Transform Analysis and Asset Pricing
for Affine Jump-Diffusions

Darrell Duffie, Jun Pan, and Kenneth Singleton*

First draft : November 26, 1997
This draft : August 1, 1999

Graduate School of Business, Stanford University

Abstract:

In the setting of "affine" jump-diffusion state processes, this paper provides an analytical treatment of a class of transforms, including various Laplace and Fourier transforms as special cases, that allow an analytical treatment of a range of valuation and econometric problems. Example applications include fixed-income pricing models, with a role for intensity-based models of default, as well as a wide range of option-pricing applications. An illustrative example examines the implications of stochastic volatility and jumps for option valuation. This example hightlights the impact on option 'smirks' of the joint distribution of jumps in volatility and jumps in the underlying asset price, through both jump amplitude as well as jump timing.

1

1 Introduction

In valuing financial securities in an arbitrage-free environment, one inevitably faces a trade-off between the analytical and computational tractability of pricing and estimation, and the complexity of the probability model for the state vector X. In light of this trade-off, academics and practitioners alike have found it convenient to impose sufficient structure on the conditional distribution of X to give closed- or nearly closed-form expressions for securities prices. An assumption that has proved to be particularly fruitful in developing tractable, dynamic asset pricing models is that X follows an *affine jump-diffusion* (*AJD*), which is, roughly speaking, a jump-diffusion process for which the drift vector, "instantaneous" covariance matrix, and jump intensities all have affine dependence on the state vector. Prominent among *AJD* models in the term-structure literature are the Gaussian and square-root diffusion models of Vasicek [1977] and Cox, Ingersoll, and Ross [1985]. In the case of option pricing, there is a substantial literature building on the particular affine stochastic-volatility model for currency and equity prices proposed by Heston [1993].

This paper synthesizes and significantly extends the literature on affine asset-pricing models by deriving a closed-form expression for an "extended transform" of an *AJD* process X, and then showing that this transform leads to analytically tractable pricing relations for a wide variety of valuation problems. More precisely, fixing the current date t and a future payoff date T, suppose that the stochastic "discount rate" $R(X_t)$, for computing present values of future cash flows, is an affine function of X_t. Also, consider the generalized terminal payoff function $(v_0 + v_1 \cdot X_T) e^{u \cdot X_T}$ of X_T, where v_0 is scalar and the n elements of each of v_1 and u are scalars. These scalars may be real, or more generally, complex. We derive a closed-form expression for the transform

$$E_t \left(\exp \left(- \int_t^T R(X_s, s) \, ds \right) (v_0 + v_1 \cdot X_T) e^{u \cdot X_T} \right), \qquad (1.1)$$

where E_t denotes expectation conditioned on the history of X up to t. Then, using this transform, we show that the tractability offered by extant, specialized affine pricing models extends to the entire family of *AJD*s. Additionally, by selectively choosing the payoff $(v_0 + v_1 \cdot X_T) e^{u \cdot X_T}$, we significantly extend the set of pricing problems (security payoffs) that can be tractably addressed with X following an *AJD*. To motivate the usefulness of our extended trans-

2

form in theoretical and empirical analyses of affine models, we briefly outline three applications.

1.1 Affine, Defaultable Term Structure Models

There is a large literature on the term structure of default-free bond yields that presumes that the state vector underlying interest rate movements follows an AJD under risk-neutral probabilities (see, for example, Dai and Singleton [1999] and the references therein). Assuming that the instantaneous riskless short-term rate r_t is affine with respect to an n-dimensional AJD process X_t (that is $r_t = \rho_0 + \rho_1 \cdot X_t$) Duffie and Kan [1996] show that the $(T - t)$-period zero-coupon bond price,

$$E \left(\exp \left(- \int_t^T r_s \, ds \right) \mid X_t \right), \tag{1.2}$$

is known in closed form, where expectations are computed under the risk-neutral measure.[1]

Recently, considerable attention has been focused on extending these models to allow for the possibility of default in order to price corporate bonds and other credit-sensitive instruments.[2] To illustrate the new pricing issues that may arise with the possibility of default, suppose that, with respect to given risk-neutral probabilities, X is an AJD, the arrival of default is at a stochastic intensity λ_t, and upon default the holder recovers a constant fraction w of face value. Then, from results in Lando [1998], the initial price of a T-period zero-coupon bond is given under technical conditions by

$$E \left(\exp \left(- \int_0^T (r_t + \lambda_t) \, dt \right) \right) + w \int_0^T q_t \, dt, \tag{1.3}$$

where $q_t = E \left[\lambda_t \exp \left(- \int_0^t (r_u + \lambda_u) \, du \right) \right]$. The first term in (1.3) is the value of a claim that pays 1 contingent on survival to maturity T. We may view q_t as the price density of a claim that pays 1 if default occurs in the "interval" $(t, t + dt)$. Thus the second term in (1.3) is the price of any proceeds from default before T. These expectations are to be taken with respect to the

---

[1]The entire class of affine term structure models is obtained as the special case of (1.1) found by setting $R(X_t) = r_t$, $u = 0$, $v_0 = 1$, and $v_1 = 0$.

[2]See, for example, Jarrow, Lando, and Turnbull [1997] and Duffie and Singleton [1999].

given risk-neutral probabilites. Both the first term of (1.3) and, for each t, the price density q_t can be computed in closed form using our extended transform. Specifically, assuming that both r_t and λ_t are affine with respect to X_t, the first term in (1.3) is the special case of (1.1) obtained by letting $R(X_t) = r_t + \lambda_t$, $u = 0$, $v_0 = 1$ and $v_1 = 0$. Similarly, q_t is obtained as a special case of (1.1) by setting $u = 0$, $R(X_t) = r_t + \lambda_t$, and $v_0 + v_1 \cdot X_t = \lambda_t$. Thus, using our extended transform, the pricing of defaultable zero-coupon bonds with constant fractional recovery of par reduces to the computation of a one-dimensional integral of a known function. Similar reasoning can be used to derive closed-form expressions for bond prices in environments for which the default arrival intensity is affine in X along with "gapping" risk associated with unpredictable transitions to different credit categories, as shown by Lando [1998].

A different application of the extended transform is pursued by Piazzesi [1998], who extends the AJD model in order to treat term-structure models with releases of macro-economic information and with central-bank interest-rate targeting. She considers jumps at both random and at deterministic times, and allows for an intensity process and interest-rate process that have linear-quadratic dependence on the underlying state vector, extending the basic results of this paper.

1.2   Estimation of Affine Asset Pricing Models

Another useful implication of (1.1) is that, by setting $R = 0$, $v_0 = 1$, and $v_1 = 0$, we obtain a closed-form expression for the conditional characteristic function ϕ of X_T given X_t, defined by $\phi(u, X_t, t, T) = E\left(e^{iu \cdot X_T} | X_t\right)$, for real u. Because knowledge of ϕ is equivalent to knowledge of the joint conditional density function of X_T, this result is useful in estimation and all other applications involving the transition densities of an AJD.

For instance, Singleton [1999] exploits knowledge of ϕ to derive maximum likelihood estimators for AJDs based on the conditional density $f(\cdot \mid X_t)$ of X_{t+1} given X_t, obtained by Fourier inversion of ϕ as

$$f(X_{t+1} \mid X_t) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{-iu \cdot X_{t+1}} \phi(u, X_t, t, t+1) \, du. \qquad (1.4)$$

Das [1998] exploits (1.4) for the specific case of a Poisson-Gaussian AJD to compute method-of-moments estimators of a model of interest rates.

Method-of-moments estimators can also be constructed directly in terms of the conditional characteristic function. From the definition of ϕ,

$$E\left[e^{iu \cdot X_{t+1}} - \phi(u, X_t, t, t+1) \,|\, X_t\right] = 0, \qquad (1.5)$$

so any measurable function of X_t is orthogonal to the "error" $(e^{iu \cdot X_{t+1}} - \phi(u, X_t, t, t+1))$. Singleton [1999] uses this fact, together with the known functional form of ϕ, to construct generalized method-of-moments estimators of the parameters governing *AJD*s and, more generally, the parameters of asset pricing models in which the state follows an *AJD*. These estimators are computationally tractable and, in some cases, achieve the same asymptotic efficiency as the maximum likelihood estimator. Jiang and Knight [1999] and Chacko and Viceira [1999] propose related, characteristic-function based estimators of the stochastic volatility model of asset returns with volatility following a square-root diffusion.[3]

1.3   Affine Option-Pricing Models

In an influential paper in the option-pricing literature, Heston [1993] showed that the risk-neutral exercise probabilities appearing in the call option-pricing formulas for bonds, currencies, and equities can be computed by Fourier inversion of the conditional characteristic function, which he showed is known in closed form for his particular affine, stochastic volatility model. Building on this insight,[4] a variety of option-pricing models have been developed for state vectors having at most a single jump type (in the asset return), and whose behavior between jumps is that of a Gaussian or "square-root" diffusion.[5]

Knowing the extended transform (1.1) in closed-form, we can extend this option pricing literature to the case of general multi-dimensional *AJD* pro-

---

[3]Liu and Pan [1997] and Liu [1997] propose alternative estimation strategies that exploit the special structure of affine diffusion models.

[4]Among the many recent papers examining option prices for the case of state variables following square-root diffusions are Bakshi, Cao, and Chen [1997], Bakshi and Madan [1999], Bates [1996], Bates [1997], Chen and Scott [1993], Chernov and Ghysels [1998], Pan [1998], Scott [1996], and Scott [1997], among others.

[5]More precisely, the short-term interest rate has been assumed to be an affine function of independent square-root diffusions and, in the case of equity and currency option pricing, spot-market returns have been assumed to follow stochastic-volatility models in which volatility processes are independent "square-root" diffusions that may be correlated with the spot-market return shock.

cesses with much richer dynamic inter-relations among the state variables and much richer jump distributions. For example, we provide an analytically tractable method for pricing derivatives with payoffs at a future time T of the form $\left(e^{b \cdot X_T} - c\right)^+$, where c is a constant strike price, $b \in \mathbb{R}^n$, X is an AJD, and $y^+ \equiv \max(y, 0)$. This leads directly to pricing formulas for plain-vanilla options on currencies and equities, quanto options (such as an option on a common stock or bond struck in a different currency), options on zero-coupon bonds, caps, floors, chooser options, and other related derivatives. Furthermore, we can price payoffs of the form $(b \cdot X_T - c)^+$ and $(e^{a \cdot X_T} b \cdot X_T - c)^+$, allowing us to price "slope-of-the-yield-curve" options and certain Asian options.[6]

In order to visualize our approach to option pricing, consider the price p at date 0 of a call option with payoff $\left(e^{d \cdot X_T} - c\right)^+$ at date T, for given $d \in \mathbb{R}^n$ and strike c, where X is an n-dimensional AJD, with a short-term interest-rate process that is itself affine in X. For any real number y and any a and b in $\mathbb{R}^n$, let $G_{a,b}(y)$ denote the price of a security that pays $e^{a \cdot X_T}$ at time T in the event that $b \cdot X_T \leq y$. As the call option is in the money when $-d \cdot X_T \leq -\ln c$, and in that case pays $e^{d \cdot X_T} - ce^{0 \cdot X_T}$, we have the option priced at

$$p = G_{d,-d}(-\ln c) - c \, G_{0,-d}(-\ln c). \tag{1.6}$$

Because it is an increasing function, $G_{a,b}(\cdot)$ can be treated as a measure. Thus, it is enough to be able to compute the Fourier transform $\mathcal{G}_{a,b}(\cdot)$ of $G_{a,b}(\cdot)$, defined by

$$\mathcal{G}_{a,b}(z) = \int_{-\infty}^{+\infty} e^{izy} \, dG_{a,b}(y),$$

for then well-known Fourier-inversion methods can be used to compute terms of the form $G_{a,b}(y)$ in (1.6).

There are many cases in which the Fourier transform $\mathcal{G}_{a,b}(\cdot)$ of $G_{a,b}(\cdot)$ can be computed explicitly. We extend the range of solutions for the transform $\mathcal{G}_{a,b}(\cdot)$ from those already in the literature to include the entire class of

---

[6]In a complementary analysis of derivative security valuation, Bakshi and Madan [1999] show that knowledge of the special case of (1.1) with $v_0 + v_1 \cdot X_T = 1$ is sufficient to recover the prices of standard call options, but they do not provide explicit guidance as to how to compute this transform. Their applications to Asian and other options presumes that the state vector follows square-root or Heston-like stochastic-volatility models for which the relevant transforms had already been known in closed form.

*AJD*s by noting that $\mathcal{G}_{a,b}(z)$ is given by (1.1), for the complex coefficient vector $u = a + izb$, with $v_0 = 1$ and $v_1 = 0$. This, because of the affine structure, implies under regularity conditions that

$$\mathcal{G}_{a,b}(z) = e^{\alpha(0) + \beta(0) \cdot X_0}, \tag{1.7}$$

where α and β solve known, complex-valued ordinary differential equations (*ODE*s) with boundary conditions at T determined by z. In some cases, these ODEs have explicit solutions. These include independent square-root diffusion models for the short-rate process, as in Chen and Scott [1995], and the stochastic-volatility models of asset prices studied by Bates [1997] and Bakshi, Cao, and Chen [1997]. Using our *ODE*-based approach, we derive other explicit examples, for instance stochastic-volatility models with correlated jumps in both returns and volatility. In other cases, one can easily solve the *ODE*s for α and β numerically, even for high-dimensional applications.

Similar transform analysis provides a price for an option with a payoff of the form $(d \cdot X_T - c)^+$, again for the general *AJD* setting. For this case, we provide an equally tractable method for computing the Fourier transform of $\tilde{G}_{a,b,d}(\,\cdot\,)$, where $\tilde{G}_{a,b,d}(y)$ is the price of a security that pays $e^{d \cdot X_T} a \cdot X_T$ at T in the event that $b \cdot X_T \leq y$. This transform is again of the form (1.1), now with $v_1 = a$. Given this transform, we can invert to obtain $\tilde{G}_{a,b,d}(y)$ and the option price p' as

$$p' = \tilde{G}_{a,-a,0}(-\ln c) - c\,G_{0,-a}(-\ln c). \tag{1.8}$$

As shown in Section 3, these results can be used to price slope-of-the-yield-curve options and certain Asian options.

Our motivation for studying the general *AJD* setting is largely empirical. The *AJD* model takes the elements of the drift vector, "instantaneous" covariance matrix, and jump measure of X to be affine functions of X. This allows for conditional variances that depend on all of the state variables (unlike the Gaussian model), and for a variety of patterns of cross-correlations among the elements of the state vector (unlike the case of independent square-root diffusions). Dai and Singleton [1999], for instance, found that both time-varying conditional variances and negatively correlated state variables were essential ingredients to explaining the historical behavior of term structures of U.S. interest rates.

Furthermore, for the case of equity options, Bates [1997] and Bakshi, Cao, and Chen [1997] found that their affine stochastic-volatility models did not fully explain historical changes in the volatility smiles implied by $S\&P500$ index options. Within the affine family of models, one potential explanation for their findings is that they unnecessarily restricted the correlations between the state variables driving returns and volatility. Using the classification scheme for affine models found in Dai and Singleton [1999], one may nest these previous stochastic-volatility specifications within an AJD model with the same number of state variables that allows for potentially much richer correlation among the return and volatility factors.

The empirical studies of Bates [1997] and Bakshi, Cao, and Chen [1997] also motivate, in part, our focus on multivariate jump processes. They concluded that their stochastic-volatility models (with jumps in spot-market returns only) do not allow for a degree of volatility of volatility sufficient to explain the substantial "smirk" in the implied volatilities of index option prices. Both papers conjectured that jumps in volatility, as well as in returns, may be necessary to explain option-volatility smirks. Our AJD setting allows for correlated jumps in both volatility and price. Jumps may be correlated because their amplitudes are drawn from correlated distributions, or because of correlation in the jump times. (The jump times may be simultaneous, or have correlated stochastic arrival intensities.)

In order to illustrate our approach, we provide an example of the pricing of plain-vanilla calls on the $S\&P500$ index. A cross-section of option prices for a given day are used to calibrate AJDs with simultaneous jumps in both returns and volatility. Then we compare the implied-volatility smiles to those observed in the market on the chosen day. In this manner we provide some preliminary evidence on the potential role of jumps in volatility for resolving the volatility puzzles identified by Bates [1997] and Bakshi, Cao, and Chen [1997].

The remainder of this paper is organized as follows. Section 2 reviews the class of affine jump-diffusions, and shows how to compute some relevant transforms, and how to invert them. Section 3 presents our basic option-pricing results. The example of the pricing of plain-vanilla calls on the $S\&P500$ index is presented in Section 4. Additional appendices provide various technical results and extensions.

8

2 Transform Analysis for *AJD* State-Vectors

This section presents the *AJD* state-process model and the basic transform calculations that will later be useful in option pricing.

2.1 The Affine Jump-Diffusion

We fix a probability space $(\Omega, \mathcal{F}, P)$ and an information filtration[7] $(\mathcal{F}_t)$, and suppose that X is a Markov process in some state space $D \subset \mathbb{R}^n$, solving the stochastic differential equation

$$dX_t = \mu(X_t)\, dt + \sigma(X_t)\, dW_t + dZ_t \,, \tag{2.1}$$

where W is an $(\mathcal{F}_t)$-standard Brownian motion in $\mathbb{R}^n$; $\mu : D \to \mathbb{R}^n$, $\sigma : D \to \mathbb{R}^{n \times n}$, and Z is a pure jump process whose jumps have a fixed probability distribution ν on $\mathbb{R}^n$ and arrive with intensity $\{\lambda(X_t) : t \geq 0\}$, for some $\lambda : D \to [0, \infty)$. To be precise, we suppose that X is a Markov process whose transition semi-group has an infinitesimal generator[8] $\mathcal{D}$ of the Lévy type, defined at a bounded C^2 function $f : D \to \mathbb{R}$, with bounded first and second derivatives, by

$$\mathcal{D}f(x,t) = f_t(x,t) + f_x(x,t)\mu(x) + \frac{1}{2}\mathrm{tr}\left[f_{xx}(x,t)\sigma(x)\sigma(x)^\top \right]$$
$$+ \lambda(x) \int_{\mathbb{R}^n} \left[f(x+z,t) - f(x,t) \right] d\nu(z), \tag{2.2}$$

Intuitively, this means that, conditional on the path of X, the jump times of Z are the jump times of a Poisson process with time-varying intensity $\{\lambda(X_s) : 0 \leq s \leq t\}$, and that the size of the jump of Z at a jump time T is independent of $\{X_s : 0 \leq s < T\}$ and has the probability distribution ν.

For notational convenience, we assume that X_0 is "known" (has a trivial distribution). Appendices provide additional technical details, as well as generalizations to multiple jump types with different arrival intensities, and to time-dependent $(\mu, \sigma, \lambda, \nu)$.

---

[7]The filtration $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$ is assumed to satisfy the usual conditions, and X is assumed to be Markov relative to $(\mathcal{F}_t)$. For technical details, see, for example, Ethier and Kurtz [1986].

[8]The generator $\mathcal{D}$ is defined by the property that $\{f(X_t, t) - \int_0^t \mathcal{D}f(X_s, s)\, ds : t \geq 0\}$ is a martingale for any f in its domain. See Ethier and Kurtz [1986] for details.

We impose an "affine" structure on μ, $\sigma\sigma^\top$, and λ, in that all of these functions are assumed to be affine on D. In order for X to be well defined, there are joint restrictions on $(D, \mu, \sigma, \lambda, \nu)$, as discussed in Duffie and Kan [1996] and Dai and Singleton [1999].

2.2 Transforms

First, we show that the Fourier transform of X_t and of certain related random variables is known in closed form up to the solution of an ordinary differntial equation (ODE). Then, we show how the distribution of X_t and the prices of options can be recovered by inverting this transform.

We fix an affine discount-rate function $R : D \to \mathbb{R}$. The affine dependence of μ, $\sigma\sigma^\top$, λ, and R are determined by coefficients (K, H, l, ρ) defined by:

- $\mu(x) = K_0 + K_1 x$, for $K = (K_0, K_1) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}$.

- $(\sigma(x)\sigma(x)^\top)_{ij} = (H_0)_{ij} + (H_1)_{ij} \cdot x$, for $H = (H_0, H_1) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n \times n}$.

- $\lambda(x) = l_0 + l_1 \cdot x$, for $l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^n$.

- $R(x) = \rho_0 + \rho_1 \cdot x$, for $\rho = (\rho_0, \rho_1) \in \mathbb{R} \times \mathbb{R}^n$.

For $c \in \mathbb{C}^n$, the set of n-tuples of complex numbers, we let $\theta(c) = \int_{\mathbb{R}^n} \exp(c \cdot z) \, d\nu(z)$ whenever the integral is well defined. This "jump transform" θ determines the jump-size distribution.

The "coefficients" (K, H, l, θ) of X completely determine its distribution, given an initial condition $X(0)$. A "characteristic" $\chi = (K, H, l, \theta, \rho)$ captures both the distribution of X as well as the effects of any discounting, and determines a transform $\psi^\chi : \mathbb{C}^n \times D \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{C}$ of X_T conditional on $\mathcal{F}_t$, when well defined at $t \leq T$, by

$$\psi^\chi(u, X_t, t, T) = E^\chi \left(\exp\left(-\int_t^T R(X_s) \, ds \right) e^{u \cdot X_T} \,\bigg|\, \mathcal{F}_t \right), \qquad (2.3)$$

where E^χ denotes expectation under the distribution of X determined by χ. Here, ψ^χ differs from the familiar (conditional) characteristic function of the distribution of X_T because of the discounting at rate $R(X_t)$.

The key to our applications is that, under technical regularity conditions given in Proposition 1 below,

$$\psi^\chi(u, x, t, T) = e^{\alpha(t) + \beta(t) \cdot x}, \qquad (2.4)$$

where β and α satisfy the complex-valued $ODEs$[9]

$$\dot{\beta}(t) = \rho_1 - K_1^\top \beta(t) - \frac{1}{2}\beta(t)^\top H_1 \beta(t) - l_1\left(\theta(\beta(t)) - 1\right), \qquad (2.5)$$

$$\dot{\alpha}(t) = \rho_0 - K_0 \cdot \beta(t) - \frac{1}{2}\beta(t)^\top H_0 \beta(t) - l_0\left(\theta(\beta(t)) - 1\right), \qquad (2.6)$$

with boundary conditions $\beta(T) = u$ and $\alpha(T) = 0$. The ODE (2.5)-(2.6) is easily conjectured from an application of Ito's Formula to the candidate form (2.4) of ψ^X. In order to apply our results, we would need to compute solutions α and β to these $ODEs$. In some applications, as for example in Section 4, explicit solutions can be found. In other cases, solutions would be found numerically, for example by Runge-Kutta. This suggests a practical advantage of choosing a jump distribution ν with an explicitly known or easily computed jump transform θ.

The following technical conditions will justify this method of calculating the transform.

Definition: A characteristic (K, H, l, θ, ρ) is *well-behaved* at $(u, T) \in \mathbb{C}^n \times [0, \infty)$ if (2.5)-(2.6) are solved uniquely by β and α; and if

(i) $E\left(\int_0^T |\gamma_t| \, dt\right) < \infty$, where $\gamma_t = \Psi_t\left(\theta(\beta(t)) - 1\right)\lambda(X_t)$,

(ii) $E\left[\left(\int_0^T \eta_t \cdot \eta_t \, dt\right)^{1/2}\right] < \infty$, where $\eta_t = \Psi_t\, \beta(t)^\top \sigma(X_t)$, and

(iii) $E\left(|\Psi_T|\right) < \infty$,

where $\Psi_t = \exp\left(-\int_0^t R(X_s)\, ds\right) e^{\alpha(t) + \beta(t)\cdot X(t)}$.

Proposition 1. Suppose (K, H, l, θ, ρ) is well-behaved at (u, T). Then the transform ψ^X of X defined by (2.3) is given by (2.4).

Proof: It is enough to show that Ψ is a martingale, for then $\Psi_t = E(\Psi_T \mid \mathcal{F}_t)$, and we can multiply Ψ_t by $\exp\left(\int_0^t R(X_s)\, ds\right)$ to get the result. By Ito's Formula,[10]

$$\Psi_t = \Psi_0 + \int_0^t \Psi_s \mu_\Psi(s)\, ds + \int_0^t \eta_s\, dW_s + J_t\,, \qquad (2.7)$$

---

[9]Here, $c^\top H_1 c$ denotes the vector in $\mathbb{C}^n$ with k-th element $\sum_{i,j} c_i (H_1)_{ijk} c_j$.

[10]See Protter [1990] for a complex version of Ito's Formula.

11

where, using the fact that α and β satisfy the *ODE* (2.5)-(2.6), we have $\mu_\Psi = 0$, and where

$$J_t = \sum_{0 < \tau(i) \leq t} \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-} \right) - \int_0^t \gamma_s \, ds,$$

with $\tau(i) = \inf\{t : N_t = i\}$ denoting the i-th jump time of X. Under the integrability condition (i), Lemma 1 of Appendix A implies that J is a martingale. Under integrability condition (ii), $\int \eta \, dW$ is a martingale. Thus Ψ is a martingale and we are done. ∎

Anticipating the application to option pricing, for each given $(d, c, T) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}_+$, our next goal is to compute (when well defined) the "expected present value"

$$C\left(d, c, T, \chi\right) = E^\chi \left(\exp\left(-\int_0^T R(X_s) \, ds \right) \left(e^{d \cdot X_T} - c \right)^+ \right). \qquad (2.8)$$

We have

$$C\left(d, c, T, \chi\right) = E^\chi \left(\exp\left(-\int_0^T R(X_s) \, ds \right) \left(e^{d \cdot X_T} - c \right) \mathbf{1}_{d \cdot X_T \geq \ln(c)} \right)$$
$$= G_{d,-d}\left(-\ln(c); X_0, T, \chi \right) - c \, G_{0,-d}\left(-\ln(c); X_0, T, \chi \right), \qquad (2.9)$$

where, given some $(x, T, a, b) \in D \times [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$, $G_{a,b}(\,\cdot\,; x, T, \chi) : \mathbb{R} \to \mathbb{R}_+$ is given by

$$G_{a,b}(y; X_0, T, \chi) = E^\chi \left(\exp\left(-\int_0^T R(X_s) \, ds \right) e^{a \cdot X_T} \mathbf{1}_{b \cdot X_T \leq y} \right). \qquad (2.10)$$

The Fourier-Stieltjes transform $\mathcal{G}_{a,b}(\,\cdot\,; X_0, T, \chi)$ of $G_{a,b}(\,\cdot\,; X_0, T, \chi)$, if well defined, is given by

$$\mathcal{G}_{a,b}(v; X_0, T, \chi) = \int_{\mathbb{R}} e^{ivy} \, dG_{a,b}(y; X_0, T, \chi)$$
$$= E^\chi \left(\exp\left(-\int_0^T R(X_s) \, ds \right) \exp\left[(a + ivb) \cdot X_T \right] \right)$$
$$= \psi^\chi(a + ivb, X_0, 0, T).$$

We may now extend the Lévy inversion formula[11] (from the typical case of a proper cumulative distribution function) to obtain the following result.

Proposition 2 (Transform Inversion).   Suppose, for fixed $T \in [0, \infty)$, $a \in \mathbb{R}^n$, and $b \in \mathbb{R}^n$, that $\chi = (K, H, l, \theta, \rho)$ is well-behaved at $(a + ivb, T)$ for any $v \in \mathbb{R}$, and that

$$\int_{\mathbb{R}} |\psi^\chi(a + ivb, x, 0, T)| \, dv < \infty. \qquad (2.11)$$

Then $G_{a,b}(\,\cdot\,; x, T, \chi)$ is well defined by (2.10) and given by

$$G_{a,b}(y; X_0, T, \chi) = \frac{\psi^\chi(a, X_0, 0, T)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\mathrm{Im}\left[\psi^\chi(a + ivb, X_0, 0, T)e^{-ivy}\right]}{v} \, dv \,,$$
$$(2.12)$$

where $\mathrm{Im}(c)$ denotes the imaginary part of $c \in \mathbb{C}$.

A proof is given in Appendix A. For $R = 0$, this formula gives us the probability distribution function of $b \cdot X_T$. The associated transition density of X is obtained by differentiation of $G_{a,b}$. More generally, this provides the transition function of X with "killing" at rate[12] R.

2.3   Extended Transform

As noted in the introduction, certain pricing problems in our setting, for example Asian option valuation or default-time distributions, call for the calculation of the expected present value of the product of affine and exponential-affine functions of X_T. Accordingly, we define the "extended" transform $\phi^\chi : \mathbb{R}^n \times \mathbb{C}^n \times D \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{C}$ of X_T conditional on $\mathcal{F}_t$, when well defined for $t \leq T$ by

$$\phi^\chi(v, u, X_t, t, T) = E\left(\exp\left(-\int_t^T R(X_s) \, ds\right)(v \cdot X_T) \, e^{u \cdot X_T} \,\bigg|\, \mathcal{F}_t\right). \quad (2.13)$$

The extended transform ϕ^χ can be computed by differentiation of the transform ψ^χ, just as moments can be computed from a moment-generating

---

[11]See, for example, Gil-Pelaez [1951] and Williams [1991] for a treatment of the Lévy inversion formula.

[12]A negative R is sometimes called a "creation" rate in Markov-process theory.

function (under technical conditions justifying differentiation through the expectation.) In practice, computing the derivatives of the transform calls for solving a new set of *ODEs*, as indicated below. Specifically, under technical conditions, including the differentiability of the jump transform θ, we show that

$$\phi^\chi(v, u, x, t, T) = \psi^\chi(u, x, t, T)\left(A(t) + B(t) \cdot x\right), \qquad (2.14)$$

where ψ^χ is given by (2.4), and where B and A satisfy the linear ordinary differential equations

$$-\dot{B}(t) = K_1(t)^\top B(t) + \beta(t)^\top H_1(t)B(t) + l_1\,\nabla\theta\left(\beta(t)\right)B(t) \qquad (2.15)$$

$$-\dot{A}(t) = K_0(t) \cdot B(t) + \beta(t)^\top H_0(t)B(t) + l_0\nabla\theta\left(\beta(t)\right)B(t), \qquad (2.16)$$

with the boundary conditions $B(T) = v$ and $A(T) = 0$, and where $\nabla\theta(c)$ is gradient of $\theta(c)$ with respect to $c \in \mathbb{C}^n$.

Proposition 3: Suppose $\chi = (K, H, l, \theta, \rho)$ is *"extended" well-behaved* at (v, u, T), a technical condition stated in Appendix A. Then the extended transform ϕ^χ defined by (2.13) is given by (2.14).

One could further extend this approach so as to calculate higher-order moments, as in Pan [1998].

3   Option Pricing Theory

This section applies our basic transform analysis to the pricing of options. In all cases, we assume that the price process S of the asset underlying the option is of the form $S_t = \left(\bar{a}(t) + \bar{b}(t) \cdot X_t\right)e^{a(t) + b(t) \cdot X_t}$, for deterministic $\bar{a}(t), \bar{b}(t), a(t)$, and $b(t)$. This is the case for many applications in affine settings, including underlying assets that are equities, currencies, and zero-coupon bonds.

  Two traditional formulations[13] of the asset-pricing problem are:

---

[13]A popular variant was developed in a Gaussian setting by Jamshidian [1989]. In a setting in which X is an affine jump-diffusion under the equivalent martingale measure Q, one normalizes the underlying exponential-affine asset price by the price of a zero-coupon bond maturing on the option expiration date T. Then, in the new numeraire, the short-rate process is of course zero, and there is a new equivalent martingale measure $Q(T)$,

1. Model the "risk-neutral" behavior of X under an equivalent martingale measure Q. That is, take X to be an affine jump-diffusion under Q with given characteristic χ_Q. Then apply (2.9) and (2.12).

2. Model the behavior of X as an affine jump-diffusion under the actual (that is, the "data-generating") measure P. If one then supposes that the state-price density (also known as the "pricing kernel" or "marginal-rate-of-substitution" process) is an exponential-affine form in X, then X is also an affine jump-diffusion under Q, and one can either:

   (a) calculate, as in Appendix C, the implied equivalent martingale measure Q and associated characteristic χ_Q of X under Q, and proceed as in the first alternative above, or

   (b) simply apply the definition of the state-price density, which determines the price of an option directly in terms of $G_{a,b}$, computed using our transform analysis. This alternative is sketched in Section 3.2 below.

Of course the two approaches are consistent, and indeed the second formulation implies the first, as indicated. The second approach is more complete, and would be indicated for empirical time-series applications, for which the "actual" distribution of the state process X as well as the parameters determining risk-premia must be specified and estimated.

Applications of these approaches to call-option pricing are briefly sketched in the next two sub-sections. Other derivative pricing applications are provided in Section 3.3.

3.1 Risk-Neutral Pricing

Here, we take Q to be an equivalent martingale measure associated with a short-term interest rate process defined by $R(X_t) = \rho_0 + \rho_1 \cdot X_t$. This means that the market value at time t of any contingent claim that pays an

---

often called the "forward measure," under which prices are exponential affine. Application of Girsanov's Theorem uncovers new affine behavior for the underlying state process X under $Q(T)$, and one can proceed as before. The change-of-measure calculations for this approach can be found in Appendix C.

$\mathcal{F}_T$-measurable random variable V at time T is, by definition,

$$E^Q \left(\exp \left(- \int_t^T R(X_s) \, ds \right) V \; \Big| \; \mathcal{F}_t \right), \qquad (3.1)$$

where, under Q, the state vector X is assumed to be an AJD with coefficients $(K^Q, H^Q, l^Q, \theta^Q)$. The relevant characteristic for risk-neutral pricing is then $\chi_Q = \left(K^Q, H^Q, l^Q, \theta^Q, \rho \right)$. It need not be the case that markets are complete. The existence of some equivalent martingale measure and the absence of arbitrage are in any case essentially equivalent properties, under technical conditions, as pointed out by Harrison and Kreps [1979]. For recent technical conditions, see for example Delbaen and Schachermayer [1994].

We let S denote the price process for the security underlying the option, and suppose for simplicity[14] that $\ln S_t = X_t^{(i)}$, an element of the state vector $X = \left(X^{(1)}, \ldots, X^{(n)} \right)$. Other components of the state process X may jointly specify the arrival intensity of jumps, the behavior of stochastic volatility, the behavior of other asset returns, interest-rate behavior, and so on. The given asset is assumed to have a dividend-yield process $\{\zeta(X_t) : t \geq 0\}$ defined by

$$\zeta(x) = q_0 + q_1 \cdot x, \qquad (3.2)$$

for given $q_0 \in \mathbb{R}$ and $q_1 \in \mathbb{R}^n$. For example, if the asset is a foreign currency, then $\zeta(X_t)$ is the foreign short-term interest rate.

Because Q is an equivalent martingale measure, the coefficients $K_i^Q = ((K_0^Q)_i, (K_1^Q)_i)$ determining[15] the "risk-neutral" drift of $X^{(i)} = \ln S$ are given

---

[14]The more general case of $S_t = \exp(a_t + b_t \cdot X_t)$ can be similarly treated. Possibly after some innocuous affine change of variables in the state vector, possibly involving time dependencies in the characteristic χ, we can always reduce to the assumed case.

[15]Under (3.3)-(3.4), we have

$$S_t - S_0 = \int_0^t S_u \left[R(X_u) - \zeta(X_u) \right] du + \int_0^t S_u \sigma^{(i)} (X_u)^\top \, dW_u^Q$$

$$+ \sum_{0 < u \leq t} S_{u-} \left(\exp(\Delta X_u^{(i)}) - 1 \right) - \int_0^t S_u \left(\theta^Q(\epsilon(i)) - 1 \right) (l_0^Q + l_1^Q \cdot X_u) \, du,$$

where W^Q is an $(\mathcal{F}_t)$-standard Brownian motion in $\mathbb{R}^n$ under Q. (Here, $\Delta X_t = X_t - X_{t-}$ denotes the jump of X at t.) As the sum of the last 3 terms is a local Q-martingale, this indeed implies consistency with the definition of an equivalent martingale measure.

by

$$(K_0^Q)_i = \rho_0 - q_0 - \frac{1}{2}(H_0^Q)_{ii} - l_0^Q \left(\theta^Q(\epsilon(i)) - 1 \right) \tag{3.3}$$

$$(K_1^Q)_i = \rho_1 - q_1 - \frac{1}{2}(H_1^Q)_{ii} - l_1^Q \left(\theta^Q(\epsilon(i)) - 1 \right), \tag{3.4}$$

where $\epsilon(i) \in \mathbb{R}^n$ has 1 as its i-th component, and every other component equal to 0.

Unless other security price processes are specified, the risk-neutral characteristic χ_Q is otherwise unrestricted by arbitrage considerations. There are analogous no-arbitrage restrictions on χ_Q for each additional specified security price process of the form $e^{a+b\cdot X_t}$.

By the definition of an equivalent martingale measure and the results of Section 2.2, a plain-vanilla European call option with expiration time T and strike c has a price p at time 0 which is given by (2.9) to be

$$p = G_{\epsilon(i),-\epsilon(i)} \left(-\ln(c); X_0, T, \chi_Q \right) - c\, G_{0,-\epsilon(i)} \left(-\ln(c); X_0, T, \chi_Q \right). \tag{3.5}$$

To be precise, we can exploit Propositions 1 and 2 and summarize this option-pricing tool as follows, extending Heston [1993], Bates [1996], Scott [1997], Bates [1997], Bakshi and Madan [1999], and Bakshi, Cao, and Chen [1997].

Proposition 4. The option-pricing formula (2.9) applies, where G is computed by (2.12), provided:

(a) χ is well-behaved at $(d - ivd, T)$ and at $(-ivd, T)$, for all $v \in \mathbb{R}_+$, and

(b) $\int_{\mathbb{R}} |\psi^\chi(d - ivd, x, 0, T)|\, dv < \infty$, and $\int_{\mathbb{R}} |\psi^\chi(-ivd, x, 0, T)|\, dv < \infty$.

3.2 State-Price Density

Suppose the state vector X is an affine jump-diffusion with coefficients (K, H, l, θ) under the actual (data-generating) measure P. Let ξ be an $(\mathcal{F}_t)$-adapted "state-price density," defined by the property that the market value at time t of any security that pays an $\mathcal{F}_T$-measurable random variable V at time T is given by

$$\frac{1}{\xi(t)} E \left(V\xi(T) \mid \mathcal{F}_t \right).$$

We assume for convenience that $\xi_t = e^{a(t)+b(t)\cdot X_t}$, for some bounded measurable $a : [0,\infty) \to \mathbb{R}$ and $b : [0,\infty) \to \mathbb{R}^n$. Without loss of generality, we take it that $\xi(0) = 1$.

Suppose the price of a given underlying security at time T is $e^{d\cdot X(T)}$, for some $d \in \mathbb{R}^n$. By the definition of a state-price density, a plain-vanilla European call option struck at c with exercise date T has a price at time 0 of

$$p = E\left[e^{a(T)+b(T)\cdot X(T)}(e^{d\cdot X(T)} - c)^+\right].$$

This leaves the option price

$$p = e^{a(T)}G_{b(T)+d,-d}(-\ln c; X_0, T, \chi^0) - ce^{a(T)}G_{b(T),-d}(-\ln c; X_0, T, \chi^0),$$

where $\chi^0 = (K, H, l, \theta, 0)$. (One notes that the short-rate process plays no role beyond that already captured by the state-price density.)

As mentioned at the beginning of this section, and detailed in Appendix C, an alternative is to translate the option-pricing problem to a "risk-neutral" setting.

3.3 Other Option-Pricing Applications

This section develops as illustrative examples several additional applications to option pricing. For convenience, we adopt the risk-neutral pricing formulation. That is, we suppose that the short rate is given by $R(X)$, where R is affine, and X is an affine jump-diffusion under an equivalent martingale measure Q. The associated characteristic χ_Q is fixed. While we treat the case of call options, put options can be treated by the same method, or by put-call parity.

3.3.1 Bond Derivatives

Consider a call option, struck at c with exercise date T, on a zero-coupon bond maturing at time $s > T$. Let $\Lambda(T,s)$ denote the time-T market price of the underlying bond. From Duffie and Kan [1996], under the regularity conditions given in Section 2.2,

$$\Lambda(T,s) = \exp\left(\alpha(T, s, 0) + \beta(T, s, 0) \cdot X_T\right),$$

where, from this point, for any u we write $\beta(t, T, u)$ and $\alpha(t, T, u)$ for the solution to (2.5)-(2.6), adding the arguments (T, u) so as to indicate the

dependence on the terminal time T and boundary condition u for β, which will vary in what follows. At time T, the option pays

$$
\begin{aligned}
(\Lambda(T,s)-c)^+ &= \left(e^{\alpha(T,s,0)+\beta(T,s,0)\cdot X(T)}-c\right)^+ & (3.6)\\
&= e^{\alpha(T,s,0)}\left(e^{\beta(T,s,0)\cdot X(T)}-e^{-\alpha(T,s,0)}c\right)^+. & (3.7)
\end{aligned}
$$

The value of the bond option can therefore be obtained from (2.9) and (2.12). The same approach applies to caps and floors, which are simply portfolios of zero-coupon bond options with payment in arrears, as reviewed in Appendix D. This extends the results of Chen and Scott [1995] and Scott [1996]. Chacko and Das [1998] work out the valuation of Asian interest-rate options for a large class of affine models. They provide numerical examples based on a multi-factor Cox-Ingersoll-Ross state vector.

3.3.2 Quantos

Consider a quanto of exercise date T and strike c on an underlying asset with price process $S = \exp(X^{(i)})$. The time-T payoff of the quanto is $(S_T M(X_T) - c)^+$, where $M(x) = e^{m\cdot x}$, for some $m \in \mathbb{R}^n$. The quanto scaling $M(X_T)$ could, for example, be the price at time T of a given asset, or the exchange rate between two currencies. The initial market value of the quanto option is then

$$
G_{m+\epsilon(i),-\epsilon(i)}\left(-\ln(c);x,T,\chi_Q\right) - c\,G_{0,-\epsilon(i)}\left(-\ln(c);x,T,\chi_Q\right).
$$

An alternative form of the quanto option pays $M(X_T)(S_T - c)^+$ at T, and has the price $G_{m+\epsilon(i),-\epsilon(i)}\left(-\ln(c);x,T,\chi_Q\right) - c\,G_{m,-\epsilon(i)}\left(-\ln(c);x,T,\chi_Q\right)$.

3.3.3 Foreign Bond Options

Let $\exp(X^{(i)})$ be a foreign-exchange rate, $R(X)$ be the domestic short interest rate, and $\zeta(X)$ be the foreign short rate, for affine ζ. Consider a foreign zero-coupon bond maturing at time s, whose payoff at maturity, in domestic currency, is therefore $\exp(X_s^{(i)})$. The risk-neutral characteristic χ_Q is restricted by (3.3)-(3.4). From Proposition 1, the domestic price at time t of the foreign bond is $\Lambda^f(t,s) = \exp(\alpha(t,s,\epsilon(i)) + \beta(t,s,\epsilon(i))\cdot X_t)$.

We now consider an option on this bond with exercise date $T < s$ and domestic strike price c on the foreign s-year zero-coupon bond, paying $(\Lambda^f(T,s)-c)^+$ at time T, in domestic currency. The initial market value of this option can therefore be obtained as for a domestic bond option.

3.3.4 Chooser Options

Let $S^{(i)} = \exp(X^{(i)})$ and $S^{(j)} = \exp(X^{(j)})$ be two security price processes. An exchange, or "chooser," option with exercise date T, pays $\max(S_T^{(i)}, S_T^{(j)})$. Depending on their respective dividend payout rates, the risk-neutral characteristic χ_Q is restricted by (3.3)-(3.4), applied to both i and j. The initial market value of this option is

$$G_{\epsilon(i),\epsilon(i)-\epsilon(j)}(0; x, T, \chi_Q) + G_{\epsilon(j),0}(0, x, T, \chi_Q) - G_{\epsilon(j),\epsilon(j)-\epsilon(i)}(0; x, T, \chi_Q).$$

3.3.5 Asian Options

Under the assumption of a deterministic short rate and dividend-yield process, that is, $\rho_1 = q_1 = 0$, we may also use the extended transform analysis of Section 2.3 to price Asian options. Let $X^{(i)}$ be the underlying price process of an Asian option with strike price c and expiration date T. The option pays $\left(\frac{1}{T}\int_0^T X_t^{(i)} dt - c\right)$ at the expiration date T. If Q is an equivalent martingale measure, we must have

$$dX_t^{(i)} = (R(X_t, t) - \zeta(X_t, t)) X_t \, dt + dM_t^{(i)},$$

where M^i is a Q-martingale. For any $0 \le t \le T$, let $Y_t = \int_0^t X_s^{(i)} ds$. For short rate ρ_0, we can let $\tilde{\rho}_0 = (\rho_0, 0)$ and $\tilde{\rho}_1 = (0, 0) = 0$, and see that, under Q, $\tilde{X} = (X, Y)$ is an $(n+1)$-dimensional affine jump diffusion with characteristic $\tilde{\chi} = \left(\tilde{K}, \tilde{H}, \tilde{l}, \tilde{\theta}, \tilde{\rho}\right)$ that can be easily derived from using the fact that $dY_t = X_t^{(i)} dt$. We thus obtain the initial market value of the Asian option, under technical regularity, as[16]

$$\frac{1}{T}\tilde{G}_{\epsilon(n+1),-\epsilon(n+1),0}\left(-cT; \tilde{X}_0, T, \tilde{\chi}\right) - c\, G_{0,-\epsilon(n+1)}\left(-cT; \tilde{X}_0, T, \tilde{\chi}\right),$$

where $G(\,\cdot\,)$ is given by (2.12) and where, for a, b, and d in $\mathbb{R}^n$,

$$\tilde{G}_{a,b,d}(y; x, T, \chi) = \frac{\phi^\chi(a, d, x, 0, T)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}\left[\phi^\chi(a, d + ivb, x, 0, T)e^{-ivy}\right]}{v}\, dv.$$

$$(3.8)$$

---

[16]In this context, $\epsilon(i) \in \mathbb{R}^{n+1}$ has 1 as its i-th component, and every other component equal to 0.

This calculation of $\tilde{G}_{a,b,d}$ and the Asian option price is in parallel with the calculation (2.12) of $G_{a,b}$, using Fourier-inversion of the extended transform ϕ^χ, and is justified provided that $\tilde{\chi}$ is extended well behaved at $(a, d+ivb, T)$ for any $v \in \mathbb{R}$, and that $\int_{\mathbb{R}} |\phi^{\tilde{\chi}}(a, d+ivb, x, 0, T)| \, dv < \infty$.

As zero-coupon bond yields in an AJD setting are affine, we can also apply the extended-transform approach to the valuation of slope-of-the-yield-curve options.

4 A "Double-Jump" Illustrative Model

As an illustration of the methodology, this section provides explicit transforms for a 2-dimensional affine jump-diffusion model. We suppose that S is the price process, strictly positive, of a security that pays dividends at a constant proportional rate $\bar{\zeta}$, and we let $Y = \ln(S)$. The state process is $X = (Y, V)^\top$, where V is the volatility process.

We suppose for simplicity that the short rate is a constant r, and that there exists an equivalent martingale measure Q, under which[17]

$$
d \begin{pmatrix} Y_t \\ V_t \end{pmatrix} = \begin{pmatrix} r - \bar{\zeta} - \bar{\lambda}\bar{\mu} - \frac{1}{2}V_t \\ \kappa_v(\bar{v} - V_t) \end{pmatrix} dt + \sqrt{V_t} \begin{pmatrix} 1 & 0 \\ \bar{\rho}\sigma_v & \sqrt{1-\bar{\rho}^2}\,\sigma_v \end{pmatrix} dW_t^Q + dZ_t,
$$
(4.1)

where W^Q is an $(\mathcal{F}_t)$-standard Brownian motion under Q in $\mathbb{R}^2$, and Z is a pure jump process in $\mathbb{R}^2$ with constant mean jump-arrival rate $\bar{\lambda}$, whose bivariate jump-size distribution ν has the transform θ. A flexible range of distributions of jumps can be explored through the specification of θ. The risk-neutral coefficient restriction (3.3) is satisfied if and only if $\bar{\mu} = \theta(1,0) - 1$.

Before we move on to special examples, we lay out the formulation for option pricing as a straightforward application of our earlier results. At time t, the transform[18] ψ of the log-price state variable Y_T can be calculated using the ODE approach in (2.6) as:

$$
\psi(u, (y, v), t, T) = \exp\left(\bar{\alpha}(T-t, u) + uy + \bar{\beta}(T-t, u)v\right), \qquad (4.2)
$$

---

[17]Unless otherwise stated, the distributional properties of (Y, V) described in this section are in a "risk-neutral" sense, that is, under Q.

[18]That is, $\psi(u, (y, v)', t, T) = \psi^\chi((u, 0)', (y, v)', t, T)$, where χ is the characteristic under Q of X associated with the short rate defined by $(\rho_0, \rho_1) = (r, 0)$.

where, letting $b = \sigma_v \bar{\rho} u - \kappa_v$, $a = u(1-u)$, and[19] $\gamma = \sqrt{b^2 + a\sigma_v^2}$, we have

$$\bar{\beta}(\tau, u) = -\frac{a\left(1 - e^{-\gamma\tau}\right)}{2\gamma - (\gamma + b)\left(1 - e^{-\gamma\tau}\right)}, \tag{4.3}$$

$$\bar{\alpha}(\tau, u) = \alpha_0(\tau, u) - \bar{\lambda}\tau(1 + \bar{\mu}u) + \bar{\lambda}\int_0^\tau \theta\left(u, \bar{\beta}(s, u)\right) ds, \tag{4.4}$$

where[20]

$$\alpha_0(\tau, u) = -r\tau + (r - \bar{\zeta})u\tau - \kappa_v \bar{v}\left(\frac{\gamma + b}{\sigma_v^2}\tau + \frac{2}{\sigma_v^2}\ln\left[1 - \frac{\gamma + b}{2\gamma}(1 - e^{-\gamma\tau})\right]\right),$$

and where the term $\int_0^\tau \theta(u, \bar{\beta}(s, u))\, ds$ depends on the specific formulation of bivariate jump transform $\theta(\,\cdot\,, \cdot\,)$.

4.1 A Concrete Example

As a concrete example, consider the jump transform θ defined by

$$\theta(c_1, c_2) = \bar{\lambda}^{-1}\left(\lambda^y \theta^y(c_1) + \lambda^v \theta^v(c_2) + \lambda^c \theta^c(c_1, c_2)\right), \tag{4.5}$$

where $\bar{\lambda} = \lambda^y + \lambda^v + \lambda^c$, and where

$$\theta^y(c) = \exp\left(\mu_y c + \frac{1}{2}\sigma_y^2 c^2\right),$$

$$\theta^v(c) = \frac{1}{1 - \mu_v c},$$

$$\theta^c(c_1, c_2) = \frac{\exp\left(\mu_{c,y}c_1 + \frac{1}{2}\sigma_{c,y}^2 c_1^2\right)}{1 - \mu_{c,v}c_2 - \rho_J \mu_{c,v}c_1}.$$

What we incorporate in this example is in fact three types of jumps:

- Jumps in Y, with arrival intensity λ^y and normally distributed jump size with mean μ_y and variance σ_y^2,

---

[19]To be more precise, $\gamma = |\gamma^2|^{1/2}\exp\left(\frac{i\arg(\gamma^2)}{2}\right)$, where $\gamma^2 = b^2 + a\sigma_v^2$. Note that for any $z \in \mathbb{C}$, $\arg(z)$ is defined such that $z = |z|\exp(i\arg(z))$, with $-\pi < \arg(z) \leq \pi$.

[20]For any $z \in \mathbb{C}$, $\ln(z) = \ln|z| + i\arg(z)$, as defined on the "principal branch."

- Jumps in V, with arrival intensity λ^v and exponentially distributed jump size with mean μ_v,

- Simultaneous correlated jumps in Y and V, with arrival intensity λ^c. The marginal distribution of the jump size in V is exponential with mean $\mu_{c,v}$. Conditional on a realization, say z_v, of the jump size in V, the jump size in Y is normally distributed with mean $\mu_{c,y} + \rho_J z_v$, and variance $\sigma_{c,y}^2$.

In Bakshi, Cao, and Chen [1997] and Bates [1997], the *SVJ-Y* model, defined by $\lambda^v = \lambda^c = 0$, was studied using cross sections of options data to fit the "volatility smirk." They find that allowing for negative jumps in Y is useful insofar as it increases the skewness of the distribution of Y_T, but that this does not generate the level of skewness implied by the volatility smirk observed in market data. They call for a model with jumps in volatility. Using this concrete "double-jump" example (4.5), we can address this issue, and provide some insights into what a richer specification of jumps may imply.

Before leaving this section to explore the implications of jumps for "volatility smiles," we provide explicit option pricing through the transform formula (4.2), by exploiting the bivariate jump transform θ specified in (4.5). We have

$$\int_0^\tau \theta(u, \bar{\beta}(s, u))\, ds = \overline{\lambda}^{-1}\left(\lambda^y f^y(u, \tau) + \lambda^v f^v(u, \tau) + \lambda^c f^c(u, \tau)\right),$$

where

$$f^y(u, \tau) = \tau \exp\left(\mu_y u + \frac{1}{2}\sigma_y^2 u^2\right),$$

$$f^v(u, \tau) = \frac{\gamma - b}{\gamma - b + \mu_v a}\tau - \frac{2\mu_v a}{\gamma^2 - (b - \mu_v a)^2}\ln\left(1 - \frac{(\gamma + b) - \mu_v a}{2\gamma}(1 - e^{-\gamma\tau})\right),$$

$$f^c(u, \tau) = \exp\left(\mu_{c,y} u + \sigma_{c,y}^2\frac{u^2}{2}\right)d,$$

where $a = u(1 - u)$, $b = \sigma_v \overline{\rho} u - \kappa_v$, $c = 1 - \rho_J \mu_{c,v} u$, and

$$d = \frac{\gamma - b}{(\gamma - b)c + \mu_{c,v} a}\tau - \frac{2\mu_{c,v} a}{(\gamma c)^2 - (bc - \mu_{c,v} a)^2}\ln\left[1 - \frac{(\gamma + b)c - \mu_{c,v} a}{2\gamma c}(1 - e^{-\gamma\tau})\right].$$
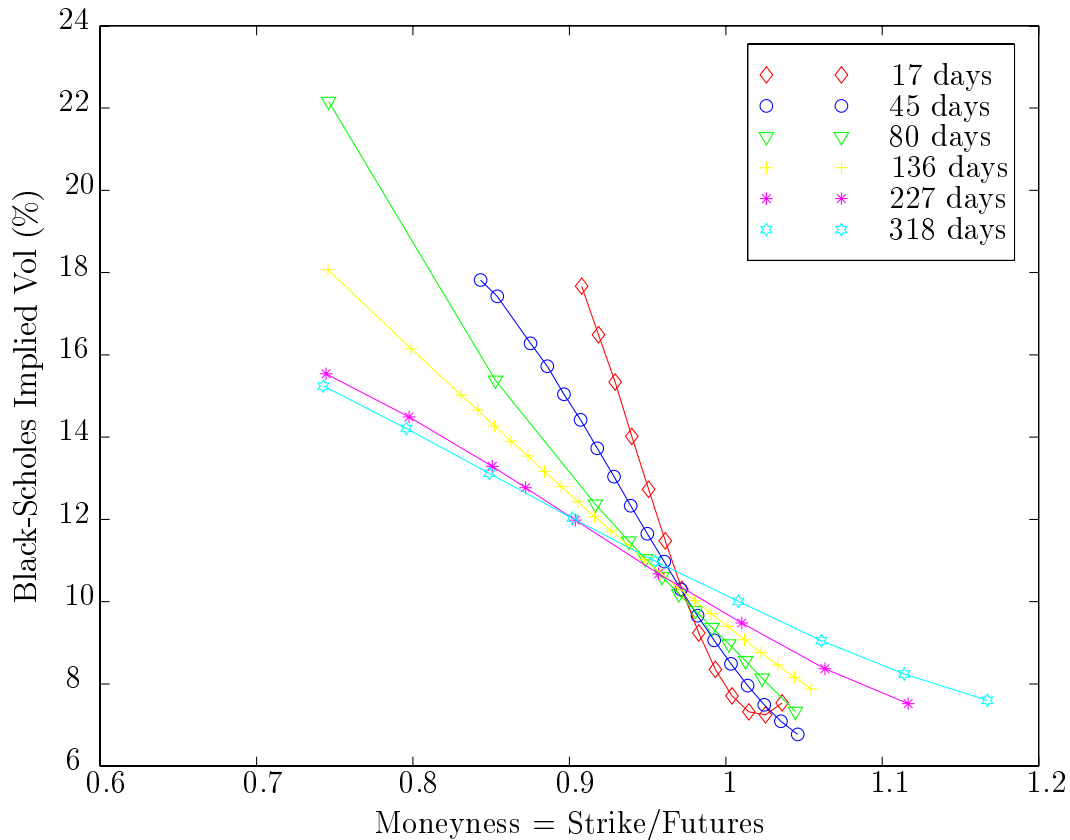
Figure 1: "Smile curves" implied by S&P 500 Index options of 6 different maturities. Option prices are obtained from market data of November 2, 1993.

4.2 Jump Impact on "Volatility Smiles"

As an illustration of the implications of jumps for the volatility smirk, we first select three special cases of the "double-jump" example just specified,

SV: Stochastic volatility model with no jumps, obtained by letting $\overline{\lambda} = 0$.

SVJ-Y: Stochastic volatility model with jumps in price only, obtained by letting $\lambda^y > 0$, and $\lambda^v = \lambda^c = 0$.

SVJJ: Stochastic volatility with simultaneous and correlated jumps in price and volatility, obtained by letting $\lambda^c > 0$ and $\lambda^y = \lambda^v = 0$.

In order to choose plausible values for the parameters governing these three special cases, we calibrated these three benchmark models to the actual "market-implied" smiles on November 2, 1993, plotted in Figure 1.[21] For each model, calibration was done by minimizing (by choice of the unrestricted parameters) the mean-squared pricing error (MSE), defined as the simple average of the squared differences between the observed and the modeled option prices across all strikes and maturities. The risk-free rate r is assumed to be 3.19%, and the dividend yield $\overline{\zeta}$ is assumed to be zero.
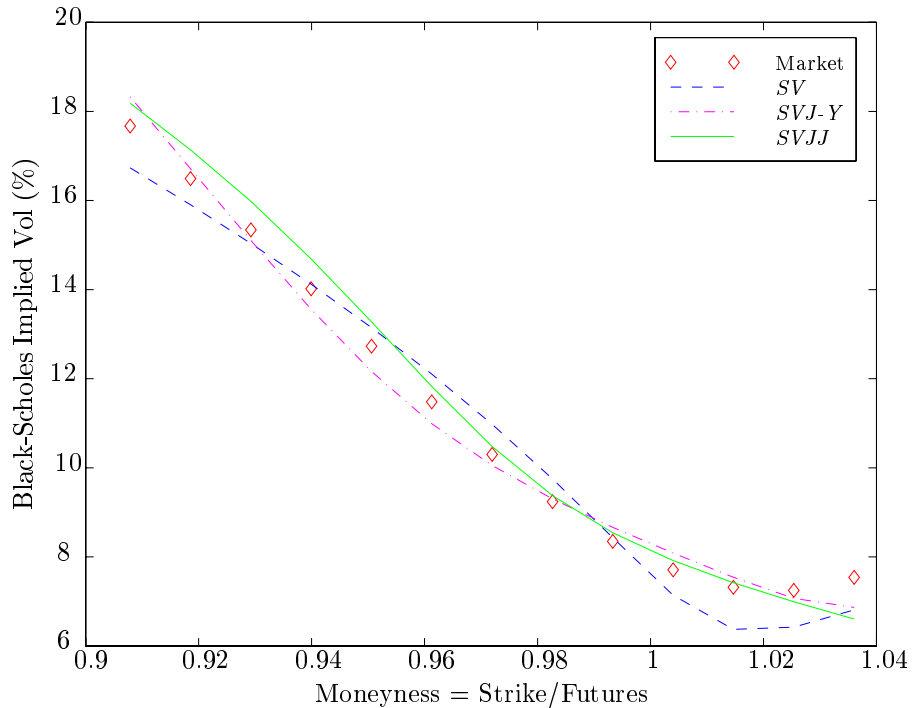

Figure 2: "Smile curves" implied by S&P 500 Index options with 17 days to expiration. Diamonds show observed Black-Scholes implied volatilities on November 2, 1993. *SV* is the Stochastic Volatility Model, *SVJ-Y* is the Stochastic Volatility Model with Jumps in Returns, and *SVJJ* is the Stochastic Volatility Model with Simultaneous and Correlated Jumps in Returns and Volatility. Model parameters were calibrated with options data of November 2, 1993.

---

[21]The options data are downloaded from the home page of Yacine Ait-Sahalia. There is a total of 87 options with maturities (times to exercise date) ranging from 17 days to 318 days, and strike prices ranging from 0.74 to 1.17 times the underlying futures price.

Table 1 displays the calibrated parameters of the models. Interestingly, for this particular day, we see that adding a jump in volatility to the *SVJ-Y* model, leading to the model *SVJJ* model, causes a substantial decline in the level of the parameter σ_v determining the volatility of the diffusion component of volatility. Thus, the volatility puzzle identified by Bates and Bakshi, Cao, and Chen, namely that the volatility of volatility in the diffusion component of V seems too high, is potentially explained by allowing for jumps in volatility. At the same time, the return jump variance σ_y^2 declines to approximately zero as we replace the *SVJ-Y* model with the *SVJJ* model. The instantaneous correlation among the jumps in return and volatility in the *SVJJ* model is $\mu_v \rho_J (\sigma_y^2 + \mu_v^2 \rho_J^2)^{-1/2}$. Thus, one consequence of the small σ_y^2 is that the jump sizes of Y and of V are nearly perfectly anti-correlated. This correlation reinforces the negative skew typically found in estimation of the *SV* model for these data,[22] as jumps down in return are associated with simultaneous jumps up in volatility.

In order to gain additional insight into the relative fit of the models to the option data used in our calibration, Figures 2 and 3 show the volatility smiles for the shortest (17-day) and longest (318-day) expiration options. For both maturities, there is a notable improvement of fit with the inclusion of jumps. Furthermore, the addition of a jump in volatility leads to a more pronounced smirk at both maturities and one that, based on the relative values of the *MSE* in Table 1, produces a better overall fit on this day.

Next, we go beyond this fitting exercise, and study how the introduction of a volatility jump component to the *SV* and *SVJ-Y* models might affect the "volatility smile," and how correlation between jumps in Y and V affects the "volatility smirk." We investigate the following three additional special cases:

1. The *SVJ-V* model: We extend the fitted *SV* model by letting $\lambda^v = 0.1$ and $\lambda^y = \lambda^c = 0$. We measure the degree of contribution of the jump component of volatility by the fraction $\lambda^v \mu_v^2 / (\sigma_v^2 V_0 + \lambda^v \mu_v^2)$ of the initial instantaneous variance of the volatility process V that is due to the jump component. By varying μ_v, the mean of the volatility jumps, three levels of this volatility "jumpiness" fraction are considered: 0, 15%, and 30%. For each case, the time-0 instantaneous drift, variance,

---

[22]In addition to the "calibration" results in the literature, see the time-series results of Chernov and Ghysels [1998] and Pan [1998]. For related work, see Poteshman [1998] and Benzoni [1998].

Table 1: Fitted Parameter Values for SV, SVJ-Y, and $SVJJ$ Models

	SV	SVJ-Y	$SVJJ$
$\overline{\rho}$	-0.70	-0.79	-0.82
$\overline{v}$	0.019	0.014	0.008
κ_v	6.21	3.99	3.46
σ_v	0.61	0.27	0.14
λ^c	0	0.11	0.47
$\overline{\mu}$	n/a	-0.12	-0.10
σ_y	n/a	0.15	0.0001
μ_v	n/a	0	0.05
ρ_J	n/a	n/a	-0.38
$\sqrt{V_0}$	10.1%	9.4%	8.7%
MSE	0.0124	0.0071	0.0041

The parameters are estimated by minimizing mean squared errors (MSE). A total of 87 options, observed on November 2, 1993, are used. $\sqrt{V_0}$ is the estimated value of stochastic volatility on the sample day. The risk-free rate is assumed to be fixed at $r = 3.19\%$, and the dividend yield at $\overline{\zeta} = 0$. From "risk neutrality," $\overline{\mu} = \theta(1, 0) - 1$.

and correlation are fixed at those implied by the fitted SV model by varying σ_v, $\overline{v}$, and $\overline{\rho}$.

2. The SVJ-Y-V model: We extend the fitted SVJ-Y model by letting $\lambda^v = \lambda^y$, $\lambda^c = 0$, and λ^y be fixed as given in Table 1. Again, the volatility "jumpiness" is measured by the fraction of the instantaneous variance of V that is due to the jump component. Three jumpiness levels, 0, 15%, and 30% are again considered. For each case, the instantaneous drift, variance, and correlation are matched to the fitted SVJ-Y model.

3. Finally, we modify the fitted $SVJJ$ model by varying the correlation between simultaneous jumps in Y and V. Five levels of correlation are considered: -1.0, -0.5, 0, 0.5, and 1.0. For each case, the means and variances of jumps in V and Y are calibrated to the fitted $SVJJ$ model.
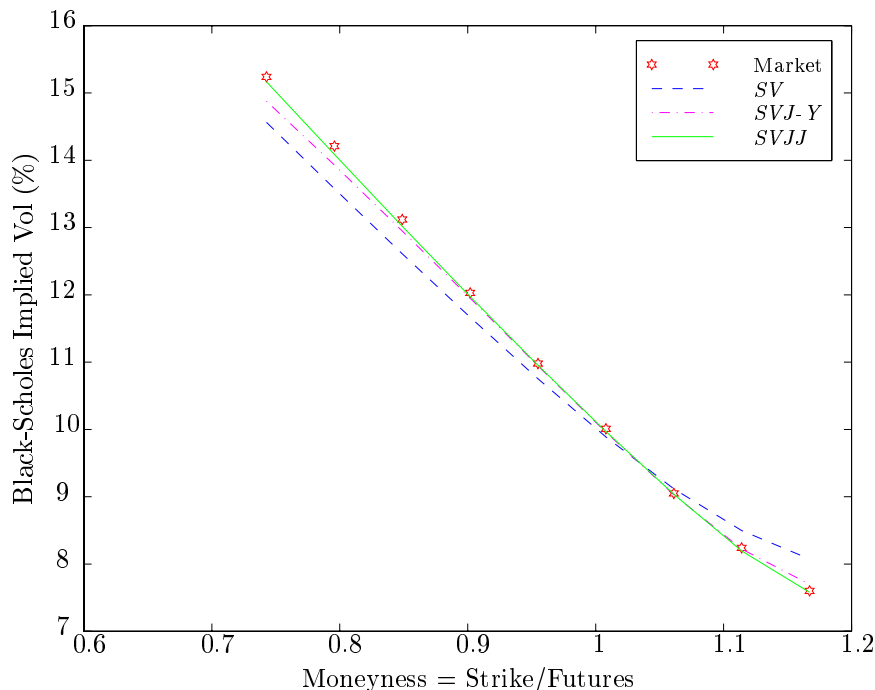
Figure 3: "Smile curves" implied by S&P 500 Index options with 318 days to expiration. "Stars" show observed implied volatility of November 2, 1993. *SV* is the Stochastic Volatility Model, *SVJ-Y* is the Stochastic Volatility Model with Jumps in Returns, and *SVJJ* is the Stochastic Volatility Model with Simultaneous and Correlated Jumps in Returns and Volatility. Model parameters were calibrated with options data of November 2, 1993.

The implied 30-day "volatility smiles" for the above three variations are plotted in Figures 4, 5, and 6.

The results for the *SVJ-V* model show that, for out-of-the-money (*OTM*) calls, the introduction of a jump in volatility lowers Black-Scholes implied volatilities. Bakshi, Cao, and Chen [1997] found that their *SVJ* model (jumps in returns, but not in volatility) systematically overpriced *OTM* calls. So our analysis suggests that adding jumps in volatility may attenuate the overpricing in the *SVJ* model, at least for options that are not too far out of the money. The addition of a jump in volatility actually exacerbates the over pricing for far-out-of-the-money calls.

Model *SVJ-Y-V* is one illustrative formulation of a model with jumps in both Y and V. Figure 5 shows that the addition of a jump in V to the *SVJ* model also attenuates the over-pricing of *OTM* calls. Whether our

parameterization of the jump distributions is enough to resolve the empirical puzzles relative to the SVJ model is an empirical issue that warrants further investigation.

Finally, Figure 6 shows that, in the presence of simultaneous jumps, the levels of implied volatilities for OTM calls depend on the sign and magnitudes of the correlation between the jump amplitudes. From our calibration of the $SVJJ$ model, the data suggest that ρ_J is negative (see Table 1). Thus, for this day, simultaneous jumps tend to reduce the Black-Scholes implied volatilities of OTM calls compared to the model with simultaneous jumps with uncorrelated amplitudes.
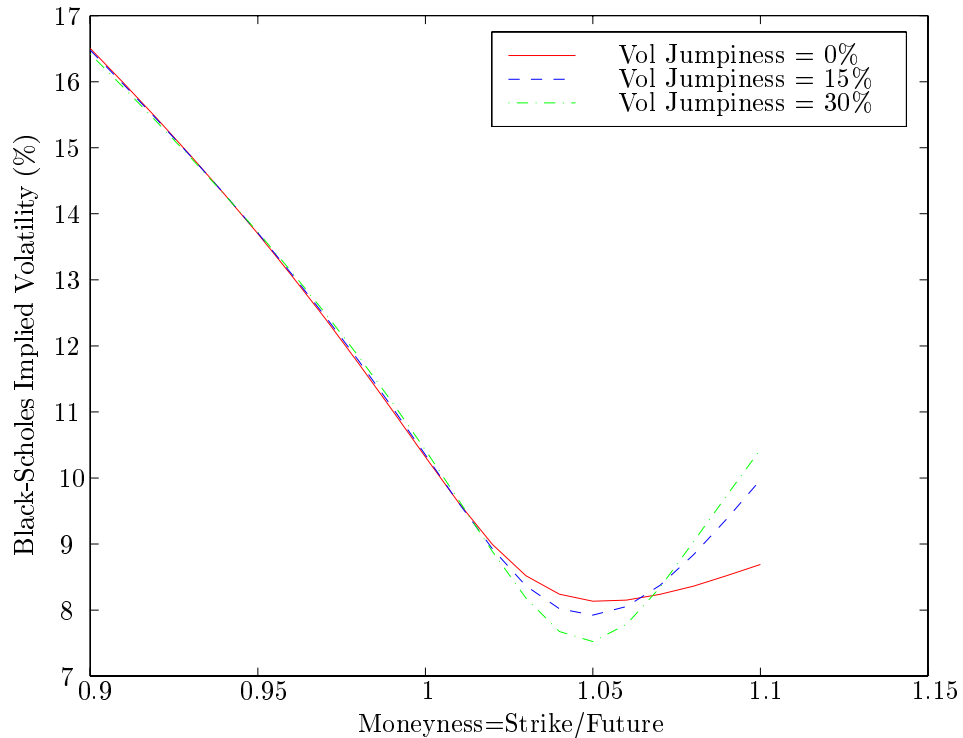

Figure 4: 30-day smile curve, varying volatility jumpiness, and no jumps in returns.

4.3 Multi-factor Volatility Specifications

Though our focus in this section has been on jump distributions, we are also interested in multi-factor models of the diffusion component of stochas-
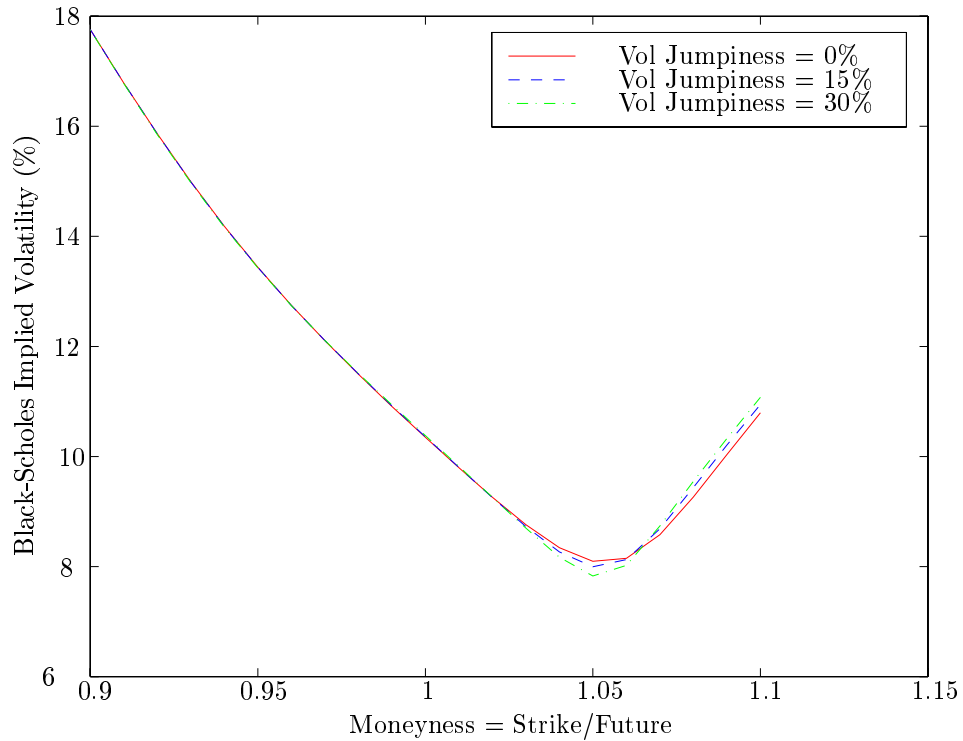
Figure 5: 30-day smile curve, varying volatility jumpiness. Independent arrivals of jumps in returns and volatility, with independent jump sizes.

tic volatility. Bates [1997] has emphasized the potential importance of more than one volatility factor for explaining the "term structure" of return volatilities, and included two, independent volatility factors in his model. Similarly, the empirical analysis in Gallant, Hsu, and Tauchen [1998] of a non-affine, 3-factor model of asset returns, with two of the three state coordinates dedicated to volatility behavior, suggests that more than one volatility factor improves the goodness of fit for $S\&P500$ returns.

Our transform analysis applies directly to any affine formulation of multi-factor stochastic volatility models, including Bates' model. Here, we also propose an examination of multi-factor volatility models in which there is a "long-term" stochastic trend component $\overline{V}_t$ in volatility. For example, we propose consideration of a three-factor model for $X = (Y, V, \overline{V})^\top$, given in
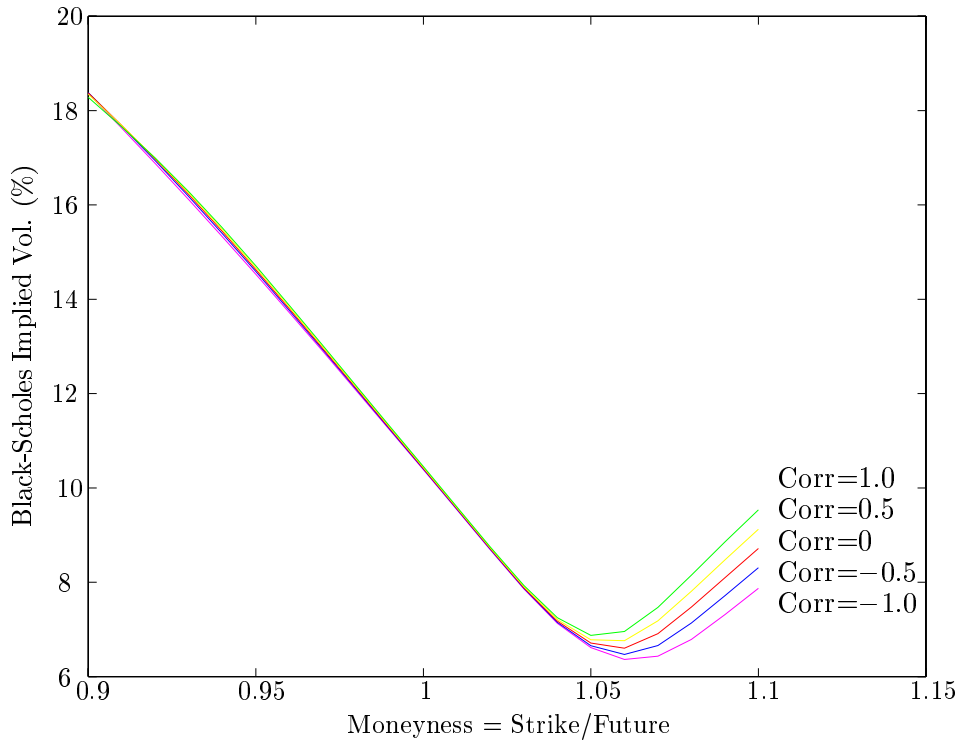
Figure 6: 30-day smile curve, varying the correlation between the sizes of simultaneous jumps in return and in volatility.

its risk-neutral form by

$$
d \begin{pmatrix} Y_t \\ V_t \\ \overline{V}_t \end{pmatrix} = \begin{pmatrix} r - \overline{\zeta} - \frac{1}{2}V_t \\ \kappa(\overline{V}_t - V_t) \\ \kappa_0(\overline{v} - \overline{V}_t) \end{pmatrix} dt + \begin{pmatrix} \sqrt{V_t} & 0 & 0 \\ \sigma\rho\sqrt{V_t} & \sigma\sqrt{1-\rho^2}\sqrt{V_t} & 0 \\ 0 & 0 & \sigma_0\sqrt{\overline{V}_t} \end{pmatrix} dW_t^Q,
$$

$$(4.6)$$

where W^Q is an $(\mathcal{F}_t)$-standard Brownian motion in $\mathbb{R}^3$ under Q.

A one-factor volatility model, such as the SV model, may well over-simplify the term structure of volatility. In particular, the SV model has an auto-correlation of returns (over successive periods of length Δ) of $\exp(-\kappa_v\Delta)$, which decreases exponentially with Δ. For the estimated values of κ typically found in practice, the autocorrelations of discretely sampled V decay too quickly relative to what is found in the data. Bollerslev and Mikkelsen [1996] argue, based on their analysis of LEAPs, for a "long memory" model

31

of volatility to capture this slow decay. The correlation of $(V_t, V_{t+\Delta})$ (with respect to the ergodic distribution of $(V, \overline{V})$) implied by model (4.6) is

$$\mathrm{corr}(V_t, V_{t+\Delta}) = e^{-\kappa\Delta} + \left(e^{-\kappa_0\Delta} - e^{-\kappa\Delta} \right) \frac{\kappa\sigma_0^2/(\kappa - \kappa_0)}{(\kappa + \kappa_0)\sigma^2/\kappa + \kappa\sigma_0^2/\kappa_0} .$$

By suitable choice of the parameter values, this correlation decays more slowly with Δ than the exponential rate in the one-factor model. In a different context, Gallant, Hsu, and Tauchen [1998] found that the correlogram for V was well approximated, at least over moderate horizons, by their two-factor volatility model, and we conjecture that the same is true of models like (4.6). In subsequent work, we plan to further investigate multi-factor volatility specifications.

Appendices

A   Technical Conditions and Arguments

This appendix contains technical results and conditions used in the body of the paper.

Lemma 1:   Under the assumptions of Proposition 1, J is a martingale.

Proof: Letting E_t denote $\mathcal{F}_t$-conditional expectation under P, for $0 \leq t \leq s \leq T$, we have

$$E_t \left(\sum_{t < \tau(i) \leq s} \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-} \right) \right) = E_t \left(\sum_{t < \tau(i) \leq s} E \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-} \;\middle|\; X_{\tau(i)-}, \tau(i) \right) \right)$$

$$= E_t \left(\sum_{t < \tau(i) \leq s} \Psi_{\tau(i)-} \left(\theta(b(\tau(i))) - 1 \right) \right)$$

$$= E_t \left(\sum_{t < \tau(i) \leq s} \int_{\tau(i-1)+}^{\tau(i)} \Psi_{u-} \left(\theta(b(u)) - 1 \right) dN_u \right)$$

$$= E_t \left(\int_t^T \Psi_{u-} \left(\theta(b(u)) - 1 \right) dN_u \right) .$$

Because $\{\Psi_{t-}\left(\theta(b(t)) - 1\right) : t \geq 0\}$ is an $(\mathcal{F}_t)$-predictable process, and the jump-counting process N has intensity $\{\lambda(X_t, t) : t \leq T\}$, integrability condition (i) implies that[23]

$$E_t\left(\int_t^s \Psi_{u-}\left(\theta(b(u)) - 1\right) dN_u\right) = E_t\left(\int_t^s \Psi_u\left(\theta(b(u)) - 1\right)\lambda(X_u, u)\, du\right).$$

Hence J is a martingale. ∎

Proposition 2 is proved as follows.

For $0 < \tau < \infty$, and a fixed $y \in \mathbb{R}$,

$$\frac{1}{2\pi}\int_{-\tau}^{\tau} \frac{e^{ivy}\psi^\chi(a - ivb, x, 0, T) - e^{-ivy}\psi^\chi(a + ivb, x, 0, T)}{iv}\, dv$$

$$= \frac{1}{2\pi}\int_{-\tau}^{\tau}\int_{\mathbb{R}} \frac{e^{-iv(z-y)} - e^{iv(z-y)}}{iv}\, dG_{a,b}(z; x, T, \chi)\, dv$$

$$= -\frac{1}{2\pi}\int_{\mathbb{R}}\int_{-\tau}^{\tau} \frac{e^{-iv(z-y)} - e^{iv(z-y)}}{iv}\, dv\, dG_{a,b}(z; x, T, \chi),$$

where Fubini is applicable[24] because

$$\lim_{y \to +\infty} G_{a,b}(y; x, T, \chi) = \psi^\chi(a, x, 0, T) < \infty,$$

given that χ is well-behaved at (a, T).

Next we note that, for $\tau > 0$,

$$\int_{-\tau}^{\tau} \frac{e^{-iv(z-y)} - e^{iv(z-y)}}{iv}\, dv = -\frac{\mathrm{sgn}(z - y)}{\pi}\int_{-\tau}^{\tau} \frac{\sin(v|z - y|)}{v}\, dv$$

is bounded simultaneously in z and τ, for each fixed y.[25] By the bounded convergence theorem,

$$\lim_{\tau \to \infty} \frac{1}{2\pi}\int_{-\tau}^{\tau} \frac{e^{ivy}\psi^\chi(a - ivb, x, 0, T) - e^{-ivy}\psi^\chi(a + ivb, x, 0, T)}{iv}\, dv$$

$$= -\int_{\mathbb{R}} \mathrm{sgn}(z - y)\, dG_{a,b}(z; x, T, \chi)$$

$$= -\psi^\chi(a, x, 0, T) + \left(G_{a,b}(y; x, T, \chi) + G_{a,b}(y-; x, T, \chi)\right),$$

---

[23]See, for example, page 27 of Brémaud [1981]. We are applying the result for the real and imaginary components of the integrand, separately.

[24]Here, we also use the fact that, for any $u, v \in \mathbb{R}$, $|e^{iv} - e^{iu}| \leq |v - u|$.

[25]We define $\mathrm{sgn}(x)$ to be 1 if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$.

where $G_{a,b}(y-;x,T,\chi) = \lim_{z\to y,z\le y} G_{a,b}(z;x,T,\chi)$. Using the integrability condition (2.11), by the dominated convergence theorem we have

$$G_{a,b}(y;x,T,\chi) = \frac{\psi^\chi(a,x,0,T)}{2}$$
$$+ \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{e^{ivy}\psi^\chi(a-ivb,x,0,T) - e^{-ivy}\psi^\chi(a+ivb,x,0,T)}{iv}\,dv.$$

Because $\psi^\chi(a-ivb,x,0,T)$ is the complex conjugate of $\psi^\chi(a+ivb,x,0,T)$, we have (2.12). ∎

Proposition 3, regarding the extended transform, relies on the following technical condition. For differentiability of θ at u, it is enough that θ is well defined and finite in a neighborhood of u.

Definition: (K,H,l,θ,ρ) is *"extended" well-behaved* at (v,u,T), if (2.5)-(2.6) are solved uniquely by β and α, if the jump transform θ is differentiable at $\beta(t)$ for all $t \le T$, if (2.15) is solved uniquely by B and A, and if the following integrability conditions (*i*)-(*iii*) are satisfied, where $\Phi_t = \Psi_t(A(t) + B(t) \cdot X_t)$:

(*i*) $E\left(\int_0^T |\tilde{\gamma}_t|\,dt\right) < \infty$, where $\tilde{\gamma}_t = \lambda(X_t)\left(\Phi_t\left(\theta(\beta(t)) - 1\right) + \Psi_t\nabla\theta(\beta(t))B(t)\right)$.

(*ii*) $E\left[\left(\int_0^T \tilde{\eta}_t \cdot \tilde{\eta}_t\,dt\right)^{1/2}\right] < \infty$, where $\tilde{\eta}_t = \Phi_t\left(\beta(t)^\top + B(t)^\top\right)\sigma(X_t)$.

(*iii*) $E\left(|\Phi_T|\right) < \infty$.

B Multiple Jump Types and Time Dependence

We can relax the jump behavior of X to accommodate time dependencies in the coefficients and different types of jumps, each arrving with a different stochastic intensity.

We re-define D to be a subset of $\mathbb{R}^n \times [0,\infty)$, and treat the state process X defined so that (X_t,t) is in D for all t. It is assumed that, for each

t, $\{x : (x,t) \in D\}$ contains an open subset of $\mathbb{R}^n$. The time-dependent generator is now defined by

$$\mathcal{D}f(x,t) = f_t(x,t) + f_x(x,t)\mu(x,t) + \frac{1}{2}\mathrm{tr}\left[f_{xx}(x,t)\sigma(x,t)\sigma(x,t)^\top\right]$$

$$+ \sum_i \lambda_i(x,t) \int_{\mathbb{R}^n} \left[f(x+z,t) - f(x,t)\right] d\nu_t^i(z), \qquad \text{(B.1)}$$

for sufficiently regular $f : D \to \mathbb{R}$. That is, jump type i has jump-conditional distribution ν_t^i at time t, depending only on t, and stochastic intensity $\{\lambda_i(X_t,t) : t \geq 0\}$, for $i \in \{1,\ldots,m\}$, where $\lambda_i : D \to \mathbb{R}_+$ is defined by $\lambda_i(x,t) = l_0^i(t) + l_1^i(t) \cdot x$, for functions $(l_0^1, l_1^1), \ldots, (l_0^m, l_1^m)$ on $[0,\infty)$ into $\mathbb{R} \times \mathbb{R}^n$. The jump transforms $\theta = (\theta^1,\ldots,\theta^m)$ are defined by $\theta^i(c,t) = \int_{\mathbb{R}^n} \exp(c \cdot z) \, d\nu_t^i(z)$, $c \in \mathbb{C}^n$. We take

$$\mu(x,t) = K_0(t) + K_1(t)\,x$$

$$\sigma(x,t)\sigma(x,t)^\top = H_0(t) + \sum_{k=1}^n H_1^{(k)}(t)\,x_k,$$

where for each $t \geq 0$, $K_0(t)$ is $n \times 1$, $K_1(t)$ is $n \times n$, $H_0(t)$ is $n \times n$ and symmetric, and $H_1(t)$ is a tensor[26] of dimension $n \times n \times n$, with symmetric $H^{(k)}(t)$ (for $k = 1,\ldots,n$). The time-dependent coefficients $K = (K_0, K_1)$, $H = (H_0, H_1)$, and $l = (l_0, l_1)$ are assumed to be bounded continuous functions on $[0,\infty)$.

In this more general setting, Propositions 1, 2, and 3 apply after introducing these time-dependent coefficients into (2.5) and (2.6), and replacing the last terms in the right-hand sides of these ODEs with $\sum_{i=1}^m l_1^i(t)\left(\theta^i(c,t) - 1\right)$ and $\sum_{i=1}^m l_0^i(t)\left(\theta^i(c,t) - 1\right)$, respectively.

We can further extend to the case of an infinite number of jump types by allowing for a general Lévy jump measure that is affine in the state vector. (See Theorem 42, page 32, of Protter [1990].)

C   Change of Measure

This appendix provides the impact of a change of measure defined by a density process or a state-price-density process that is of the exponential-affine form in an affine jump-diffusion state process X.

---

[26]Let H be an $n \times n \times n$ tensor, fix its third index to k, the tensor is reduced to an $n \times n$ matrix $H^{(k)}$ with elements, $H_{ij}^{(k)} = H(i,j,k)$.

Fixing $T > 0$, suppose, under the measure P, that a given characteristic $\chi = (K, H, l, \theta, \rho)$ is well-behaved at (b, T) for some $b \in \mathbb{R}^n$. Let

$$\xi_t = \exp\left(-\int_0^t R(X_s, s)\,ds\right) \exp\left(\alpha(t, T, b) + \beta(t, T, b) \cdot X_t\right). \qquad \text{(C.1)}$$

Under the conditions of Proposition 1, ξ is a positive martingale. We may then define an equivalent probability measure Q by $\frac{dQ}{dP} = \xi_T/\xi_0$. In this section, we show how to compute the transform of X after a change of measure with density process ξ. Many other densities could be considered, as in Bühlmann, Delbaen, Embrechts, and Shiryaev [1996]. We have chosen this density as it preserves the affine behavior of X under the change of measure, and because it arises naturally when re-normalizing prices by the price of a zero-coupon bond maturing on a particular date. (This is sometimes called "forward measure.") A more general way to choose an equivalent measure Q^* that would suffice for our purposes would have

$$\frac{dQ^*}{dP} = \frac{1}{k} \sum_{i=1}^{m} \exp\left(-\int_0^T R_i(X_s, s)\,ds\right) \exp\left(b_i \cdot X(t_i)\right), \qquad \text{(C.2)}$$

where, for each $i \in \{1, \ldots, m\}$, $R_i(x, t)$ is affine in x, t_i is a fixed time, and $b_i \in \mathbb{R}^n$; and where $k \in (0, \infty)$ is a normalizing scalar chosen so that $E(\frac{dQ^*}{dP}) = 1$.

Proposition 5 (Transform under Change of Measure):
Let $\chi(Q) = \left(K^Q, H^Q, l^Q, \theta^Q\right)$ be defined by

$$K_0^Q(t) = K_0(t) + H_0(t)\beta(t, T, b), \quad K_1^Q(t) = K_1(t) + H_1(t)\beta(t, T, b), \quad \text{(C.3)}$$
$$l_0^Q(t) = l_0(t)\theta(\beta(t, T, b), t), \quad l_1^Q(t) = l_1(t)\theta(\beta(t, T, b), t), \qquad \text{(C.4)}$$
$$\theta^Q(c, t) = \theta(c + \beta(t, T, b), t)/\theta(\beta(t, T, b), t), \quad H^Q(t) = H(t), \qquad \text{(C.5)}$$

where $H_1(t)b(t)$ denotes the $n \times n$ matrix with k-th column $H_1^{(k)}(t)b(t)$. Let $R^Q(x, t) = \rho_0^Q(t) + \rho_1^Q(t) \cdot x$, for some bounded measurable $\rho_0^Q : [0, \infty) \to \mathbb{R}$ and $\rho_1^Q : [0, \infty) \to \mathbb{R}^n$. Let $\rho^Q = \left(\rho_0^Q, \rho_1^Q\right)$ be such that $\chi(Q)$ is well-behaved at some (u, T). Then, for $t \leq T$,

$$E^Q\left(\exp\left(-\int_0^T R^Q(X_s, s)\,ds\right) \exp\left(u \cdot X_T\right) \,\Big|\, \mathcal{F}_t\right) = \psi^{\chi(Q)}(u, X_t, t, T),$$
$$\text{(C.6)}$$

where $\psi^{\chi(Q)}$ is defined by (2.4).

Proof: Let

$$W_t^Q = W_t - \int_0^t \sigma(X_s, s)^\top \beta(s, T, b)\, ds\,, \quad t \geq 0. \tag{C.7}$$

Lemma 2, below, shows that ξW^Q is a P-local martingale. It follows that W^Q is a Q-local martingale. Because $\int_0^t \sigma^\top(X_s, s)\beta(s, T, b)\, ds$ is a continuous finite-variation process, $[W_i^Q, W_j^Q]_t = [W_i^P, W_j^P]_t = \delta(i, j)\, t$, where $\delta(\cdot)$ is the Kronecker delta. By Lévy's Theorem, W^Q is a standard Brownian motion in $\mathbb{R}^n$ under Q.

Next, we let

$$M_t^Q = N_t - \int_0^t \theta(\beta(s, T, b))\lambda(X_s, s)\, ds\,, \quad t \geq 0. \tag{C.8}$$

Lemma 3, below, shows that ξM^Q is a P-local martingale. It follows that M^Q is a Q-local martingale. By the martingale characterization of intensity,[27] we conclude that, under Q, N is a counting process with the intensity $\left\{\lambda^Q(X_t, t) : t \geq 0\right\}$ defined by $\lambda^Q(x, t) = l_0^Q(t) + l_1^Q(t) \cdot x$.

Using the fact that, under Q, W^Q is a standard Brownian and the jump counting process N has intensity $\left\{\lambda^Q(X_t, t) : t \geq 0\right\}$, we may mimic the proof of Proposition 1, and obtain (C.6), replacing in the proof of Lemma 1 $E_t \left(\sum_{t < \tau(i) \leq T} \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-}\right) \right)$ with

$$E_t^Q \left(\sum_{t < \tau(i) \leq T} \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-}\right) \right) = \frac{1}{\xi_t} E_t \left(\sum_{t < \tau(i) \leq T} \xi_{\tau(i)} \left(\Psi_{\tau(i)} - \Psi_{\tau(i)-}\right) \right).$$

This completes the proof. ∎

Lemma 2: Under the assumptions of Proposition 1, ξW^Q is a P-local martingale.

---

[27]See, for example, page 28 of Brémaud [1981].

Proof: By Ito's Formula, with $0 \leq s \leq t \leq T$,

$$\xi_t W_t^Q = \xi_s W_s^Q + \int_s^t \xi_{u-} \, dW_u^Q + \int_s^t W_{u-}^Q \, d\xi_u$$

$$+ \sum_{s < u \leq t} (\xi_u - \xi_{u-}) \left(W_t^Q - W_{t-}^Q \right) + \int_s^t d[\xi, W^Q]_u^c$$

$$= \xi_s W_s^Q + \int_s^t \xi_{u-} \left(dW_u - \sigma^\top(X_u, u) b(u) \, du \right)$$

$$+ \int_s^t W_u^Q \, d\xi_u + \int_s^t \xi_u \sigma^\top(X_u, u) b(u) \, du$$

$$= \xi_s W_s^Q + \int_s^t \xi_{u-} \, dW_u + \int_s^t W_u^Q \, d\xi_u \,,$$

where $[\xi, W^Q]^c$ denotes the continuous part of the "square-brackets" process $[\xi, W^Q]$. As W and ξ are P-martingales, both $\left\{ \int_0^t \xi_{u-} \, dW_u : t \geq 0 \right\}$ and $\left\{ \int_0^t W_u^Q \, d\xi_u : t \geq 0 \right\}$ are P-local martingales. Hence, ξW^Q is a P-local martingale. ∎

Lemma 3: Under the assumptions of Proposition 1, ξM^Q is a P-local martingale.

Proof: By Ito's Formula, with $0 \leq s \leq t \leq T$,

$$\xi_t M_t^Q = \xi_s M_s^Q + \int_s^t \xi_{u-} \, dM_u^Q + \int_s^t M_{u-}^Q \, d\xi_u + \sum_{s < u \leq t} (\xi_u - \xi_{u-})(N_u - N_{u-})$$

$$= \xi_s M_s^Q + \int_s^t \xi_{u-} \, dM_u + \int_s^t M_{u-}^Q \, d\xi_u + J^\xi,$$

where $M_t = N_t - \int_0^t \lambda(X_s, s) \, ds$, and where

$$J^\xi = \sum_{s < u \leq t} (\xi_u - \xi_{u-}) - \int_s^t \xi_u (\theta(\beta(u, T, b), u) - 1) \lambda(X_u, u) \, du.$$

As M and ξ are P-martingales, $\left\{ \int_0^t \xi_{u-} \, dM_u : t \geq 0 \right\}$ and $\left\{ \int_0^t M_{u-}^Q \, d\xi_u : t \geq 0 \right\}$ are P-local martingales. By a proof similar to that of Lemma 1, and using

38

the Integration Theorem (γ) in Brémaud [1981], we can show that J^ξ is a P-local martingale. ∎

For the remainder of this appendix, we denote Q by $Q(b)$, emphasizing the role of b in defining the change of probability measure given by (C.1). We let $\chi(b) = \left(K^{Q(b)}, H^{Q(b)}, l^{Q(b)}, \theta^{Q(b)}, \rho\right)$ denote the associated characteristic. The previous result shows in effect that, under $Q(b)$, the state vector X is still an affine jump-diffusion whose characteristics can be computed in terms of the characteristics of X under the measure P. This result provides us with an alternative approach to option pricing. We suppose that $Q(0)$ is an equivalent martingale measure. The price $\Gamma\left(X_0, a, d, c, T\right)$ of an option paying $\left(e^{a+d \cdot X_T} - c\right)^+$ at T is given by

$$
\begin{aligned}
\Gamma\left(X_0, a, d, c, T\right) &= E^{Q(0)}\left(\exp\left(-\int_0^T R(X_s, s)\, ds\right)\left(e^{a+d \cdot X_T} - c\right)^+\right) \\
&= e^a E^{Q(0)}\left(\exp\left(-\int_0^T R(X_s, s)\, ds\right) e^{d \cdot X_T} \mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right) \\
&\quad - c\, E^{Q(0)}\left(\exp\left(-\int_0^T R(X_s, s)\, ds\right) \mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right).
\end{aligned}
$$

Provided the characteristic (K, H, l, θ, ρ) is well-behaved at (d, T) and $(0, T)$, we may introduce the equivalent probability measure $Q(d)$, and write

$$
\begin{aligned}
\Gamma\left(X_0, a, d, c, T\right) &= e^a \exp\left(\alpha(0, T, d) + \beta(0, T, d) \cdot X_0\right) E^{Q(d)}\left(\mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right) \\
&\quad - c \exp\left(\alpha(0, T, 0) + \beta(0, T, 0) \cdot X_0\right) E^{Q(0)}\left(\mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right).
\end{aligned}
$$

Let $\chi(1) = \left(K^{Q(d)}, H^{Q(d)}, l^{Q(d)}, \theta^{Q(d)}, 0\right)$ and $\chi(0) = \left(K^{Q(0)}, H^{Q(0)}, l^{Q(0)}, \theta^{Q(0)}, 0\right)$ be defined by (C.3)-(C.5) for $b = d$ and $b = 0$. We suppose that $\chi(1)$ and $\chi(0)$ are well behaved at (ivd, T) for any $v \in \mathbb{R}$. Then

$$
E^{Q(d)}\left(\mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left[\psi^{\chi(1)}(ivd, x, 0, T) e^{-iv(\ln(c)-a)}\right]}{v}\, dv,
$$

$$
E^{Q(0)}\left(\mathbf{1}_{d \cdot X_T \geq \ln(c)-a}\right) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\operatorname{Im}\left[\psi^{\chi(0)}(ivd, x, 0, T) e^{-iv(\ln(c)-a)}\right]}{v}\, dv,
$$

provided $\int_{\mathbb{R}} \left|\psi^{\chi(1)}(ivd, X_0, 0, T)\right| dv < \infty$ and $\int_{\mathbb{R}} \left|\psi^{\chi(0)}(ivd, X_0, 0, T)\right| dv < \infty$. These quantities may now be substituted into the previous relation in order to obtain the option price.

D Cap Pricing

A cap is a loan with face value, say 1, at a variable interest rate that is capped at some level $\bar{r}$. At time t, let τ, 2τ, ..., $n\tau$ be the fixed dates for future interest payments. At each fixed date $k\tau$, the $\bar{r}$-capped interest payment, or "caplet," is given by $\tau \left(\mathcal{R}((k-1)\tau, k\tau) - \bar{r} \right)^+$, where $\mathcal{R}((k-1)\tau, k\tau)$ is the τ-year floating interest rate at time $(k-1)\tau$, defined by

$$\frac{1}{1 + \tau \mathcal{R}\left((k-1)\tau, k\tau\right)} = \Lambda\left((k-1)\tau, k\tau\right).$$

The market value at time 0 of the caplet paying at date $k\tau$ can be expressed as

$$\mathrm{Caplet}(k) = E^Q \left[\exp\left(-\int_0^{k\tau} R(X_u, u)\, du \right) \tau \left(\mathcal{R}((k-1)\tau, k\tau) - \bar{r} \right)^+ \right]$$

$$= (1 + \tau\bar{r})\, E^Q \left[\exp\left(-\int_0^{(k-1)\tau} R(X_u, u)\, du \right) \left(\frac{1}{1 + \tau\bar{r}} - \Lambda((k-1)\tau, k\tau) \right)^+ \right].$$

Hence, the pricing of the k-th caplet is equivalent to the pricing of an in-$(k-1)\tau$-for-τ put struck at $1/(1+\tau\bar{r})$, which can be readily obtained by using Proposition 3 and put-call parity as $\mathrm{Caplet}(k) = (1 + \tau\bar{r})\overline{C}(k)$, where

$$\overline{C}(k) = \Gamma\left(X_0, \bar{\alpha}, \bar{\beta}, \frac{1}{1 + \tau\bar{r}}, (k-1)\tau \right) - \Lambda(0, k\tau) + \frac{\Lambda(0, (k-1)\tau)}{1 + \tau\bar{r}},$$

where $\Gamma(X_0, a, d, c, T)$ is the price of a claim to $(e^{a + d \cdot X(T)} - c)^+$ paid at T, and where $\bar{\alpha} = \alpha((k-1)\tau, k\tau, 0)$ and $\bar{\beta} = \beta((k-1)\tau, k\tau, 0)$.

References

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical Performance of Alternative Option Pricing Models. *Journal of Finance 52*, 2003–49.

Bakshi, G. and D. Madan (1999). Spanning and Derivative-Security Valuation. *Journal of Financial Economics, forthcoming*.

Bates, D. (1996). Jump and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *The Review of Financial Studies 9*, 69–107.

Bates, D. (1997). Post-'87 Crash fears in S&P 500 Futures Options. Working Paper 5894, National Bureau of Economic Research.

Benzoni, L. (1998). Pricing Options under Stochastic Volatility: An Econometric Analysis. Working Paper, J.L. Kellogg Graduate School of Management, Northwestern University.

Bollerslev, T. and H. O. Mikkelsen (1996). Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics 73*, 151–184.

Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics*. New York: Springer-Verlag.

Bühlmann, H., F. Delbaen, P. Embrechts, and A. N. Shiryaev (1996). No-arbitrage, Change of Measure and Conditional Esscher Transforms. *CWI Quarterly 9*, 291–317.

Chacko, G. and S. Das (1998). Pricing Average Interest Rate Options: A General Approach. Working Paper, Harvard Business School, Harvard University.

Chacko, G. and L. Viceira (1999). Dynamic Consumption and Portfolio Choice with Stochastic Volatility. Working Paper, School of Business, Harvard University.

Chen, R. and L. Scott (1993). Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *The Journal of Fixed Income 3*, 14–31.

Chen, R. and L. Scott (1995). Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure. *Journal of Derivatives 3*, 53–72.

Chernov, M. and E. Ghysels (1998). What Data Should be Used to Price Options? Working Paper, Department of Finance, Pennsylvania State Unniversity.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53* (2), 385–407.

Dai, Q. and K. Singleton (1999). Specification Analysis of Affine Term Structure Models. *Journal of Finance, forthcoming*.

Das, S. (1998). Poisson-Gaussian Processes and The Bond Markets. Working Paper, Harvard Business School, Harvard University.

Delbaen, F. and W. Schachermayer (1994). A General Version of the Fundamental Theorem of Asset Pricing. *Math. Ann. 300*, 463–520.

Duffie, D. and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance 6* (4), 379–406.

Duffie, D. and K. Singleton (1999). Modeling Term Structures of Defaultable Bonds. *Review of Financial Studies, forthcoming*.

Ethier, S. and T. Kurtz (1986). *Markov Processes, Characterization and Convergence*. New York: John Wiley & Sons.

Gallant, R., C.-T. Hsu, and G. Tauchen (1998). Calibrating Volatility Diffusions and Extracting Integrated Volatility. Working Paper, Department of Economics, University of North Carolina.

Gil-Pelaez, J. (1951). Note on the Inversion Theorem. *Biometrika 38*, 481–482.

Harrison, M. and D. Kreps (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory 20*, 381–408.

Heston, S. (1993). A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies 6*, 327–343.

Jamshidian, F. (1989). An Exact Bond Option Formula. *Journal of Finance 44*, 205–209.

Jarrow, R., D. Lando, and S. Turnbull (1997). A Markov Model for the Term Structure of Credit Spreads. *Journal of Finance 10*, 481–523.

Jiang, G. and J. Knight (1999). Efficient Estimation of the Continuous Time Stochastic Volatility Model Via the Empirical Characteristic Function. Working Paper, University of Western Ontario.

Lando, D. (1998). On Cox Processes and Credit Risky Securities. *Review of Derivatives Research 2*, 99–120.

Liu, J. (1997). Generalized Method of Moments Estimation of Affine Diffusion Processes. Working Paper, Graduate School of Business, Stanford Unversity.

Liu, J. and J. Pan (1997). Asymptotic Maximum Likelihood Estimation of Affine Models. Working Paper, Graduate School of Business, Stanford Unversity.

Pan, J. (1998). "Integrated" Time-Series Analysis of Spot and Options Prices. Working Paper, Graduate School of Business, Stanford University.

Piazzesi, M. (1998). Monetary Policy and Macroeconomic Variables in a Model of the Term Structure of Interest Rates. Working Paper, Stanford University.

Poteshman, A. M. (1998). Estimating a General Stochastic Variance Model from Options Prices. Working Paper, Graduate School of Business, University of Chicago.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. New York: Springer-Verlag.

Scott, L. (1996). The Valuation of Interest Rate Derivatives in a Multi-Factor Cox-Ingersoll-Ross Model that Matches the Initial Term Structure. Working Paper, University of Georgia.

Scott, L. (1997). Pricing Stock Options in a Jump-diffusion Model with Stochastic Volatility and Interest Rates: Application of Fourier Inversion Methods. *Mathematical Finance 7*, 345–358.

Singleton, K. (1999). Estimation of Affine Diffusion Models Based on the Empirical Characteristic Function. Working Paper, Graduate School of Business, Stanford University.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics 5*, 177–188.

Williams, D. (1991). *Probability with Martingales*. Cambridge University Press.