1}} h_i(t,x,s)\, ds} \left(1 - e^{\int_{T_n}^{T_{n+1}} h_i(t,x,s)\, ds}\right) \lambda_Q^{(i)}(dx, t)\, dt$$

$$= -\sum_{i=1}^{r} \int_{0}^{\infty} L_n(t) H_{ni}(x) \prod_{k=\eta(t)}^{n} \frac{1 + \delta L_k(t-)}{1 + \delta L_k(t-)(1 + H_{ki}(x))} \lambda_{\mathsf{Q}}^{(i)}(dx, t)\, dt,$$

so (A.12) ensures that the intensity term in (A.13) matches (3.4). Finally, (A.14) shows that

$$\frac{1}{\delta}(1 + \delta L_n(t-)) \sum_{i=1}^{r} \int_{0}^{\infty} \left(e^{\int_{T_n}^{T_{n+1}} h_i(t,x,u)\, du} - 1 \right) \mu^{(i)}(dx, dt) = L_n(t-) J_n(t)$$

with $J_n(t)$ as in (3.1) and (3.2). □

Observe that Theorem A.2 applies simultaneously for all maturities $0 \le T \le T^*$. In Theorem 3.1 we restrict T to a finite set $\{T_1, \ldots, T_M\}$ solely to ensure that we can choose σ and h_i to satisfy (A.11) and (A.12). Extending Theorem 3.1 to hold simultaneously for all maturities would require finding suitable conditions ensuring solutions to (A.11) and (A.12). A thorough investigation of this issue in the absence of jumps is undertaken in Brace et al. (1997).

The assumption that the volatilities $\sigma(t, \cdot)$ and jump-size functions $h_i(t, x, \cdot)$ are 0 for maturities shorter than δ is but one way to embed a model of simple forwards in a model of instantaneous forwards. As emphasized by Jamshidian (1997) in the pure-diffusion case, the prices of derivatives sensitive only to a discrete tenor of forward rates are independent of the volatility assumed for maturities shorter than the basic accrual periods. This becomes more transparent when one uses a numeraire asset tailored to the tenor structure, rather than the money market account which is best suited to instantaneous rates. We return to this point after establishing the necessary theoretical background on changes of numeraire in Appendix B.

We conclude this appendix with comments on technical conditions that have not been made explicit in the statements of our results. We established Theorem A.2 as a consequence of Theorem 3.13 of Björk et al. (1997), which relies on their Assumptions 2.1 and 3.1. Thus, Theorem A.2 also relies on those assumptions. We established Theorem 3.1 as a consequence of Theorem A.2; we must therefore require that Assumption 2.1 of Björk et al. hold for the σ and h_i constructed from the γ_n and H_{ni} given in Theorem 3.1.

APPENDIX B: CHANGE OF NUMERAIRE

Our main objective here is to prove Theorem 3.2. Before proving the result, we need to develop some preliminary material. Recall that $B(t, T)$ denotes the time-t price of a zero coupon bond maturing at T.

LEMMA B.1. *With f as in (A.2) and α_f as in (A.3), $B(t, T)$ satisfies*

(B.1)

$$\frac{dB(t, T)}{B(t-, T)} = r(t)\, dt - \sigma^*(t, T)\, dW_{\mathsf{Q}}(t)$$

$$+ \int_{0}^{\infty} \sum_{i=1}^{r} \left(\exp \left\{ -\int_{t}^{T} h_i(t, x, s)\, ds \right\} - 1 \right) \left(\mu^{(i)}(dx, dt) - \lambda_{\mathsf{Q}}^{(i)}(dx, t)\, dt \right)$$

with W_{Q} a standard Brownian motion under Q and $\lambda_{\mathsf{Q}}^{(i)}$ the intensity of $\mu^{(i)}$ under Q.

Proof. The proof follows from (A.1) and Proposition A.2, and essentially restates Proposition 2.2(3) of Björk et al. (1997). □

To change numeraire from $\beta(t)$ to a zero coupon bond, for any maturity T define

$$Z(t, T) = \frac{B(t, T)}{\beta(t)} = \frac{B(t, T)}{\exp\left(\int_0^t r(s)\,ds\right)}.$$

Itô's formula for jump processes applied to $Z(t, T)$ shows, via (B.1), that

$$dZ(t, T) = \frac{1}{\beta(t)}(-r(t)\,dt\,B(t, T) + dB(t, T))$$

$$= -Z(t, T)\sigma^*(t, T)dW_Q(t)$$

$$+ Z(t-, T) \sum_{i=1}^r \int_0^\infty \left(\exp\left\{-\int_t^T h_i(t, x, s)\,ds\right\} - 1\right)$$

$$\times \left(\mu^{(i)}(dx, dt) - \lambda_Q^{(i)}(dx, t)\,dt\right).$$

Thus,

$$\frac{dZ(t, T+\delta)}{Z(t-, T+\delta)} = -\left(\int_t^{T+\delta} \sigma(t, s)\,ds\right) dW_Q(t)$$

$$+ \sum_{i=1}^r \int_0^\infty \left(e^{-\int_t^{T+\delta} h_i(t,x,s)\,ds} - 1\right)\left(\mu^{(i)}(dx, dt) - \lambda_Q^{(i)}(dx, t)\,dt\right)$$

is a martingale. We can now define the measure $P_{T+\delta}$ through the likelihood ratio

$$\left(\frac{dP_{T+\delta}}{dQ}\right)_t = \frac{Z(t, T+\delta)}{B(0, T+\delta)}.$$

The next result follows from the version of Girsanov's theorem (Thm. 3.13) in Björk et al. (1997).

LEMMA B.2. *Under* $P_{T+\delta}$, *the intensity of* $\mu^{(i)}(dx, dt)$ *is given by*

$$\lambda_{T+\delta}^{(i)}(dx, t) = e^{-\int_t^{T+\delta} h_i(t,x,s)\,ds} \lambda_Q^{(i)}(dx, t),$$

and the process $W_{T+\delta}(t), dW_{T+\delta}(t) = dW_Q(t) + \sigma^*(t, T+\delta)\,dt$, *is a standard Brownian motion.*

Proof of Theorem 3.2. As in the proof of Theorem 3.1 we embed the model of simple forward rates in an arbitrage-free model of instantaneous forward rates; indeed, we enforce the conditions (A.11) and (A.12) imposed for the previous embedding.

For any $n = 1, \ldots, M$, we obtain the risk-neutral dynamics of $L_n(t)$ from (A.7):

$$dL_n(t) = \frac{1}{\delta}(1 + \delta L_n(t))\left[\sigma^*(t, T_{n+1}) - \sigma^*(t, T_n)\right]L_n(t) \cdot \left\{dW_Q(t) + \sigma^*(t, T_{n+1})^\top dt\right\}$$

$$+ \frac{1}{\delta}(1 + \delta L_n(t)) \sum_{i=1}^r \int_0^\infty e^{-\int_t^{T_n} h_i(t,x,s)\,ds}\left(e^{-\int_{T_n}^{T_{n+1}} h_i(t,x,s)\,ds} - 1\right)\lambda_Q^{(i)}(dx, t)\,dt$$

$$+ \frac{1}{\delta}(1 + \delta L_n(t-)) \sum_{i=1}^r \int_0^\infty \left(e^{\int_{T_n}^{T_{n+1}} h_i(t,x,u)\,du} - 1\right)\mu^{(i)}(dx, dt).$$

Using (A.11) and Lemma B.2, we can rewrite this as

(B.1)

$$
dL_n(t) = \gamma_n(t)L_n(t) \cdot \left\{ dW_{M+1}(t) + \left[\sigma^*(t, T_{n+1}) - \sigma^*(t, T_{M+1})\right]^\top dt \right\}
$$
$$
+ \frac{1}{\delta}(1 + \delta L_n(t)) \sum_{i=1}^{r} \int_0^\infty e^{\int_{T_n}^{T_{M+1}} h_i(t,x,s)\,ds} \left(e^{-\int_{T_n}^{T_{n+1}} h_i(t,x,s)\,ds} - 1 \right) \lambda_{M+1}^{(i)}(dx, t)\,dt
$$
$$
+ \frac{1}{\delta}(1 + \delta L_n(t-)) \sum_{i=1}^{r} \int_0^\infty \left(e^{\int_{T_n}^{T_{n+1}} h_i(t,x,u)\,du} - 1 \right) \mu^{(i)}(dx, dt).
$$

Repeated application of (A.11) shows that

$$
\sigma^*(t, T_{n+1}) - \sigma^*(t, T_{M+1}) = - \sum_{j=n+1}^{M} \frac{\gamma_j(t)\delta L_j(t)}{1 + \delta L_j(t)},
$$

so the first dt term in (B.1) matches the term in (3.6). That the jump term in (B.1) matches the term $dJ(t, T - k\delta)$ in (3.5) follows directly from (A.14). Next we consider the intensity terms. For each $i = 1, \dots, r$,

$$
\frac{1}{\delta}(1 + \delta L_n(t)) e^{\int_{T_n}^{T_{M+1}} h_i(t,x,s)\,ds} \left(e^{-\int_{T_n}^{T_{n+1}} h_i(t,x,s)\,ds} - 1 \right)
$$
$$
= \frac{1}{\delta}(1 + \delta L_n(t)) e^{\int_{T_{n+1}}^{T_{M+1}} h_i(t,x,s)\,ds} \left(1 - e^{\int_{T_n}^{T_{n+1}} h_i(t,x,s)\,ds} \right)
$$
$$
= -L_n(t) H_{n,i}(x) e^{\int_{T_{n+1}}^{T_{M+1}} h_i(t,x,s)\,ds}
$$
$$
= -L_n(t) H_{n,i}(x) \prod_{j=n+1}^{M} \frac{1 + \delta L_j(t-)(1 + H_{ji}(x))}{1 + \delta L_j(t-)},
$$

using (A.14) for the second equality and (A.12) for the third. Summing over i produces the intensity term in (3.6).   □

Close inspection of this proof reveals that it does not rely on the conditions $\sigma(t, u) = 0$ and $h_i(t, x, u) = 0$, $t < u < t + \delta$, used in the proof of Theorem 3.1; indeed, the values of σ and h_i in this range of arguments is immaterial to Theorem 3.2. This point merits comment. We have worked primarily in the risk-neutral measure in order to be able to apply the general framework of Björk et al. (1997). This is convenient but entails an arbitrary choice of parameters for maturities shorter than δ. This shortcoming ensues from the reliance of the risk-neutral measure on instantaneous rates through its numeraire.

An alternative approach parallels the type of construction in Jamshidian (1997) and Musiela and Rutkowski (1997a), starting from a numeraire involving only simple rates. For example, starting under the measure P_{M+1} and imposing the requirement that each $B_n(t)/B_{M+1}(t)$ be a martingale produces the model in Theorem 3.2 without reference to an underlying term structure of instantaneous rates. Similarly, starting from a discretely compounded money market account (accruing simple interest at rate $L_n(T_n)$ over the interval $[T_n, T_{n+1}]$) produces the model in Theorem 3.1, again without reference to instantaneous rates. A thorough development of this approach would involve extending the general results in Jamshidian (1997) and Musiela and Rutkowski (1997a) to processes with discontinuities. This has recently been undertaken in a very general semimartingale setting by Jamshidian (1999).

Proof of Proposition 3.1. The superposition of the MPPs in I_n has risk-neutral intensity $\sum_{i \in I_n} \lambda_Q^{(i)}$. By Lemma B.2 and equations (A.12) and (3.19), under the T_{n+1}-forward measure the new intensity is

$$\prod_{k=\eta(t)}^{n} \frac{1 + \delta L_k(t-)}{1 + x\delta L_k(t-)} \sum_{i \in I_n} \lambda_Q^{(i)}(x, t) \, dx,$$

which reduces to $\hat{\lambda}_n f_n(x) \, dx$, in light of (3.20). It follows from this intensity that under the T_{n+1}-forward measure the superposition of the MPPs in I_n constitute a marked Poisson process with constant arrival rate $\hat{\lambda}_n$ and i.i.d. marks with density f_n. Our choice of functions h_{ni} now produces (3.11). □

APPENDIX C: SWAP RATE MODEL

The proof of Theorem 5.1 is similar to that of Theorem 3.2 so we omit the details and instead provide some intuition for the form of the result. From (5.1) we see that S_n is the ratio of one linear combination of asset prices to another. As in Jamshidian (1997) analysis of the pure-diffusion case, this suggests that it is convenient to take as numeraire asset the denominator $B_{n,M}(t) \equiv \sum_{i=n+1}^{M+1} \delta B_i(t)$. Let $\mathsf{P}_{n,M}$ denote the probability measure associated with this numeraire, defined by

(C.1)
$$\left. \frac{d\mathsf{P}_{n,M}}{d\mathsf{P}_{M+1}} \right|_t = \frac{B_{n,M}(t)}{B_{M+1}(t)} \cdot \frac{B_M(0)}{B_{n,M}(0)}.$$

Under this measure, S_n must be a martingale because it is the ratio of $B_n - B_{M+1}$ to the numeraire asset price. This requires that S_n have the form

(C.2)

$$\frac{dS_n(t)}{S_n(t-)} = -\sum_{i=1}^{r} \int_0^{\infty} G_{ni}(x) \lambda_{n,M}^{(i)}(dx, t) + \theta_n(t) \, dW_{n,M}(t) + d\left(\sum_{i=1}^{r} \sum_{j=1}^{N_t^{(i)}} G_{ni}(X_j^{(i)})\right),$$

with $W_{n,M}$ a standard Brownian motion under $\mathsf{P}_{n,M}$ and $\lambda_{n,M}^{(i)}$ the intensity of the ith MPP under $\mathsf{P}_{n,M}$.

This specifies the dynamics of each S_n under a measure $\mathsf{P}_{n,M}$ that depends on n. However, we would like to specify the dynamics of all S_n simultaneously under a single measure P_{M+1}. For this we need to determine how the dynamics of S_n change when we change measures from $\mathsf{P}_{n,M}$ to P_{M+1}—that is, when we change numeraires from $B_{n,M}$ to B_{M+1}. As in Jamshidian (1997), we have

(C.3)
$$\frac{B_{n,M}(t)}{B_{M+1}(t)} = \sum_{k=n+1}^{M+1} \delta \prod_{j=n+1}^{k} (1 + \delta S_j(t));$$

from (C.1) we see that, up to a proportionality constant, this is $d\mathsf{P}_{n,M}/d\mathsf{P}_{M+1}$. We must therefore apply the Girsanov theorem (see Lemma A.1) with this ratio for R_t to determine the change of dynamics under this change of measure.

Applying Itô's formula to the logarithm of (C.3) shows that the volatility of this ratio is

$$\frac{\sum_{k=n+1}^{M+1} \delta \sum_{l=n+1}^{k} \delta \theta_l(t) S_l(t) \prod_{j=n+1, j \neq l}^{k} (1 + \delta S_j(t))}{\sum_{k=n+1}^{M+1} \delta \prod_{j=n+1}^{k} (1 + \delta S_j(t))}.$$

Premultiplying by $\theta_n(t)$ yields $-\alpha_n^o(t)$ after interchanging the order of summation, so

$$\theta_n(t)\,dW_{n,M}(t) = \alpha_n^o(t)\,dt + \theta_n(t)\,dW_{M+1}(t),$$

with $W_{n,M}$ and W_{M+1} standard Brownian motions under $\mathsf{P}_{n,M}$ and P_{M+1}, respectively. This explains the first contribution to the drift in (5.4). For the second term we need to consider the jumps in (C.3). Given the arrival of a mark x in the ith MPP at time t, the percentage jump in (C.3) is

$$\frac{\sum_{k=n+1}^{M+1} \delta \prod_{j=n+1}^{k}(1+\delta S_j(t-)\{1+G_{ni}(x)\})}{\sum_{k=n+1}^{M+1} \delta \prod_{j=n+1}^{k}(1+\delta S_j(t-))} - 1,$$

because $S_j(t) = S_j(t-)\{1+G_{ni}(x)\}$. From the Girsanov theorem it follows that the intensities $\lambda_{n,M}^{(i)}$ and $\lambda_{M+1}^{(i)}$ under $\mathsf{P}_{n,M}$ and P_{M+1} are related as in (5.6). The integral in (5.4) is thus just another way of writing the one in (C.2). We have therefore outlined the key steps in going from (C.2) under $\mathsf{P}_{n,M}$ to (5.3) and (5.4) in Theorem 5.1.

Proof of Proposition 5.1. With $\mathsf{E}_{T_{M+1}}$ denoting expectation under the P_{M+1} forward measure (i.e., with numeraire asset B_{M+1}), the swaption price is

$$(\text{C.4}) \qquad C_n(t) = B_{M+1}(t)\mathsf{E}_{T_{M+1}}\left[\delta\sum_{j=n}^{M} B_j(T)\frac{(S_n(T)-K)^+}{B_{M+1}(T)}\,\middle|\,\mathcal{F}_t\right].$$

As in Jamshidian (1997), changing numeraire from P_{M+1} to $\mathsf{P}_{n,M}$, (C.4) becomes

$$C_n(t) = \delta \sum_{j=n+1}^{M+1} B_j(t)\mathsf{E}_{n,M}\big[(S_n(T)-K)^+\,\big|\,\mathcal{F}_t\big].$$

It follows from (5.1) that S_n is a martingale under $\mathsf{P}_{n,M}$, since it is the ratio of a linear combination of bond prices to the numeraire price. Moreover, as discussed above, the ith MPP has intensity $\lambda_{n,M}^{(i)}$ under $\mathsf{P}_{n,M}$. With (5.7), we conclude that S_n has the dynamics of a Merton jump-diffusion under $\mathsf{P}_{n,M}$, with volatility $\|\theta_n(t)\|$, jump rate $\hat{\lambda}_n$, and jump size density f_n. The expectation above for C_n can now be evaluated and yields (5.8). $\qquad\square$

REFERENCES

ANDERSEN, L., and J. ANDREASEN (2000): Volatility Skews and Extensions of the LIBOR Market Model, *Appl. Math. Finance* 7, 1–32.

BABBS, S., and N. WEBBER (1997): Term Structure Modelling under Alternative Official Regimes; in *Mathematics of Derivative Securities*, eds., M. A. H. Dempster and S. R. Pliska. Cambridge University Press, Cambridge, UK.

BAKSHI, G., C. CAO, and Z. CHEN (1997): Empirical Performance of Alternative Option Pricing Models, *J. Finance* 52, 2003–2049.

BATES, D. (2000): Post '87 Crash Fears in the S&P 500 Futures Option Market, *J. Econometrics* 94, 181–238.

BJÖRK, T., Y. KABANOV, and W. RUNGGALDIER (1997): Bond Market Structure in the Presence of Marked Point Processes, *Math. Finance* 7, 211–239.

BLACK, F. (1976): The Pricing of Commodity Contracts, *J. Financial Econ.* 3, 167–179.

BRACE, A., D. GATAREK, and M. MUSIELA (1997): The Market Model of Interest Rate Dynamics, *Math. Finance* 7, 127–155.

BURNETAS, A., and P. RITCHKEN (1997): On Rational Jump Diffusion Models: An Approach Using Potentials, *Rev. Derivat. Res*. 1, 325–349.

DAS, S. R. (1999a): A Direct Discrete-Time Approach to Poisson-Gaussian Bond Option Pricing in Heath-Jarrow-Morton Model, *J. Econ. Dynam. Control* 23, 333–369.

DAS, S. R. (1999b): The Surprise Element: Jumps in Interest Rate Diffusions, Working paper, Harvard Business School.

DAS, S. R., and S. FORESI (1996): Exact Solutions for Bond and Option Prices with Systematic Jump Risk, *Rev. Derivat. Res*. 1, 7–24.

DAS, S. R., and R. K. SUNDARAM (1999): Of Smiles and Smirks: A Term Structure Perspective, *J. Financial Quant. Anal*. 34, 211–240.

DUFFIE, D., and R. KAN (1996): A Yield-Factor Model of Interest Rates, *Math. Finance* 6, 379–406.

DUFFIE, D., J. PAN, and K. SINGLETON (2000): Transform Analysis and Option Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

DUFFIE, D., and K. SINGLETON (1999): Modeling Term Structures of Defaultable Bonds, *Rev. Financial Stud*. 12, 687–720.

EL-JAHEL, L., H. LINDBERG, and W. PERRAUDIN (1997): Interest Rate Distributions, Yield Curve Modelling and Monetary Policy; in *Mathematics of Derivative Securities*, eds. M. A. H. Dempster and S. R. Pliska. Cambridge University Press, Cambridge, UK.

ELLIOTT, R. J. (1982): *Stochastic Calculus and Applications*. New York: Springer-Verlag.

GLASSERMAN, P., and N. MERENER (2001): Numerical Solution of Jump Diffusion LIBOR Market Models, *Finance Stoch*. 7, 1–27.

HEATH, D., R. JARROW, and A. MORTON (1992): Bond Pricing and the Term Structure of Interest Rates: a New Methodology for Contingent Claims Valuation, *Econometrica* 60, 77–105.

JAMSHIDIAN, F. (1997): LIBOR and Swap Market Models and Measures, *Finance Stoch*. 1, 293–330.

JAMSHIDIAN, F. (1999): LIBOR Market Model with Semimartingales, Working paper, NetAnalytic Ltd., London.

JARROW, R., and D. MADAN (1995): Option Pricing Using the Term Structure of Interest Rates to Hedge Systematic Discontinuities in Asset Returns, *Math. Finance* 5, 311–336.

JARROW, R., and D. MADAN (1999): Hedging Contingent Claims on Semimartingales, *Finance Stoch*. 3, 111–134.

JARROW, R., and S. TURNBULL (1995): Pricing Options on Financial Securities Subject to Default Risk, *J. Finance* 50, 53–86.

JOHANNES, M. S. (2003): The Statistical and Economic Role of Jumps in Interest Rates, *J. Finance*. to appear.

KOU, S. G. (2002): A Jump Diffusion Model for Option Pricing, *Mgemt. Sci*. 48, 1086–1101.

MERTON, R. (1976): Option Pricing when Underlying Stock Returns Are Discontinuous, *J. Financial Econ*. 3, 125–144.

MILTERSEN, K. R., K. SANDMANN, and D. SONDERMANN (1997): Closed-Form Solutions for Term Structure Derivatives with Lognormal Interest Rates, *J. Finance* 52, 409–430.

MUSIELA, M., and M. RUTKOWSKI (1997a): Continuous-Time Term Structure Models: Forward Measure Approach, *Finance Stoch*. 1, 261–292.

MUSIELA, M., and M. RUTKOWSKI (1997b): *Martingale Methods in Financial Modeling*. New York: Springer.

SHIRAKAWA, H. (1991): Interest Rate Option Pricing with Poisson-Gaussian Forward Rate Curve Processes, *Math. Finance* 1, 77–94.