Department of Economics

Stochastic Volatility and Jumps Driven by Continuous Time Markov Chains

Kyriakos M. Chourdakis

Working Paper No. 430

December 2000

ISSN 1473-0278

Queen Mary
University of London

Stochastic Volatility and Jumps Driven by Continuous Time Markov Chains

Kyriakos M. Chourdakis*
Queen Mary, University of London

December 2000

Abstract

This paper considers a model where there is a single state variable that drives the state of the world and therefore the asset price behavior. This variable evolves according to a multi-state continuous time Markov chain, as the continuous time counterpart of Hamilton (1989) model. It derives the moment generating function of the asset log-price difference under very general assumptions about its stochastic process, incorporating volatility and jumps that can follow virtually any distribution, both of them being driven by the same state variable. For an illustration, the extreme value distribution is used as the jump distribution. The paper shows how GMM and conditional ML estimators can be constructed, generalizing Hamilton's filter for the continuous time case. The risk neutral process is constructed and contingent claim prices under this specification are derived, in the lines of Bakshi and Madan (2000). Finally, an empirical example is set up, to illustrate the potential benefits of the model.

Keywords: Option Pricing, Markov Chain, Moment Generating Function.

JEL Classification: C51, G12, G13.

**Address:* Department of Economics, Queen Mary, University of London, Mile End, E1 4NS, U.K. *e-mail:* k.chourdakis@qmw.ac.uk *URL address:* <http://www.qmw.ac.uk/~te9001>

1 Introduction

There is increasing evidence that models augmented with some kind of regime switching behavior will capture in a more realistic and intuitive way the time-series properties of many economic variables. The seminal paper of Hamilton (1989)¹ uses a simple switching regime model and successfully identifies the business cycles. Since then, virtually all economic and econometric specifications have been subject to some extension involving regime switches, especially the ones that attempt to capture sudden but cyclical behavior changes. For example Schwert (1990) notes the higher volatility that stock markets exhibit during recessions.² Most of the regime switching models are given in a discrete time setting, mainly because the majority of filtering and smoothing procedures is available for the discrete case only.

On the other hand, continuous time models have been widely used in order to price derivative instruments and construct hedging strategies.³ Since its introduction, the Black-Scholes pricing formula in Black and Scholes (1973) has been augmented in order to capture the stylized facts of the stock and

¹Hamilton (1989) gives a ML procedure which enables one to estimate regime switching models and make inference about the state of the world when this is not directly observed. In that sense it can be thought of as a discrete state version of the Kalman-Bucy filter.

²Hamilton and Susmel (1994) and Susmel (1999) consider ARCH type switching regressions to model the volatility which depends on past news and the state of the economy; where Gray (1996) uses a regime switching specification to explain the interest rate behavior. A time-varying transition probability model is discussed in Filardo (1994). In a general equilibrium framework Cecchetti, Lam, and Mark (1990) utilize Markov chains to examine the autocorrelation structure of the stock returns; Turner, Startz, and Nelson (1989) consider a learning mechanism that agents incorporate, in order to filter and identify the unobserved state of the world; while Chourdakis and Tzavalis (2000a) retrieve the implied risk aversion coefficients when discrete shifts in the production are incorporated.

³Examples of models in a discrete time setting include Chourdakis and Tzavalis (2000b) who use a RNVR in the spirit of Brennan (1979), Duan (1995), and Heston and Nandi (2000) who use recursive procedures for the characteristic function to arrive to the option pricing formula.

option price behavior. Merton (1976) derives a solution to the option pricing problem when jumps are allowed in the stock price diffusion, while Hull and White (1987) and Heston (1993) consider a model where the [unobserved] volatility evolves according to a mean reverting diffusion. Other researchers have combined those two variants, constructing models that incorporates two or three unobserved random variables, namely the Poisson counter that determines the occurrence of jumps, a random variable representing the magnitude of the jumps, and the diffusion responsible for volatility changes [see for example Bakshi, Cao, and Chen (1997)]. Lo and Wang (1995) consider a random coefficient model, in order to examine the effect of the asset price predictability on derivative pricing. A model with continuous time regime shifts has been proposed by Naik (1993), where closed form solutions are available when the state variable takes two distinct values; the option pricing formulas are derived using the results of Pedler (1971) on binary Markov chains for this particular case. Bakshi and Madan (2000) describe how options with trigonometric payoffs span the contingent claim space and give explicit solutions when the Fourier transform of the risk neutral density is available in closed form.

This paper considers a model where there is a unique state variable that drives both the state of the world and the jumps. This variable evolves according to a multi-state continuous time Markov chain making the model a continuous time counterpart of Hamilton's discrete time model. The state of the world will determine the parameters of the diffusion process, while changes of this state will result in discontinuities taking the form of jumps. This model can be seen as an augmentation of Naik's work in two directions: (i) It produces closed form solutions for the option pricing without restricting the state variable in being a binary one, a number of values are allowed adding flexibility but at the cost of higher complexity,⁴ and (ii) It allows the jump

⁴Naik (1993) and Chourdakis and Tzavalis (2000b) give closed form solutions when the Markov chain is a binary proces. They use the results of Pedler (1971) on binary Markov chains to express the value of a contingent claim as a weighted

component to be of a stochastic magnitude.⁵ The jump model of Merton (1976) is a special case of the model that this paper proposes, while the stochastic volatility models of Hull and White (1987) and Heston (1993) can be thought as limiting cases of the regime dependent model, as the number of states increases and a particular structure is imposed on their alternations, thus making the unobserved state process to converge to a mean reverting diffusion.

The paper is organized as follows: Section 2 introduces the model and derives the moment generating [and the characteristic] function of the asset return under very general assumptions about its stochastic process, incorporating mean reversion and jumps that can follow virtually any distribution. Some useful result of matrix calculus are presented, which allow one to compute all conditional and unconditional moments of the asset return, for any time horizon.⁶ Several ways of estimating the parameters of a continuous time Markov chain driven process are discussed in section 3.⁷ The conditional moments can be used in order to obtain the continuous time parameters of such models, using a GMM procedure. Alternatively, the paper discusses how the average of conditional [on the sojourn time] prices, in the spirit of Merton (1976). In this paper the dimension of the Markov chain is not restricted to be equal to two, and therefore the results of Pedler (1971) are not applicable.

⁵Naik (1993) introduces a deterministic jump when the state of the Markov chain changes, in order to introduce a correlation pattern between asset price and volatility changes.

⁶The moments of the discrete time regime switching models are discussed in Timmermann (2000) for the unconditional [one-period] and in Chourdakis and Tzavalis (2000b) for the conditional [multi-period] case.

⁷Naik (1993) and David and Veronesi (1998) among others implicitly argue that by using high frequency data one can adequately approximate the continuous time parameters by utilizing Hamilton's filter. On the other hand, the essence of diffusions that are driven by Markov chains is to capture long term cyclical phenomena, such as the business cycles or long term exchange rate fluctuations. In such a context, the frequency of the data sample might not suffice for one to assume a good approximation of the continuous time parameters from their corresponding discrete time estimators.

likelihood function can be derived by inverting the characteristic function, allowing one to utilize ML estimation methods. In that case, it is shown how inference about the unobserved state can be made in the lines of Hamilton (1989). Section 4 derives the corresponding risk neutral process of the asset price and discusses its features. It consequently gives the European call option price when the underlying asset price follows such a specification. In section 5 a small empirical exercise illustrates the potential of the proposed model to fit stock return series and to explain some stylized facts of the option pricing literature. Calibration shows that the effect of regime changes on European call prices is relatively small, allowing one to ignore the regime risk for empirical purposes. Section 6 concludes.

2 The model

This section lays down the assumptions and the stochastic differential equations associated with the diffusion that is considered in this paper. The Laplace and Fourier transforms of the density are then retrieved. It is shown how the conditional moments of the process can be computed, giving some useful results on matrix function differentiation. Some estimation procedures are discussed, and derivative prices are given in closed form.

2.1 A diffusion that depends on a discretely displaced stochastic process

Denote the underlying asset price with $S(t)$, and a state process $\mathbf{x}(t)$, for $t \geq 0$. The setup of the option pricing model is based on the following

Assumption Consider the stock returns to follow the diffusion

$$\begin{aligned} \frac{dS(t)}{S(t)} = & \left\{ \mu(\mathbf{x}(t)) - \mathbf{x}^T(t) [\mathbf{Q} \odot \mathbf{M}_\eta] \boldsymbol{\iota} \right\} dt \\ & + \sigma(\mathbf{x}(t)) d\varepsilon(t) + \tilde{\eta}(d\mathbf{x}(t)), \quad (1) \end{aligned}$$

where:

- $\varepsilon(t)$ is a standardized Wiener process;
- the instantaneous return has a drift and a volatility parameter which depend explicitly on the stochastic process $\mathbf{x}(t)$. In addition, jumps occur which are dependent on the behavior of $\mathbf{x}(t)$. The process $\mathbf{x}(t)$ follows a continuous time Markov chain [CTMC], with state space the set $\mathcal{S} = \{\mathbf{e}_1, \dots, \mathbf{e}_N\}$, which is the normal basis⁸ of $\mathbf{R}^{N \times N}$;
- the jumps are given by the [stochastic] function $\tilde{\eta}(\cdot)$, which is defined as $\tilde{\eta}(d\mathbf{x}(t)) = [\delta(d\mathbf{x}(t), -\boldsymbol{\iota})]^T (\mathbf{H} - \mathbf{I}_N^*) [\delta(d\mathbf{x}(t), \boldsymbol{\iota})]$, where $\mathbf{H}$ is a random matrix of independent jumps; the (i, j) -th element of $\mathbf{H}$ is always positive and follows a distribution $\tilde{\mathcal{D}}_{ij}$, with the diagonal elements equal to zero by construction, and for the notation: $\boldsymbol{\iota}$ is the $(N \times 1)$ vector of ones, $\delta(\cdot, \cdot)$ is a Kronecker function applied elementwise⁹, and $\mathbf{I}_N^*$ is an $(N \times N)$ matrix with all elements equal to one, but the main diagonal equal to zero, $[\mathbf{I}_N^*]_{ij} = 1 - \delta(i, j)$;
- the instantaneous drift can take N distinct values, $\mu(\mathbf{x}(t)) = \mathbf{x}^T(t) \boldsymbol{\mu}$. In addition, $\mathbf{E}[\mathbf{H}] = \mathbf{M}_\eta + \mathbf{I}_N^*$, and the rate matrix $\mathbf{Q}$ is defined below; and finally
- $\sigma(\mathbf{x}(t)) = \mathbf{x}^T(t) \boldsymbol{\sigma}$.

⁸Therefore $\mathbf{e}_i$ is an $(N \times 1)$ vector of zeros, with its i -th entry equal to one.

⁹Note that if there is not a state change, then $d\mathbf{x}(t) = \mathbf{0}$. If there is a change from state $\mathbf{e}_i$ to $\mathbf{e}_j$, then $dx_i(t) = -1$ and $dx_j(t) = 1$, with the other elements being equal to zero.

The above equation can be rewritten by setting $s(t) = \ln S(t)$, and suppressing the time variable¹⁰, as

$$ds = \left\{ \mu(\mathbf{x}) - \mathbf{x}^T [\mathbf{Q} \odot \mathbf{M}_\eta] \boldsymbol{\iota} - \frac{1}{2} \mathbf{x}^T [\boldsymbol{\sigma} \odot \boldsymbol{\sigma}] \right\} dt + \sigma(\mathbf{x}) d\varepsilon + \eta(d\mathbf{x}), \quad (2)$$

with $\eta(d\mathbf{x}) = [\delta(d\mathbf{x}, -\boldsymbol{\iota})]^T \mathbf{J} [\delta(d\mathbf{x}, \boldsymbol{\iota})]$, with the elements of $\mathbf{J} = \ln_\odot(\mathbf{H})$ following the appropriate distribution $\mathcal{D}_{ij}$, and $\ln_\odot$ is understood to denote the logarithms to be taken elementwise. For illustration purposes it will be later assumed that $\mathbf{H}$ will be a collection of Weibull distributions with parameters (α, γ) , $\tilde{\mathcal{D}}_{ij} = \mathcal{W}(\alpha_{ij}, \gamma_{ij})$, implying the extreme value [or Fisher-Tippett] distribution for $\mathbf{J}$ with the same parameters¹¹, $\mathcal{D}_{ij} = \mathcal{FT}(\mu_{ij}, \sigma_{ij})$.

Consider also the σ -fields $\mathcal{F}_t^{S, \mathbf{x}} = \sigma(\{S(u), \mathbf{x}(u)\} : -\infty \leq u \leq t)$, and $\mathcal{F}_t^S = \sigma(S(u) : -\infty \leq u \leq t)$ $t \geq 0$, serving as the information available at time t . Observe that the asset price process is trivially adapted to both filtrations, while the state process is not $\mathcal{F}_t^S$ -adapted; its conditional expectation has to be derived based on $\mathcal{F}_t^S$. This conditional expectation is denoted by $\boldsymbol{\xi}(t) = \mathbf{E}[\mathbf{x}(t) | \mathcal{F}_t^S]$. Also denote as $\boldsymbol{\xi}(t|u) = \mathbf{E}[\mathbf{x}(t) | \mathcal{F}_u^S]$ the filtered [forecasted] state probabilities for $u > t$ [$u < t$].

¹⁰This transformation is due to the Poisson stochastic differential form, see Kushner (1967) and the references therein.

¹¹The parameters of the Weibull $\mathcal{W}(\alpha, \gamma)$ distribution are: the characteristic life α , and the shape parameter γ , such that the density is $f(x) = \gamma \alpha^{-\gamma} x^{\gamma-1} \exp[-(x/\alpha)^\gamma]$; whereas the Fisher-Tippett $\mathcal{FT}(\mu, \sigma)$ density is $f(x) = \frac{1}{\sigma} \exp[(x - \mu)/\sigma - \exp[(x - \mu)/\sigma]]$, with characteristic function $\bar{f}(\theta) = \Gamma(1 + \varpi \sigma \theta) \exp[\varpi \mu \theta]$. The special case where $\gamma = 1$ gives rise to the exponential and the Gumbel distributions respectively. $\varpi = \sqrt{-1}$ and $\Gamma(\cdot)$ denotes the Gamma function [for all special functions in this paper, consult Abramowitz and Stegun (1972)]. The arguments of the Gamma function satisfy the constraint $\text{Re } z > 0$, which allows the values to be computed very easily with accuracy at least of order 10^{-12} [see Appendix for technical details]. The relationship between the parameters is the following: If $X \sim \mathcal{W}(\alpha, \gamma)$ then $\ln X \sim \mathcal{FT}(\mu = \ln \alpha, \sigma = 1/\gamma)$.

The state process is considered independent of the asset price process. The infinitesimal transition probabilities are given by

$$\mathbf{P}[\mathbf{x}(t+dt) = \mathbf{e}_j | \mathbf{x}(t) = \mathbf{e}_i] = \begin{cases} q_{ij}dt + o(dt), & \text{for } i \neq j \\ 1 + q_{ii}dt + o(dt), & \text{for } i = j \end{cases}, \quad (3)$$

where $q_{ii} = -\sum_{j=1..N, j \neq i} q_{ij}$. These probabilities can be arranged in the infinitesimal transition probability matrix which can be written in the form $\mathbf{P}(dt) = \mathbf{I}_N + \mathbf{Q}dt + \mathbf{o}(dt)$, where $\mathbf{o}(dt)$ denotes a matrix with elements of order $o(dt)$, and the rate matrix is defined as

$$\mathbf{Q} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1N} \\ q_{21} & q_{22} & \cdots & q_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1} & q_{N2} & \cdots & q_{NN} \end{pmatrix}.$$

In general the Chapman-Kolmogorov relationships imply that

$$\mathbf{P}(\tau) = \exp\{\tau \mathbf{Q}\}. \quad (4)$$

One can verify that the solution of the asset price diffusion (2) is

$$\begin{aligned} s(\tau) = s(0) + \int_0^\tau \left\{ \mathbf{x}^T \boldsymbol{\mu} - \mathbf{x}^T [\mathbf{Q} \odot \mathbf{M}_\eta] \boldsymbol{\iota} - \frac{1}{2} \mathbf{x}^T [\boldsymbol{\sigma} \odot \boldsymbol{\sigma}] \right\} dt \\ + \int_0^\tau \mathbf{x}^T \boldsymbol{\sigma} d\varepsilon(t) + \int_0^\tau [\delta(d\mathbf{x}, -\boldsymbol{\iota})]^T \mathbf{J}[\delta(d\mathbf{x}, \boldsymbol{\iota})], \end{aligned} \quad (5)$$

or in a more compact form

$$s(\tau) - s(0) = \int_0^\tau dw(t). \quad (6)$$

The distribution of the random variable $dw(t)$ depends on the transitions of the CTMC $\mathbf{x}(t)$: If the chain remains in regime $\mathbf{e}_i$ during the interval

(t_-, t) , then it is distributed normally, with the instantaneous mean $\mu_{ii}dt = \mathbf{e}_i^T \{ \boldsymbol{\mu} - \mathbf{x}^T [\mathbf{Q} \odot \mathbf{M}_\eta] \boldsymbol{\iota} - \frac{1}{2} \boldsymbol{\sigma} \odot \boldsymbol{\sigma} \} dt$ and volatility $\sigma_{ii}\sqrt{dt} = \mathbf{e}_i^T \boldsymbol{\sigma} \sqrt{dt}$. If there is a transition from state $\mathbf{e}_i$ to state $\mathbf{e}_j$, then $dw(t)$ is distributed according to some distribution $\mathcal{D}_{ij}$.

2.2 The moment generating function of the state dependent stochastic process

In order to compute the moment generating function and consequently the Fourier transform of the asset log-price density, the following definitions are needed.

Definition 1 (ε -skeleton) *Consider a CTMC $\mathbf{x}(t)$ defined on the state space $\mathcal{S}$. Define the ε -skeleton of $\mathbf{x}(t)$ to be the Discrete Time Markov Chain $\mathbf{x}_n^{(\varepsilon)}$, defined on the same state space $\mathcal{S}$, with one step transition probabilities formed in the matrix*

$$\mathbf{P}^{(\varepsilon)} = \left(\mathbf{P} \left[\mathbf{x}_{n+1}^{(\varepsilon)} = \mathbf{e}_j | \mathbf{x}_n^{(\varepsilon)} = \mathbf{e}_i \right] \right)_{\{i,j\} \in \mathcal{S}^2} \stackrel{def}{=} \mathbf{P}(\varepsilon).$$

It is obvious, from the definition of the ε -skeleton, that as $\varepsilon \downarrow 0$, the skeleton becomes finer and the discrete chain $\mathbf{x}_n$ converges to the continuous time one, $\mathbf{x}(t)$.

Definition 2 (discounted cumulative score) *Consider a realization of a CTMC $\mathbf{x}(t)$ defined on the state space $\mathcal{S}$. Construct a process in the following way: With a transition from state $\mathbf{e}_i$ to state $\mathbf{e}_j$ at time t associate a score $dw(t)$, whose distribution $F_{ij}^w(x, dt)$ depends only on the states $\mathbf{e}_i$ and $\mathbf{e}_j$. In addition make the assumption that given a state sequence, that is to say given the realization of the chain $\mathbf{x}(t)$, the random variables $dw(t_1)$ and $dw(t_2)$ are mutually independent for all times t_1 and t_2 . In addition consider a*

function¹² $\rho : \mathbf{R} \rightarrow \mathbf{R}$, which is non-increasing and continuous, with $\rho(0) = 1$. This function $\rho(t)$ serves as the discounting factor. Define the **discounted cumulative score** as the additive process represented by the Itô integral

$$y(\tau) = \int_0^\tau \rho(t) dw(t) .$$

Example 3 (sojourn time) Let the scores be defined in the following way: $dw(t) = dt$, if the transition taking place at time t terminates in state $\mathbf{e}_i$, and $dw(t) = 0$, otherwise. In addition, set $\rho(t) = 1$. Then, the cumulative score $y_i(\tau)$ is representing the total time spent in state $\mathbf{e}_i$ up to time τ .

A special case, which leads to moment generating functions in closed form, arises when one considers scores which belong to certain families: When the CTMC remains in state i over the infinitesimal time dt , the score associated, say $dw(t)$, follows a distribution with density $f_{ii}(x, dt)$ which has a Laplace

¹²Although the model presented here does not incorporate mean reversion other than the one implicit in the Markov chain, therefore $\rho(x) = 1$, the proposition is proved for this more general case. This could be proved useful if one was to model series where this mean reversion feature is present.

transform of the form^{13,14}

$$\mathcal{L}\{f_{ii}(x, dt)\}(\theta, dt) = 1 + g_{ii}(\theta) dt + o(dt), \quad (7)$$

while the score associated with a transition from a state i into a state j will be associated with a random variable which is independent of the infinitesimal time dt , with density $f_{ij}(x)$ and corresponding Laplace transform

$$\mathcal{L}\{f_{ij}(x, dt)\}(\theta, dt) = g_{ij}(\theta) + O(dt). \quad (8)$$

The normal distribution of the diffusion in (2) is of the form of equation (7). Define the product-integral of a matrix function, say $\Phi(\cdot)$, which is denoted by $\mathfrak{P}_U \{\mathbf{I} + d\Phi(u)\}$, as in Dollard and Friedman (1979). The stochastic behavior of $y(\tau)$, as it is inferred by its mgf, is established in the following

Proposition 4 (mgf of the disc. cum. score) *Conditional on the initial state $\mathbf{x}(0)$, the parameters $\mathbf{Q}$, and the functions $\rho(\cdot)$ and $g_{ij}(\cdot)$, the mgf*

¹³Alternatively one can say that conditional on the state of the world $\mathbf{x}^*$, $w(t)$ is a Lévy process.

¹⁴This class of random variables is very rich, and includes the family of the infinitely divisible distributions which result from Lévy processes. The Laplace transform of a random variable generated by a Lévy process, ω , is of the form

$$\omega(\theta, \tau) = e^{\tau\psi(\theta)}, \quad \psi(\theta, 0) = 0,$$

where ψ is the characteristic exponent of the process. Distributions such as the Normal, Gamma, Poisson etc. belong to this family. Of course in the continuous time framework the only process that will give a.s. continuous paths is the Brownian motion, where the increments are normally distributed [see for example Feller (1971, Chapter XVII) and Bertoin (1996) for the relationship between the Lévy processes and the infinitely divisible distributions, and Øksendal (1991) for the a.s. continuity of the Brownian motion].

of $y(\tau) - y(0) = y(\tau)$ will be given by ¹⁵

$$\phi^y(\theta, \tau | \mathbf{x}(0)) = [\mathbf{x}(0)]^T \mathfrak{P}_{u \in [0, \tau]} \{ \mathbf{I} + \mathbf{B}(\theta, u) du \} \boldsymbol{\nu}, \quad (9)$$

where $\mathbf{B}(\theta, u)$ has elements of the form

$$\beta_{ij}(\theta, u) = \begin{cases} q_{ii} + g_{ii}(\rho(u)\theta), & \text{when } i = j \\ q_{ij}g_{ij}(\rho(u)\theta), & \text{when } i \neq j \end{cases}.$$

PROOF: See Appendix.

Corollary 5 *If $\rho(t) = 1, \forall t \in [0, +\infty)$, then the mgf of $y(\tau) - y(0) = y(\tau)$ is given by*

$$\phi^y(\theta, \tau | \mathbf{x}(0)) = [\mathbf{x}(0)]^T \exp \{ \tau \mathbf{B}(\theta) \} \boldsymbol{\nu},$$

where $\mathbf{B}(\theta)$ has elements of the form

$$\beta_{ij}(\theta) = \begin{cases} q_{ii} + g_{ii}(\theta), & \text{when } i = j \\ q_{ij}g_{ij}(\theta), & \text{when } i \neq j \end{cases}.$$

PROOF: The proof follows directly from the properties of the product-integral.¹⁶

Corollary 6 *Conditional on the initial distribution $\boldsymbol{\xi}(0)$, the parameters $\mathbf{Q}$, and the function $g_{ij}(\cdot)$, the mgf of $y(\tau)$ will be given by*

$$\phi^y(\theta, \tau | \mathcal{F}_0^S) = [\boldsymbol{\xi}(0)]^T \exp \{ \tau \mathbf{B}(\theta) \} \boldsymbol{\nu}.$$

¹⁵ The exponential of a matrix is defined as $\exp \mathbf{A} = \sum_{i=0}^{\infty} \frac{1}{i!} \mathbf{A}^i$. In the empirical part of this paper, this has been evaluated using a (6, 6)-Padé type approximation, after rescaling the corresponding complex matrix, which achieves overall accuracy of order 10^{-20} at least.

¹⁶ In particular that for a time invariant matrix Φ , $\mathfrak{P}_{[0, \tau]} \{ \mathbf{I} + d\Phi \} = \exp \{ \tau \Phi \}$.

PROOF: Follows immediately from the above corollary, the definitions of the filtrations and Bayes' rule. $\square$

The moments of the discounted cumulative score $y(\tau)$ can be obtained by differentiating (9) at $\theta = 0$. These calculations are not straightforward, due to the matrix exponential involved and the fact that the matrix function $\mathbf{B}$ and its derivatives do not necessarily commute. The chain rule for matrix differentiation can be applied in the lines of Mathias (1997):

Definition 7 (BUTB Toeplitz matrix) *Consider the $(n \times n)$ matrices $\mathbf{M}_0, \mathbf{M}_1, \dots, \mathbf{M}_k$. Then, the block upper triangular block (BUTB) Toeplitz matrix, denoted by $T[\mathbf{M}_0, \mathbf{M}_1, \dots, \mathbf{M}_k]$, is defined as the $(n(k+1) \times n(k+1))$ block matrix with block (i, j) of size $(n \times n)$ equal to $\mathbf{M}_{j-i}$ for $j \geq i$ and zeros everywhere else. For example,*

$$T[\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_2] = \begin{pmatrix} \mathbf{M}_0 & \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{0} & \mathbf{M}_0 & \mathbf{M}_1 \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_0 \end{pmatrix}.$$

Proposition 8 (matrix exponential derivatives) *Consider a $(N \times N)$ matrix function $\Phi(\mathbf{X})$, $\mathbf{R}^{N \times N} \rightarrow \mathbf{R}^{N \times N}$, differentiable at $\mathbf{X} = \mathbf{0}$. Under general conditions, given in Mathias (1997),*

$$\exp \left\{ T \left[\Phi(0), \frac{\Phi^{(1)}(0)}{1!}, \dots, \frac{\Phi^{(k)}(0)}{k!} \right] \right\} = T \left[\exp \{ \Phi(0) \}, \frac{1}{1!} \frac{d}{d\theta} \exp \{ \Phi(0) \}, \dots, \frac{1}{k!} \frac{d}{d\theta} \exp \{ \Phi(0) \} \right].$$

PROOF: See Mathias (1997, Theorem 4.1).

Assuming, without loss of generality, that $s(0) = 0$, the k -th moment of $s(\tau)$, given in (6), is $m_k(\tau) = \frac{d}{d\theta} \phi^y(0, \tau)$. Say that one wants to evaluate the

first k_0 moments; the above proposition will give that conditional on $\mathcal{F}_0^{S,\mathbf{x}}$,

$$m_k \left(\tau | \mathcal{F}_0^{S,\mathbf{x}} \right) = k! [\mathbf{x}(0)]^T \mathbf{A}_{[1,k+1]}(\tau) \boldsymbol{\iota}, \quad (10)$$

where $\mathbf{A}_{[1,k]}(\tau)$ denotes the $(1, k)$ block of the matrix

$$\begin{aligned} \mathbf{A}(\tau) &= \\ &= \exp \left\{ \tau \times T \left[\mathbf{B}(0), \frac{1}{1!} \frac{d\mathbf{B}(\theta)}{d\theta} \Big|_{\theta=0}, \dots, \frac{1}{k_0!} \frac{d^{k_0}\mathbf{B}(\theta)}{d\theta^{k_0}} \Big|_{\theta=0} \right] \right\}. \end{aligned} \quad (11)$$

Corollary 9 (sojourn time) *For the sojourn time variable, the first and second moments conditional on the initial state $\mathbf{x}(0)$ and the parameters $\mathbf{Q}$, are given by*

$$\begin{aligned} m_1 \left(\tau | \mathcal{F}_0^{S,\mathbf{x}} \right) &= [\mathbf{x}(0)]^T \mathbf{A}_{[1,2]}(\tau) \boldsymbol{\iota}, \\ m_2 \left(\tau | \mathcal{F}_0^{S,\mathbf{x}} \right) &= 2 [\mathbf{x}(0)]^T \mathbf{A}_{[1,3]}(\tau) \boldsymbol{\iota}, \end{aligned}$$

where

$$\mathbf{A} = \exp \left\{ \begin{pmatrix} \tau \mathbf{Q} & \tau \mathbf{I}_N^{ii} & \mathbf{0} \\ \mathbf{0} & \tau \mathbf{Q} & \tau \mathbf{I}_N^{ii} \\ \mathbf{0} & \mathbf{0} & \tau \mathbf{Q} \end{pmatrix} \right\},$$

and $\mathbf{I}_N^{ij}$ is an $(N \times N)$ matrix of zeros, with only the (i, j) -th element equal to one.

PROOF: See Appendix.

Corollary 10 (pure regime shift model) *Consider the stochastic process in (2), and assume that $s(0) = 0$ and $\mathbf{J} = \mathbf{0}$. Then, the conditional on*

$\mathcal{F}_0^{S,\mathbf{x}}$ moments of the stochastic process $s(\tau)$ are given by

$$m_k \left(\tau | \mathcal{F}_0^{S,\mathbf{x}} \right) = k! [\mathbf{x}(0)]^T \mathbf{A}_{[1,k+1]}(\tau) \boldsymbol{\nu},$$

where $\mathbf{A}$ is the matrix

$$\mathbf{A} = \exp \left\{ \tau \times \begin{pmatrix} \mathbf{D}^{(0)} & \mathbf{D}^{(1)} & \mathbf{D}^{(2)} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{(0)} & \mathbf{D}^{(1)} & & \mathbf{0} \\ \vdots & & \vdots & \ddots & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}^{(0)} \end{pmatrix} \right\},$$

and the block-elements are $\mathbf{D}^{(0)} = \mathbf{Q}$, $\mathbf{D}^{(1)} = \sum_{i=1}^N \theta \mathbf{e}_i^T [\boldsymbol{\mu} - \frac{1}{2} \boldsymbol{\sigma} \odot \boldsymbol{\sigma}] \mathbf{I}_N^{ii}$, and $\mathbf{D}^{(2)} = \sum_{i=1}^N [\mathbf{e}_i^T \boldsymbol{\sigma}]^2 \mathbf{I}_N^{ii}$. The time has been suppressed for notational convenience.

PROOF: See Appendix.

Corollary 11 (regime shift model, extreme jumps) Consider the stochastic process in (2), where the matrix $\mathbf{J}$ is Fisher-Tippett distributed. Then, the conditional on $\mathcal{F}_0^{S,\mathbf{x}}$ moments of the stochastic process $y(\tau)$ are given by

$$m_k \left(\tau | \mathcal{F}_0^{S,\mathbf{x}} \right) = k! [\mathbf{x}(0)]^T \mathbf{A}_{[1,k+1]}(\tau) \boldsymbol{\nu},$$

where $\mathbf{A}$ is the upper block triangular matrix

$$\mathbf{A} = \exp \left\{ \tau \times \begin{pmatrix} \mathbf{D}^{(0)} & \mathbf{D}^{(1)} & \mathbf{D}^{(2)} & \dots & \mathbf{D}^{(k)} \\ \mathbf{0} & \mathbf{D}^{(0)} & \mathbf{D}^{(1)} & & \mathbf{D}^{(k-1)} \\ \vdots & & \vdots & \ddots & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}^{(0)} \end{pmatrix} \right\},$$

and the matrices $\mathbf{D}^{(k)}$ are given in the appendix. The time has been suppressed for notational convenience.

PROOF: See Appendix.

Denote with $f_{s(\tau)}(s|\mathcal{F})$ the [conditional of $\mathcal{F}$] density of the asset log-price difference after time τ has elapsed, $y(\tau) = s(\tau) - s(0)$. The Laplace transform of this density, which is denoted with $f_{\Delta s(\tau,0)}^{\mathcal{L}}(\theta|\mathcal{F})$ is derived using equations (6) and the definition of $\phi^y(\theta, \tau|\mathcal{F})$, as

$$f_{\Delta s(\tau,0)}^{\mathcal{L}}(\theta|\mathcal{F}) = \phi^y(\theta, \tau|\mathcal{F}). \quad (13)$$

3 Estimation methodology

This section discusses alternative methods that one can utilize in order to obtain estimates of the parameter vector $\boldsymbol{\vartheta}$, from a set of observations, $\{S(t_u)\}_{u=1,\dots,M+1}$, when the state of the world is either observed or unobserved. It also discusses the filter of Hamilton (1989), as a method of retrieving the conditional probabilities of the state, $\boldsymbol{\xi}_{t|t} = (\mathbf{P}[\mathbf{x}(t) = \mathbf{e}_i|\mathcal{F}_t^S])_{i=1,\dots,N}$, from the data.

3.1 Maximum likelihood

In this section, the assumption that $\mathcal{F} = \mathcal{F}_0^S$ is maintained, that is to say, the state of the world is considered unobservable as in most empirical applications. Consider the time series of log-asset differences, say $\{\Delta s(t_u)\}$, recorded at times $\{t_u\}$, $u = 1, \dots, M$. The derivation of the Laplace transform of the log-asset density in equation (13) allows one to calculate the corresponding Fourier transform $f_{\Delta s(\tau,0)}^{\mathcal{F}}(\theta|\mathcal{F})$, as

$$f_{\Delta s(\tau,0)}^{\mathcal{F}}(\theta|\mathcal{F}) = f_{\Delta s(\tau,0)}^{\mathcal{L}}(\varpi\theta|\mathcal{F}), \quad (14)$$

where ϖ denotes the imaginary unit. Standard Fourier transformation will give the density of the asset log-price difference $\Delta s(t_u) = s(t_{u+1}) - s(t_u)$,

over a time interval $\Delta t_u = t_{u+1} - t_u$, conditional on the initial state $\mathbf{x}(t_u)$

$$f(\Delta s(t_u) | \mathbf{x}(t_u), \boldsymbol{\vartheta}) = \frac{1}{\pi} \int_0^{+\infty} \text{Re} \left[e^{-\varpi \theta \Delta s(t_u)} f_{\Delta s(t_{u+1}, t_u)}^{\mathcal{F}}(\theta | \mathbf{x}(t_u)) \right] d\theta,$$

where $f_{\Delta s(t_{u+1}, t_u)}^{\mathcal{F}}(\theta | \mathbf{x}(t_u)) = [\mathbf{x}(t_u)]^T \exp \{ \Delta t_u \times \mathbf{B}(\varpi \theta) \} \boldsymbol{\iota}$. This implies an appealing maximum likelihood (ML) estimation algorithm in the lines of Hamilton (1993), where the conditional [on the state] distribution is not normal. The unconditional log-likelihood will be

$$\begin{aligned} L \left(\{ \Delta s(t_u) \}_{u=1, \dots, M}; \boldsymbol{\vartheta} \right) \\ = \sum_{u=1}^{M-1} \ln \left\{ \frac{1}{\pi} \int_0^{+\infty} \text{Re} \left[e^{-\varpi \theta \Delta s(t_u)} f_{\Delta s(t_{u+1}, t_u)}^{\mathcal{F}}(\theta | \boldsymbol{\xi}^\infty) \right] d\theta \right\}, \end{aligned}$$

with $\boldsymbol{\xi}^\infty = [\xi_i^\infty]_{i=1, \dots, N}$ is the vector of the ergodic probabilities, given from the diagonal elements of the limit $\lim_{\tau \rightarrow \infty} \exp \{ \tau \mathbf{Q} \}$. This likelihood can be numerically maximized in order to retrieve the estimates $\hat{\boldsymbol{\vartheta}}$.

Corollary 12 *If the jumps are Weibull distributed, with distribution $\tilde{\mathcal{D}}_{ij} = \mathcal{W}(\alpha_{ij}, \gamma_{ij})$, then the log-jumps will be Fisher-Tippett distributed, with $\mathcal{D}_{ij} = \mathcal{FT}(\ln \alpha_{ij}, \gamma_{ij}^{-1})$, and the criterion function to be minimized will become*

$$\begin{aligned} L \left(\{ \Delta s(t_u) \}_{u=1, \dots, M}; \boldsymbol{\vartheta} \right) \\ = \sum_{u=0}^M \ln \left\{ \int_0^{+\infty} \text{Re} \left[e^{-\varpi \theta \Delta s(t_u)} [\boldsymbol{\xi}^\infty]^T \exp \{ \Delta t_u \times \mathbf{B}(\theta) \} \boldsymbol{\iota} \right] d\theta \right\}, \end{aligned}$$

where the matrix $\mathbf{B}(\theta)$ will be of the form

$$\beta_{ij}(\theta) = \begin{cases} q_{ii} - [\mathbf{e}_i^T (\mathbf{Q} \odot \mathbf{M}_\eta) \boldsymbol{\iota} + \frac{1}{2} \sigma_{ii}^2] \varpi \theta - \frac{1}{2} \sigma_{ii}^2 \theta^2, & \text{when } i = j \\ q_{ij} \Gamma(1 + \varpi \theta / \gamma_{ij}) \alpha_{ij}^{\varpi \theta}, & \text{when } i \neq j \end{cases},$$

and the matrix $\mathbf{M}_\eta$ will have elements $[\mathbf{M}_\eta]_{ij} = \alpha_{ij}\Gamma(1 + \gamma_{ij}^{-1}) - 1$, the expected values of the corresponding Weibull distribution.

The algorithm described in Hamilton (1989) and Hamilton (1994) can be used in order to retrieve the filtered state probabilities, denoted with $\{\boldsymbol{\xi}_{t_u|t_u}\}_{u=1,\dots,M}$, based on the recursions:

$$\boldsymbol{\xi}_{t_{u+1}|t_{u+1}} = \frac{\left[\boldsymbol{\xi}_{t_{u+1}|t_u} \odot \left(f\left(\Delta s(t_u) | \mathbf{e}_i, \hat{\boldsymbol{\vartheta}}\right) \right)_{i=1,\dots,N} \right]}{\boldsymbol{\iota}^T \left[\boldsymbol{\xi}_{t_{u+1}|t_u} \odot \left(f\left(\Delta s(t_u) | \mathbf{e}_i, \hat{\boldsymbol{\vartheta}}\right) \right)_{i=1,\dots,N} \right]}, \quad (15a)$$

$$\boldsymbol{\xi}_{t_{u+1}|t_u}^T = \boldsymbol{\xi}_{t_u|t_u}^T \exp \left\{ (t_{u+1} - t_u) \times \hat{\mathbf{B}}(\theta) \right\} \quad (15b)$$

3.2 Method of the moments

An alternative way to obtain estimates of the parameters of interest is via the method of the moments, if the sample elements are equally spaced across time, $\Delta t_u = \Delta t$, $\forall u = 1, \dots, M$. Since via the moment generating function one can compute raw moments of any order, equation (10) can be used in order to obtain unconditional GMM estimates of the parameter vector $\boldsymbol{\vartheta}$. More specifically, denote by $\mathbf{h}(\boldsymbol{\vartheta}, \Delta s(t_u))$ the $(k \times 1)$ vector with elements

$$h_i(\boldsymbol{\vartheta}, \Delta s(t_u)) = [\Delta s(t_u)]^i - i! [\boldsymbol{\xi}^\infty]^T \mathbf{A}_{[1,i+1]}(\Delta t) \boldsymbol{\iota},$$

with $\mathbf{A}_{[1,i+1]}$ the appropriate submatrix given by (11); and by $\mathbf{g}(\boldsymbol{\vartheta})$ the corresponding $(k \times 1)$ vector of the sample moment deviations

$$g_i(\boldsymbol{\vartheta}) = \frac{1}{M} \sum_{u=0}^M [\Delta s(t_u)]^i - i! [\boldsymbol{\xi}^\infty]^T \mathbf{A}_{[1,i+1]}(\Delta t) \boldsymbol{\iota}.$$

The criterion function to be minimized will therefore be the quadratic form

$$Q(\boldsymbol{\vartheta}) = \mathbf{g}(\boldsymbol{\vartheta}) \mathbf{W}_M [\mathbf{g}(\boldsymbol{\vartheta})]^T$$

with $\mathbf{W}_M$ the weighting matrix, which can be estimated [given $\boldsymbol{\vartheta}$] from its Newey-West estimate [see Hamilton (1994) for details]. Starting from an arbitrary weighting matrix, the above criterion function can be minimized in order to obtain an initial estimate of the parameter vector. Using this estimate the weighting matrix can be recomputed, the criterion function can be reconstructed, and a new estimate of $\boldsymbol{\vartheta}$ can be calculated. This procedure can be iterated for a number of times, until the criterion function and the estimate of $\boldsymbol{\vartheta}$ become stable.

3.3 Direct characteristic function estimation

Feuerverger and McDunnough (1981) suggest that estimates which achieve the Cramer-Rao bound can be obtained, by estimating the empirical characteristic function directly using the $k - L$ procedure. The empirical characteristic function $\hat{f}(\theta) = \frac{1}{M} \sum_{u=0}^M \exp[\varpi \theta \Delta s(t_u)]$ is calculated at k points $\{\theta_i\}_{i=1, \dots, k}$, and maximum likelihood is employed in order to minimize the distance between the empirical and the theoretical characteristic function. This is analogous to the method of the moments, where the moment conditions are replaced by the $(2k \times 1)$ vector $\mathbf{g}^*(\boldsymbol{\vartheta})$, with elements

$$\begin{aligned} g_i^*(\boldsymbol{\vartheta}) &= \begin{cases} \text{Re} \left\{ \frac{1}{M} \sum_{u=0}^M \exp[\varpi \theta_i \Delta s(t_u)] - f_{\Delta s(t)}^{\mathcal{F}}(\theta_i | \boldsymbol{\xi}^\infty) \right\} & , i \leq k \\ \text{Im} \left\{ \frac{1}{M} \sum_{u=0}^M \exp[\varpi \theta_{i-k} \Delta s(t_u)] - f_{\Delta s(t)}^{\mathcal{F}}(\theta_{i-k} | \boldsymbol{\xi}^\infty) \right\} & , i > k \end{cases} . \end{aligned}$$

The criterion function is

$$Q^*(\boldsymbol{\vartheta}) = \mathbf{g}^*(\boldsymbol{\vartheta}) \Sigma^{-1} [\mathbf{g}^*(\boldsymbol{\vartheta})]^T, \quad (16)$$

with Σ the corresponding variance-covariance matrix.

It is shown that the above procedure yields arbitrarily high asymptotic efficiency, by using a sufficiently dense grid $\{\theta_i\}$, converging to the Cramer-Rao bound as k increases. The realizations of the stochastic process in (6), depend on the Markov chain $\mathbf{x}(t)$, therefore one has to construct the conditional [on $\mathcal{F}_{t_u}^{S,\mathbf{x}}$] characteristic function, giving the following moment conditions

$$\begin{aligned} & [\mathbf{g}^*(\boldsymbol{\vartheta})]_i \\ = & \begin{cases} \text{Re} \left\{ \frac{1}{M} \sum_{u=0}^M \left[\exp [\varpi \theta_i \Delta s(t_u)] - f_{\Delta s(t_u)}^{\mathcal{F}} \left(\theta_i | \mathcal{F}_{t_u}^{S,\mathbf{x}} \right) \right] \right\} & , i \leq k \\ \text{Im} \left\{ \frac{1}{M} \sum_{u=0}^M \left[\exp [\varpi \theta_{i-k} \Delta s(t_u)] - f_{\Delta s(t_u)}^{\mathcal{F}} \left(\theta_{i-k} | \mathcal{F}_{t_u}^{S,\mathbf{x}} \right) \right] \right\} & , i > k \end{cases} . \end{aligned}$$

4 Option Pricing

In this section, we maintain the assumption that $\mathcal{F} = \mathcal{F}_0^{S,\mathbf{x}}$, and suppress the conditioning. To this end, one has to construct the risk neutral measure, which in turn will define the contingent claim prices. There are three sources of uncertainty in equation (1) or (2): the uncertainty of the realizations of the future errors $d\varepsilon$ and jumps $\mathbf{H}$ [or equivalently $\mathbf{J}$], and the uncertainty of the future regime sequence $\mathbf{x}$.

Consider the diffusion of the asset log-price under the risk neutral mea-

sure, in line with Merton (1976) and Bates (1995)¹⁷

$$ds = rdt - \mathbf{x}^T (\bar{\mathbf{Q}} \odot \bar{\mathbf{M}}_\eta) \boldsymbol{\iota} dt - \frac{1}{2} \mathbf{x}^T [\boldsymbol{\sigma} \odot \boldsymbol{\sigma}] dt + \mathbf{x}^T \boldsymbol{\sigma} d\bar{\varepsilon} + [\delta(d\mathbf{x}, -\boldsymbol{\iota})]^T \bar{\mathbf{J}} [\delta(d\mathbf{x}, \boldsymbol{\iota})], \quad (17)$$

where $\bar{\varepsilon}$ follows a standardized normal distribution under the risk neutral measure and $\bar{\mathbf{J}}$ is a transformation of $\mathbf{J}$ under the risk neutral measure. This means that a jump which is due to a regime shift from $\mathbf{e}_i$ to $\mathbf{e}_j$ will be distributed according to the distribution $\bar{\mathcal{D}}_{ij}$, the corresponding density having a Laplace transform $\bar{g}_{ij}(\theta)$. The random process $\mathbf{x}$ is now a CTMC under the risk neutral measure generated by a rate matrix $\bar{\mathbf{Q}}$.

Representative agent general equilibrium models derive the risk neutral parameters as modifications of the true ones, taking into account the effect of the state variables on the diffusion of the percentage changes of the marginal utility of nominal wealth, U_W , its diffusion denoted by $d\mathcal{M} = \frac{dU_W}{U_W}$. Denote by $d\mathcal{M}_{ij}$ the former quantity, given a chain change from state $\mathbf{e}_i$ to $\mathbf{e}_j$. The relationships between the true and the risk neutral parameters are the following: The elements of the rate matrix, $\bar{\mathbf{Q}}$, are given, in the lines of Bates (1995), by

$$\bar{q}_{ij} = q_{ij} \mathbf{E}[1 + d\mathcal{M}_{ij}], \text{ for } i \neq j, \quad (18)$$

and $\bar{q}_{ii} = -\sum_{j=1..N, j \neq i} \bar{q}_{ij}$; the elements of the expected jump matrix, $\bar{\mathbf{M}}_\eta$, are given by

$$\bar{\mu}_{ij} = \mu_{ij} + \frac{\mathbf{Cov}[J_{ij}, d\mathcal{M}_{ij}]}{\mathbf{E}[1 + d\mathcal{M}_{ij}]}. \quad (19)$$

¹⁷ The expected instantaneous drift of the asset price diffusion is equal to $-\rho[s(t) - \boldsymbol{\mu}^T \mathbf{x}(t)] dt + \boldsymbol{\iota}^T (\mathbf{Q} \odot \mathbf{M}_J)^T \mathbf{x}(t) dt$. Risk neutrality demands that the discounted asset prices form martingales [Pliska (1997)], therefore the expected log-return of the asset must be set equal to the instantaneous risk free rate of return.

There are two important issues concerning equations (17) and (18)-(19):

First, the relationships between the *true* process given in equation (2) and the *risk neutral* process given in (17). The instantaneous drift of the true process has disappeared; moreover the risk neutral process does not exhibit mean reversion through the mean reverting behavior of the CTMC $\mathbf{x}$. One could argue that it might not be necessary to include such a parametrized drift in the first place. This is not true: Lo and Wang (1995) show that although the drift does not enter the option pricing formulas directly, it can have a substantial effect on the prices. The intuition is straightforward: if the true model is estimated correctly then the estimated parameters that enter the formulae directly [the volatilities for example] will be different compared to the ones estimated using the wrong specification.

Second, the computation of the prices of risk in equations (18) and (19). Given the information of the representative agent, $\mathcal{F}_t^{S,\mathbf{x}}$, the securities market is incomplete, that is to say the risk neutral parameters cannot be retrieved and therefore correct option prices cannot be computed. There are three possible ways of tackling this problem:

1. One can make the assumption that the regime risk and the jump risk cannot be diversified, or equivalently that the expected marginal utility of wealth is not affected by regime changes, giving $d\mathcal{M}_{ij} = 0$. This somewhat strict assumption results into risk neutral parameters that are equal to their true counterparts [see Chourdakis and Tzavalis (2000b) and Chourdakis and Tzavalis (2000a)]
2. Alternatively, one can assume a parametric form of the prices of risk, for example that $\mathbf{E}[d\mathcal{M}_{ij}] = \varphi_{ij}$, a constant. In addition, one has to augment the representative agent's information with the prices of the options written in the past, that is to say $\mathcal{F}_t = \mathcal{F}_t^{S,\mathbf{x}} \cup \{C(s, \cdot), s \leq t\}$, in order for those extra parameters to be estimated.

3. Finally, one can use historical estimates of the regime impact on the marginal utility of wealth in a general equilibrium framework, and use equations (18) and (19) directly.¹⁸

The risk neutral process can be written in integral form as

$$s(\tau) - s(0) = r\tau + \int_0^\tau \left\{ -\mathbf{x}^T(t) (\bar{\mathbf{Q}} \odot \bar{\mathbf{M}}_\eta) \boldsymbol{\iota} - \frac{1}{2} \mathbf{x}^T(t) [\boldsymbol{\sigma} \odot \boldsymbol{\sigma}] \right\} dt + \int_0^\tau \mathbf{x}^T(t) \boldsymbol{\sigma} d\bar{\varepsilon}(t) + \int_0^\tau [\delta(d\mathbf{x}(t), -\boldsymbol{\iota})]^T \bar{\mathbf{J}} [\delta(d\mathbf{x}(t), \boldsymbol{\iota})], \quad (20)$$

or, again, in the compact form

$$s(\tau) - s(0) = r\tau + \int_0^\tau d\bar{w}(t).$$

The Laplace and consequently the Fourier transforms of the density of the risk neutral process, $s(\tau) - s(0)$, denoted $\bar{f}_{\Delta s(\tau,0)}^{\mathcal{L}}(\theta)$ and $\bar{f}_{\Delta s(\tau,0)}^{\mathcal{F}}(\theta)$, are calculated as

$$\bar{f}_{\Delta s(\tau,0)}^{\mathcal{L}}(\theta) = e^{\theta r \tau} [\mathbf{x}(0)]^T \exp \{ \tau \bar{\mathbf{B}}(\theta) \} \boldsymbol{\iota},$$

where $\bar{\mathbf{B}}(\theta)$ has elements of the form

$$\bar{\beta}_{ij}(\theta) = \begin{cases} \bar{q}_{ii} - \boldsymbol{\iota}^T (\bar{\mathbf{Q}} \odot \bar{\mathbf{M}}_\eta)^T \mathbf{e}_i - \frac{1}{2} \sigma_{ii}^2 \theta^2, & \text{when } i = j \\ \bar{q}_{ij} \bar{g}_{ij}(\theta), & \text{when } i \neq j \end{cases},$$

¹⁸ This approach has not been utilized in the option pricing literature, mainly due to the failure of general equilibrium models to explain excess returns and equity volatilities [see Cochrane (1997)].

and $\bar{f}_{\Delta s(\tau,0)}^{\mathcal{F}}(\theta) = \bar{f}_{\Delta s(\tau,0)}^{\mathcal{L}}(\varpi\theta)$. Recall, that if the price of the regime risk and the price of the jump risk are equal to zero, then

$$\begin{aligned}\bar{\mathbf{Q}} &= \mathbf{Q}, \\ \bar{\mathbf{J}} &= \mathbf{J}, \quad [\implies \bar{\mathbf{M}}_\eta = \mathbf{M}_\eta]\end{aligned}$$

allowing the Laplace and Fourier transforms to be computed from the estimated parameters of equation (2).

The price of a European call option, maturing after time τ , will be given by the expectation of its payoffs, under this risk neutral measure, that is to say

$$C(\tau, K; \boldsymbol{\vartheta}) = e^{-rt} \bar{\mathbf{E}} \left[[S(\tau) - K]^+ | \mathcal{F}_0^{S, \mathbf{x}} \right] = e^{-rt} \bar{\mathbf{E}}_0^{S, \mathbf{x}} [S(\tau) - K]^+, \quad (22)$$

where $\bar{\mathbf{E}}[\cdot]$ denotes the expectation taken under the risk neutral measure, and the vector $\boldsymbol{\vartheta}$ includes all the parameters estimated at time $t = 0$, as well as the initial state $\mathbf{x}(0)$ and price $S(0)$. The above implies the following

Proposition 13 (European call option price) *Consider an asset whose log-price obeys the stochastic differential equation (2). Assume that the price of the state shift risk and the price of the jump risk affect the estimated parameters as in equations (18) and (19). Then, the price of the European call, maturing after time τ with strike price equal to K , will be given by*

$$C(\tau, K; \boldsymbol{\vartheta}) = S(0) \Pi_1(\tau, K; \boldsymbol{\vartheta}) - e^{-r\tau} K \Pi_2(\tau, K; \boldsymbol{\vartheta}), \quad (23)$$

where

$$\Pi_n(\tau, K; \boldsymbol{\vartheta}) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-\varpi\theta \ln[K/S(0)]} \bar{f}_n(\tau, \theta; \boldsymbol{\vartheta})}{\varpi\theta} \right] d\theta, \quad n = 1, 2$$

and the following quantities are defined:

$$f_2(\tau, \theta; \boldsymbol{\vartheta}) = \bar{f}_{\Delta s(\tau, 0)}^{\mathcal{F}}(\theta)$$

$$f_1(\tau, \theta; \boldsymbol{\vartheta}) = \frac{\bar{f}_{\Delta s(\tau, 0)}^{\mathcal{F}}(\theta - \varpi)}{\bar{f}_{\Delta s(\tau, 0)}^{\mathcal{F}}(-\varpi)}$$

the Fourier transform is

$$\bar{f}_{\Delta s(\tau, 0)}^{\mathcal{F}}(\theta) = e^{\varpi \theta r \tau} [\mathbf{x}(0)]^T \exp\{\tau \bar{\mathbf{B}}(\theta)\} \boldsymbol{\iota};$$

and the matrix $\bar{\mathbf{B}}(\theta)$ has elements of the form

$$\bar{\beta}_{ij}(\theta) = \begin{cases} \bar{q}_{ii} - [\mathbf{e}_i^T (\bar{\mathbf{Q}} \odot \bar{\mathbf{M}}_{\eta}) \boldsymbol{\iota} + \frac{1}{2} \sigma_{ii}^2] \varpi \theta - \frac{1}{2} \sigma_{ii}^2 \theta^2, & \text{when } i = j \\ \bar{q}_{ij} \bar{g}_{ij}(\varpi \theta), & \text{when } i \neq j \end{cases}.$$

PROOF: Follows Bakshi and Madan (2000, Case 2).

Corollary 14 (regime shift model, extreme jumps) *If the jumps under risk neutrality are Weibull distributed following $\mathcal{W}(\bar{\alpha}_{ij}, \bar{\gamma}_{ij})$, then the above matrix $\bar{\mathbf{B}}(\theta)$ will be of the form*

$$\bar{\beta}_{ij}(\theta) = \begin{cases} \bar{q}_{ii} - [\mathbf{e}_i^T (\bar{\mathbf{Q}} \odot \bar{\mathbf{M}}_{\eta}) \boldsymbol{\iota} + \frac{1}{2} \sigma_{ii}^2] \varpi \theta - \frac{1}{2} \sigma_{ii}^2 \theta^2, & \text{when } i = j \\ \bar{q}_{ij} \Gamma(1 + \varpi \theta / \bar{\gamma}_{ij}) \bar{\alpha}_{ij}^{\varpi \theta}, & \text{when } i \neq j \end{cases},$$

where the matrix $\bar{\mathbf{M}}_{\eta}$ will have elements $[\bar{\mathbf{M}}_{\eta}]_{ij} = \bar{\alpha}_{ij} \Gamma(1 + \bar{\gamma}_{ij}^{-1}) - 1$, the expected values of the corresponding risk neutral Weibull distribution.

Corollary 15 (pure regime shift model) *In the absence of jumps, the above matrix $\bar{\mathbf{B}}(\theta)$ is a diagonal one of the form $\bar{\mathbf{B}}(\theta) = \bar{\mathbf{Q}} - \frac{\theta^2}{2} \sum_{i=1}^N \sigma_{ii}^2 \mathbf{I}_N^{ii}$,*

that is to say

$$\bar{\beta}_{ij}(\theta) = \begin{cases} \bar{q}_{ii} - \frac{1}{2}\sigma_{ii}^2\varpi\theta - \frac{1}{2}\sigma_{ii}^2\theta^2, & \text{when } i = j \\ \bar{q}_{ij}, & \text{when } i \neq j \end{cases}.$$

PROOF: Both Corollaries follow immediately from the Proposition above.

□

5 An empirical illustration

Weekly S&P500 closing prices, spanning thirty years from 1970 to 2000 are used in order to retrieve estimates of the parameters for a two-state CTMC model with extreme jumps. The parameter vector has ten elements, namely

$$\boldsymbol{\vartheta} = \{\mu_{ii}, \sigma_{ii}, \alpha_{ii}, \gamma_{ii}, q_{ij}\}_{i=1,2, j \neq i}.$$

The estimates are computed using the most naive direct estimation of the characteristic function, as described in section 3.3. The empirical characteristic function is computed as $\hat{f}(\theta) = \frac{1}{M} \sum_{u=0}^M \exp[\varpi\theta\Delta s(t_u)]$ for various values of θ , together with its two standard deviation bounds. The value of θ , say θ^U , is found, at which the empirical characteristic function two standard deviation interval reaches zero, for both its real and imaginary parts. Five equally spaced points are then chosen in the interval $[0, \theta^U]$, and the criterion function (16) is minimized. The estimated values of interest are the following:

$$\begin{aligned} \hat{\boldsymbol{\sigma}} &= (0.20, 0.05)^T, \\ \hat{\mathbf{Q}} &= \begin{pmatrix} -0.90 & 0.90 \\ 0.15 & -0.15 \end{pmatrix}, \\ \mathbf{H} &\sim \mathcal{W}\left(\hat{\alpha} = \begin{pmatrix} 1.03 & - \\ - & 0.97 \end{pmatrix}, \hat{\gamma} = \begin{pmatrix} 30 & - \\ - & 30 \end{pmatrix}\right). \end{aligned}$$

The above parameter values are used in order to compute the moments of the log-price difference, over various horizons. Equation (13) is used in order to retrieve the raw moments, and the ordinary moments are computed afterwards. These are given in figure 9.1. Observe that both skewness and kurtosis exhibit singularities at zero, with their values going to some infinity, whereas both skewness and kurtosis remain high, even for relatively long horizons. The solid line $[-]$ represents the moments at their ergodic [with respect to the CTMC] levels, being in a sense some weighting average of the two conditional moments: the moments conditional on the high volatility state are given by the empty squares $[-\square-]$, the ones conditional on the low volatility state are given by the full squares $[-\blacksquare-]$. Obviously, these extreme cases will serve as bounds for the moments for every initial distribution of the state $\mathbf{x}(0)$.

The calculated moments signify that the specification in (6) can generate higher moments which are consistent with the empirical evidence found in the option pricing literature, namely the negative and persistent skewness and the high degree of kurtosis. The next step is to examine whether or not these features are translated into satisfactory implied volatility [IV] smiles. This exercise will be an indicator of the overall performance of the model, illustrating its potential to encompass the stylized facts. In the first stage the strong assumption that all prices of risk are equal to zero is made, to investigate the shape of the smiles. Option prices are computed using equation (23), and the BS formula is inverted, solving for the volatility. Figure 9.2 displays IV smiles for different maturities: short [one week], medium [one month] and long [six months]. The risk free rate is assumed to be $r = 3\%$; the spot price is considered to be $S(0) = 100$; and different strike prices lie on the horizontal axis. In addition actual IV values are shown, being generated from contracts made in 1997, downloaded from the web page of Yacine Aït-Sahalia.

The figure shows that the performance of the CTMC affine model is satisfactory, especially for short maturity options. The actual IVs lie quite close to the unconditional smile [—], which are computed assuming that $\xi(0) = \xi^\infty$. In addition, the majority of them lie within the upper and lower bounds, which are generated assuming $\xi(0) = (1, 0)^T$ [–□–] and $\xi(0) = (1, 0)^T$ [–■–], respectively.

One can observe that the fit is not perfect when one considers the ergodic case: actual IVs from contracts that are deep-in- or deep-out-of-the-money are higher than the theoretical ones, while long maturity options are generally underpriced. This misspecification could imply that the true parameters are different from the ones utilized, either due to the naive estimation procedure, or because they change across time. On the other hand, there is no reason to assume that during 1997 the true initial distribution of the state is the ergodic one. The fact that the majority of the IVs lie within the bounds suggests that there exists an initial distribution that will further improve the fit and perhaps give unbiased IVs. Nevertheless the pattern is encouraging: both the actual and the theoretical IV smiles seem to flatten as the maturity increases [courtesy of the decreased kurtosis] while the asymmetry remains [courtesy of the persistent skewness].

The above analysis made the assumption that the price of risk is equal to zero, implying a risk neutral probability measure that is a very simple modification of the true one. This assumption is now examined, by investigating how different prices of state-change risk will affect the IV smiles. As illustrated in equation (18), the elements of the rate matrix have to be modified according to the impact of the state change on the diffusion of the percentage changes of the marginal utility of nominal wealth, $d\mathcal{M}$. This impact can be positive or negative, the two extreme cases being illustrated in figure 9.3. The full squares [–■–] resemble the scenario where higher volatility is associated with higher marginal utility of nominal wealth [due to lower *levels* of nominal wealth], in this example it is assumed that given a transition from

the low to the high volatility regime, the marginal utility of nominal wealth will increase by 40%. The empty squares $[-\square-]$ resemble the inverse case, and the solid line where all prices of risk are equal to zero, serves as the benchmark.

The results, as given by the figure, are intuitive. Consider, for example, the case where higher volatility implies [on average] higher marginal utility of nominal wealth, that is to say *given* a transition from state e_2 to e_1 [volatility increase] the expectation $\mathbf{E}[d\mathcal{M}_{21}] > 0$ [or by similar reasoning $\mathbf{E}[d\mathcal{M}_{21}] < 0$]. Equation (18) states that in order to price the option, the representative agent will behave as if the probabilities of moving towards the high volatility state were higher than they really are, resulting into option contracts with higher IVs. The positive relationship between σ and $d\mathcal{M}$ is translated into a negative relationship between the volatility and the nominal wealth: the [risk averse] agent cannot diversify this risk in such an incomplete market, therefore the prices of call options, that offer a hedging instrument, will generally rise [see Chourdakis and Tzavalis (2000b) for a similar argument].

6 Conclusions

This paper introduced a model for the asset returns where a diffusion depends on a state variable in two ways: (i) The diffusion parameters, that is to say the mean and the volatility, depend explicitly on the state of the world, and (ii) When the state of the world changes, a jump is added into the diffusion equation. The paper derived the moment generating function in a closed form. Such a specification has the potential of generating various shapes and levels of conditional higher moments, for different number of states and horizons. In contrast with the previous work on derivative pricing, the states are not restricted in being just two. Closed forms of the Laplace and Fourier transforms of the conditional and unconditional density under mild assumptions about the distribution of jumps were then given. The nec-

essary tools in order for one to efficiently estimate the parameters of such a model were laid down, including the likelihood function and higher moments of any order, generalizing Hamilton (1989) filter for the continuous time case. The paper also showed how derivative pricing can be carried out using the results of Bakshi and Madan (2000), based on the characteristic function of the asset log-price difference. Finally, using an empirical example, it illustrated that this class of models have the potential to capture the stylized facts found in asset pricing literature, and price European option contracts in a satisfactory way.

References

- Abramowitz, Milton, and Irene A. Stegun, 1972, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover Publications Inc.: New York, NY) 9th edn.
- Aït-Sahalia, Yacine, and Andrew Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *The Journal of Finance* 53, 499–547.
- Bakshi, Gudrup, Charles Cao, and Zhiyuan Chen, 1997, Empirical performance of alternative option pricing models, *The Journal of Finance* 5, 2003–2049.
- Bakshi, Gudrup, and Dilip Madan, 2000, Spanning and derivative-security valuation, *Journal of Financial Economics* 55, 205–238.
- Bates, David S., 1995, Testing option pricing models, NBER Working Paper 5129.
- Bertoin, Jean, 1996, *Lévy Processes*. Cambridge Tracts in Mathematics (Cambridge University Press: Cambridge, U.K.).
- Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–659.
- Brennan, Michael J., 1979, The pricing of contingent claims in discrete time models, *The Journal of Finance* pp. 53–68.
- Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark, 1990, Mean reversion in equilibrium asset prices, *American Economic Review* 80, 398–418.
- Chourdakis, Kyriakos M., and Elias Tzavalis, 2000a, Equilibrium option pricing when dividend payments follow a Markov regime switching specification, Working Paper, QM College, Univ. of London.
- , 2000b, Option pricing under discrete shifts in stock returns, Working Paper, QM College, Univ. of London.
- Cochrane, John H., 1997, Where is the market going? uncertain facts and novel theories, *Economic Perspectives* 21, 3–37.

- David, A., and Pietro Veronesi, 1998, Option prices with uncertain fundamentals, mimeo: <http://ideas.uqam.ca>.
- Dollard, John D., and Charles N. Friedman, 1979, *Product Integration with Applications to Differential Equations*. Encyclopedia of Mathematics and its Applications (Addison-Wesley: Reading, MA).
- Duan, Jin-Chuan, 1995, The GARCH option pricing model, *Mathematical Finance* 5, 13–32.
- Feller, William E., 1971, *An Introduction to Probability Theory and its Applications* (John Wiley / Sons: New York, NY) 2nd edn.
- Feuerverger, Andrey, and Philip McDunnough, 1981, On the efficiency of empirical characteristic function procedures, *Journal of the Royal Statistical Society* B43, 20–27.
- Filardo, Andrew J., 1994, Business-cycle phases and their transitional dynamics, *Journal of Business and Economic Statistics* 12, 299–308.
- Gray, S.F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27–62.
- Hamilton, James D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357–384.
- , 1993, State-space models, in Robert F. Engle, and Daniel McFadden, ed.: *Handbook of Econometrics, IV* (North Holland).
- , 1994, *Time Series Analysis* (Princeton University Press: Princeton, NJ).
- , and Rauli Susmel, 1994, Autoregressive conditional heteroscedasticity and changes in regime, *Journal of Econometrics* 64, 307–333.
- Heston, Steven L., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327–344.
- , and Saikat Nandi, 2000, A closed-form GARCH option pricing model, *Review of Financial Studies* Frth, Frth.

- Hull, John, and Alan White, 1987, The pricing of options with stochastic volatilities, *The Journal of Finance* 42, 281–300.
- Kushner, Harold J., 1967, *Stochastic Stability and Control*. Mathematics in Science and Engineering (Academic Press: New York, NY).
- Lanczos, Cornelius, 1964, A precision approximation of the gamma function, *SIAM Journal of Numerical Analysis* B1, 86–96.
- Lo, Andrew W., and Jiang Wang, 1995, Implementing option pricing models when asset returns are predictable, *The Journal of Finance* 50, 87–129.
- Mathias, R., 1997, A chain rule for matrix functions and applications, mimeo: <http://www.math.wm.edu/mathias/>.
- Merton, Robert, 1976, Option pricing when the underlying stock returns are discontinuous, *Journal of Financial Economics* 4, 125–144.
- Naik, Vasant, 1993, Option valuation and hedging strategies with jumps in the volatility of asset returns, *The Journal of Finance* 48, 1969–1984.
- Øksendal, Bernt, 1991, *Stochastic Differential Equations* (Springer-Verlag: New York, NY) 3rd edn.
- Pedler, P. J., 1971, Occupation times for two-state Markov chains, *Journal of Applied Probability* 8, 381–390.
- Schwert, G. William, 1990, Stock volatility and the crash of 87, *Review of Financial Studies* 3, 77–102.
- Susmel, Rauli, 1999, Switching volatility in international equity markets, mimeo: <http://www.cba.uh.edu/rsusmel/>.
- Timmermann, Allan, 2000, Moments of Markov switching models, *Journal of Econometrics* 96, 75–111.
- Turner, Christopher M., Richard Startz, and Charles R. Nelson, 1989, A Markov model of heteroscedasticity, risk and learning in the stock market, *Journal of Financial Economics* 25, 3–22.

7 Proofs of Propositions

PROOF OF PROPOSITION 4: Denote $F_{ij}^w(x, t)$ the distribution function of the cumulative score associated with a transition from state $\mathbf{e}_i$ to $\mathbf{e}_j$ over time t . Let $\varepsilon > 0$, and built the ε -skeleton, $\mathbf{x}_n^{(\varepsilon)}$ of $\mathbf{x}(t)$. Consider a partition of the time interval $[0, \tau]$ into N subintervals of length ε , of the form $[n\varepsilon, (n+1)\varepsilon]$, for $0 \leq n \leq N$. On the boundaries of these subintervals, the CTMC and its skeleton coincide, that is to say $\mathbf{x}(n\varepsilon) = \mathbf{x}_n^{(\varepsilon)}$. Now, define the random variables $w_n^{(\varepsilon)}$, which are distributed with respect to their Laplace transforms given in (7) and (8). The cumulative score of the discrete process will then converge to $\sum_{n=1}^N \rho(n\varepsilon) w_n^{(\varepsilon)} = y^{(\varepsilon)}(N) \rightarrow y(\tau)$, as $\varepsilon \rightarrow 0$ [or $N \rightarrow \infty$]. Consider a time period n , $0 \leq n \leq N$ for the ε -skeleton, and the corresponding time $t = n\varepsilon$ for the CTMC, where the two chains coincide. Denote the conditional mgf $\phi_{ij}^y(\theta, t) = \mathbf{E}[e^{\theta y(t)} | \mathbf{x}(0) = \mathbf{e}_i, \mathbf{x}(t) = \mathbf{e}_j]$. Define the event $\mathcal{E} = \{\mathbf{x}_0^{(\varepsilon)} = \mathbf{e}_j, \mathbf{x}_{n-1}^{(\varepsilon)} = \mathbf{e}_l, \mathbf{x}_n^{(\varepsilon)} = \mathbf{e}_k\}$, and observe that it can be rewritten as $\mathcal{E} = \{\mathbf{x}(0) = \mathbf{e}_j, \mathbf{x}(t - \varepsilon) = \mathbf{e}_l, \mathbf{x}(t) = \mathbf{e}_k\}$. Conditional on the event $\mathcal{E}$, the random variables $y^{(\varepsilon)}(n-1)$ and $\rho(n\varepsilon) w_{n-1}^{(\varepsilon)}$ are mutually independent, and let their moment generating functions (mgf's) to be $\phi_{jl}^y(\theta, t - \varepsilon)$ and $\phi_{lk}^w(\theta, t, \varepsilon)$ respectively, satisfying

$$\phi_{lk}^w(\theta, t, \varepsilon) = \begin{cases} 1 + g_{kk}(\rho(t)\theta)\varepsilon + o(\varepsilon), & \text{when } l = k \\ g_{lk}(\rho(t)\theta), & \text{when } l \neq k \end{cases}.$$

Independency implies that $\mathbf{E}[e^{\theta y(t)} | \mathcal{E}] = \phi_{jl}^y(\theta, t - \varepsilon) \phi_{lk}^w(\theta, t, \varepsilon)$, or using the fact that the n -period transition of the ε -skeleton, from state $\mathbf{e}_i$ to $\mathbf{e}_j$, is given by $p_{ij}(t)$,

$$\begin{aligned} \mathbf{E}[e^{\theta y(t)} | \mathbf{x}(0) = \mathbf{e}_j, \mathbf{x}(t) = \mathbf{e}_k] \\ = \frac{1}{p_{jk}(t)} \sum_l p_{jl}(t - \varepsilon) \phi_{jl}^y(\theta, t - \varepsilon) p_{lk}(\varepsilon) \phi_{lk}^w(\theta, t, \varepsilon). \end{aligned}$$

Rewrite the above as

$$p_{jk}(t + \varepsilon) \phi_{jk}^y(\theta, t + \varepsilon) = \sum_l [p_{jl}(t) \phi_{jl}^y(\theta, t)] [p_{lk}(\varepsilon) \phi_{lk}^w(\theta, t, \varepsilon)],$$

and denote $\tilde{\phi}_{ij}(\theta, t) = p_{ij}(t) \phi_{ij}^y(\theta, t)$. Observe that, as ε gets sufficiently small, one can approximate $p_{lk}(\varepsilon) = \delta(l - k, 0) + q_{lk}\varepsilon + o(\varepsilon)$. Hence the expectation will be written as

$$\begin{aligned} \frac{\tilde{\phi}_{jk}(\theta, t + \varepsilon) - \tilde{\phi}_{jk}(\theta, t)}{\varepsilon} &= \sum_{l \neq k} \tilde{\phi}_{lk}(\theta, t) q_{lk} g_{lk}(\rho(t)\theta) + \\ &\quad + \tilde{\phi}_{jk}(\theta, t) [q_{kk} + g_{kk}(\rho(t)\theta)] \end{aligned}$$

Passing to the limit, $\varepsilon \rightarrow 0$, and collecting the elements $\tilde{\phi}_{jk}(\theta, t)$ in a matrix $\tilde{\Phi}(\theta, t)$, yields $\frac{\partial}{\partial t} \tilde{\Phi}(\theta, t) = \tilde{\Phi}(\theta, t) \mathbf{B}(\theta, t)$, where the matrix $\mathbf{B}(\theta, t)$ has elements of the form

$$\beta_{ij}(\theta, t) = \begin{cases} q_{ii} + g_{ii}(\rho(t)\theta), & \text{when } i = j \\ q_{ij} g_{ij}(\rho(t)\theta), & \text{when } i \neq j \end{cases}.$$

The matrix differential equation resembles a system of N partial differential equations, with trivial boundary conditions $\tilde{\Phi}(\theta, 0) = \mathbf{I}_N$. The solution of those PDEs is of the form $\tilde{\Phi}(\theta, \tau) = \mathfrak{P}_{[0, \tau]} \{ \mathbf{I} + \mathbf{B}(\theta, u) du \}$ [see Dollard and Friedman (1979)]. Bayes' rule implies that the sum of the i -th row of $\tilde{\Phi}(\theta, \tau)$ will be equal to the conditional on the initial state mgf, denoted $\phi_i^y(\theta, \tau) = \mathbf{E}[e^{\theta y(\tau)} | x(0) = i]$ and hence the vector $\phi(\theta, \tau) = \tilde{\Phi}(\theta, \tau) \boldsymbol{\nu}$. The mgf conditional on the information set at time 0 will be equal to $\phi^y(\theta, \tau) = [\mathbf{x}(0)]^T \phi(\theta, \tau) = [\mathbf{x}(0)]^T \tilde{\Phi}(\theta, \tau) \boldsymbol{\nu}$. $\square$ PROOF OF COROLLARY 9: The mgf of the innovation $dw(t)$ is of the form $\mathbf{E}[e^{\theta dw(t)} | \mathbf{x}(t) = \mathbf{e}_i] = 1 + \theta dt + o(dt)$, and $\mathbf{E}[e^{\theta dw(t)} | \mathbf{x}(t) \neq \mathbf{e}_i] = 1 + O(dt)$. The matrix $\mathbf{B}(\theta, u)$ will give

$$\begin{aligned} \int_0^\tau \mathbf{B}(\theta, u) du &= \int_0^\tau \begin{pmatrix} q_{11} & q_{1i} & q_{1N} \\ q_{21} & q_{2i} & q_{2N} \\ & \ddots & \vdots \\ q_{i1} & \cdots & q_{ii} - \theta & \cdots & q_{iN} \\ & & \vdots & \ddots & \\ q_{N1} & q_{Ni} & q_{NN} \end{pmatrix} du \\ &= \tau \mathbf{Q} + \tau \theta \mathbf{I}_N^{ii}, \end{aligned}$$

where $\mathbf{I}_N^{ij}$ is an $(N \times N)$ matrix of zeros, with only the (i, j) -th element equal to one. Consequently, $\phi^y(\theta, \tau) = [\mathbf{x}(0)]^T \exp\{\tau [\mathbf{Q} + \theta \mathbf{I}_N^{ii}]\} \boldsymbol{\nu}$. If one want to

evaluate the first two moments of $y(\tau)$, the exponential of the BUTB Toeplitz matrix needs to be constructed:

$$\mathbf{A} = \exp \left\{ \begin{pmatrix} \tau \mathbf{Q} & \tau \mathbf{I}_N^{ii} & \mathbf{0} \\ \mathbf{0} & \tau \mathbf{Q} & \tau \mathbf{I}_N^{ii} \\ \mathbf{0} & \mathbf{0} & \tau \mathbf{Q} \end{pmatrix} \right\}. \quad \square$$

PROOF OF COROLLARY 10: The mgf of $dw(t)$ is of the form

$$\begin{aligned} \mathbf{E} [e^{\theta dw(t)} | \mathbf{x}(t) = \mathbf{e}_i] &= 1 + \left(\theta \mu_i + \frac{1}{2} \theta^2 \sigma_i^2 \right) dt + o(dt), \text{ and} \\ \mathbf{E} [e^{\theta dw(t)} | \mathbf{x}(t) \neq \mathbf{e}_i] &= 1 + O(dt), \end{aligned}$$

which depend on the time. Therefore, the matrix

$$\int_0^\tau \mathbf{B}(\theta, u) du = \tau \mathbf{Q} + \tau \sum_{i=1}^N (\theta \mu_i + \theta^2 \sigma_i^2) \mathbf{I}_N^{ii}.$$

The derivatives are given by $\mathbf{D}^{(0)} = \mathbf{Q}$, $\mathbf{D}^{(1)} = \sum_{i=1}^N \mu_i \mathbf{I}_N^{ii}$, $\mathbf{D}^{(2)} = \sum_{i=1}^N \sigma_i^2 \mathbf{I}_N^{ii}$, and $\mathbf{D}^{(k)} = \mathbf{0}$, $\forall k > 2$. The k -th moment will require the following $(k \times k)$ BUTB Toeplitz matrix exponential given in Corollary 10. $\square$ PROOF OF COROLLARY 11: The mgf of the innovation $dw(t)$ is of the form

$$\begin{aligned} \mathbf{E} [e^{\theta dw(t)} | \mathbf{x}(t) = \mathbf{x}(t+dt) = \mathbf{e}_i] \\ = 1 + \left(\theta \mu_i + \frac{1}{2} \theta^2 \sigma_i^2 \right) dt + o(dt), \text{ and} \end{aligned}$$

$$\begin{aligned} \mathbf{E} [e^{\theta dw(t)} | \mathbf{x}(t) = \mathbf{e}_i, \mathbf{x}(t+dt) = \mathbf{e}_j \neq \mathbf{e}_i] \\ = \Gamma(1 - \beta_{ij}\theta) \exp[\alpha_{ij}\theta] + O(dt), \end{aligned}$$

which depend on the time. The elements of the matrix $\frac{d^k}{d\theta^k} \mathbf{B}(\theta, u) \Big|_{\theta=0}$, $\beta_{ij}^{(k)}$, are $\beta_{ii}^{(k)} = \frac{d^k}{d\theta^k} (q_{ii} + \theta \mu_i + \frac{1}{2} \theta^2 \sigma_i^2) \Big|_{\theta=0}$, for the diagonal elements; and $\beta_{ij}^{(k)} = \frac{d^k}{d\theta^k} q_{ij} \Gamma(1 - \beta_{ij}\theta) \exp[\alpha_{ij}\theta] \Big|_{\theta=0}$, for the off-diagonal elements, which are given in the technical appendix. The above elements will be used to construct the BUTB Toeplitz matrix exponential of corollary 11 $\square$

8 Technical appendix

8.1 Matrix exponential

The matrix exponential is calculated using a (6,6)-Padé rational approximation scheme. The exponential is approximated with a ratio of two polynomials each of degree six:

$$\exp \mathbf{A} = \left[\sum_{i=0}^6 \alpha_i \mathbf{A}^i \right] \times \left[\sum_{i=0}^6 \beta_i \mathbf{A}^i \right]^{-1} + \epsilon.$$

The above scheme results into errors of order $\log \|\epsilon\|_1 < -20$ if $\|\mathbf{A}\|_1 < 1/2$, with $\|\cdot\|_1$ denoting the L_1 or the maximum column sum norm. Therefore the matrix $\mathbf{A}$ is rescaled as $\mathbf{A} = 2^n \times \mathbf{A}^*$, with $\|\mathbf{A}^*\|_1 < 1/2$, and the exponential is calculated using the identity $\exp [2^n \mathbf{A}^*] = [\exp \mathbf{A}^*]^{2^n}$ which needs only n matrix multiplications to be calculated. The Padé coefficients are

i	α_i	β_i
1	1	1
2	1/2	-1/2
3	1/66	1/66
4	1/792	-1/792
5	1/15840	1/15840
6	1/665280	-1/665280

8.2 Gamma function for complex arguments

The Gamma function is approximated using the Lanczos (1964) method, which states that

$$\Gamma(1+z) = \sqrt{2\pi} \left(z + \frac{11}{2} \right)^{z+\frac{1}{2}} e^{-(z+\frac{11}{2})} \left[\alpha_0 + \sum_{i=1}^6 \frac{\alpha_i}{z+i} + \epsilon \right],$$

which ensures that the error $\log |\epsilon/2| < -10$, if $\text{Re } z > 0$. The Lanczos coefficients are

i	α_i	scale
0	1.000000000190015	
1	.7618009172947146	$\times 10^2$
2	-.8650532032941677	$\times 10^2$
3	.2401409824083091	$\times 10^2$
4	-.1231739572450155	$\times 10$
5	.1208650973866179	$\times 10^{-2}$
6	-.5395239384953	$\times 10^{-5}$

8.3 Derivatives of $\Gamma(x)$ and related functions

The values needed in this paper, in order to compute the BUTB Toeplitz matrix necessary for the moment computation under extreme jumps, are a result of the differentiation

$$\Psi_n(q, \alpha, \gamma) = \left. \frac{d^n}{d\theta^n} q \Gamma(1 + \theta/\gamma) \alpha^\theta \right|_{\theta=0}.$$

These derivatives are calculated analytically, the results being of the form

$$\Psi_n(q, \alpha, \gamma) = q \ln \alpha + q \sum_{k=1}^n \binom{n}{k} \frac{\psi^{(k)}(1)}{\gamma^k} (\ln \alpha)^{n-k},$$

with the function $\psi^{(k)}(1)$ being computed analytically in terms of various mathematical constants¹⁹, for every k . The first ten values are given in the

¹⁹ For example, $\psi^{(1)}(1) = -\gamma$, $\psi^{(2)}(1) = \frac{1}{6}\pi^2 + \gamma^2$, $\psi^{(3)}(1) = -\zeta(3) - \frac{1}{2}\pi^2\gamma - \gamma^3$, etc. γ here denotes Euler's constant and $\zeta(\cdot)$ is Riemann's Zeta function. Do not confuse the function $\psi^{(n)}$ with the polygamma function, which gives the derivative of the logarithm of the Gamma function.

following table

k	$\psi^{(k)}(1)$	scale
1	−.57721566490153286061	
2	.19781119906559451108	$\times 10$
3	−.54448744564853177343	$\times 10$
4	.23561474084025604497	$\times 10^2$
5	−.11783940826837742427	$\times 10^3$
6	.71506736252731885912	$\times 10^3$
7	−.50198488726298549314	$\times 10^4$
8	.40243621573335758137	$\times 10^5$
9	−.36252628911465498246	$\times 10^6$
10	.36270424127568947788	$\times 10^7$

9 Figures

9.1 Moments

The model utilized in order to generate these graphs is a two-state ctmc driven model with extreme jumps. The moments are computed following (10) and the results are plotted across time. The two states $\{\mathbf{e}_1, \mathbf{e}_2\}$ have the following characteristics: (1) the instantaneous volatility parameters are $\sigma = (19\%, 6\%)^T$; (2) the off-diagonal elements of the rate matrix $(q_{12}, q_{21})^T = (0.90, 0.15)^T$; (3) the parameters of the Weibull distribution which determines the magnitude of the jumps are $(\alpha_{12}, \gamma_{12})^T = (1.03, 30)^T$, and $(\alpha_{21}, \gamma_{21})^T = (0.98, 30)^T$; and (4) the interest rate is considered constant at 3%. The solid line $[-]$ represents the unconditional moments with initial state $\xi_{0|0} = \xi^\infty$; the empty squares $[-\square-]$ represent conditional moments with $\xi_{0|0} = (1, 0)^T$; and the full squares $[-\blacksquare-]$ represent conditional moments with $\xi_{0|0} = (0, 1)^T$. Note that the skewness and kurtosis exhibit singularities as the time interval shrinks.

9.1.1 Volatility

9.1.2 Skewness

9.1.3 Kurtosis

9.2 Implied volatility smiles

The model utilized in order to generate these graphs is a two-state ctmc driven model with extreme jumps. Option prices are computed following (23) and the results are inverted in order to retrieve the implied volatilities. The two states $\{\mathbf{e}_1, \mathbf{e}_2\}$ have the following characteristics: (1) the instantaneous volatility parameters are $\sigma = (19\%, 6\%)^T$; (2) the off-diagonal elements of the rate matrix $(q_{12}, q_{21})^T = (0.90, 0.15)^T$; (3) the parameters of the Weibull distribution which determines the magnitude of the jumps are $(\alpha_{12}, \gamma_{12})^T = (1.03, 30)^T$, and $(\alpha_{21}, \gamma_{21})^T = (0.98, 30)^T$; and (4) the interest rate is considered constant at 3%. The solid line [—] represents the unconditional ivs with initial state $\xi_{0|0} = \xi^\infty$; the empty squares [□] represent ivs with $\xi_{0|0} = (1, 0)^T$; and the full squares [■] represent ivs with $\xi_{0|0} = (0, 1)^T$. The figures also show realized implied volatilities for the corresponding maturities. The actual data are the same as in Aït-Sahalia and Lo (1998), available at the homepage of Yacine Aït-Sahalia.

9.2.1 One week

9.2.2 One month

9.2.3 Six months

9.3 Prices of risk

The figures below illustrate the impact of the price of regime change risk on option prices. The solid line [—] represents the unconditional IVs, based on the parameters used for the previous figures. The line marked with empty squares [□] represents IVs when the state change and the marginal utility of wealth are negatively correlated²⁰: $\mathbf{E}[d\mathcal{M}_{12}] = -40\%$, and $\mathbf{E}[d\mathcal{M}_{21}] = 40\%$. The line marked with the full squares [■] represents the inverse case, where $\mathbf{E}[d\mathcal{M}_{12}] = 40\%$, and $\mathbf{E}[d\mathcal{M}_{21}] = -40\%$. The magnitude of the jump itself is considered to have no impact on the marginal utility of nominal wealth, i.e. $\mathbf{Cov}[J_{12}, d\mathcal{M}_{12}] = \mathbf{Cov}[J_{21}, d\mathcal{M}_{21}] = 0$.

9.3.1 One week

²⁰ For example, a switch from the low volatility state to the high volatility state is associated with a *lower* level of nominal wealth, or equivalently with a *higher* marginal utility of nominal wealth.

9.3.2 One month

9.3.3 Six months

**This working paper has been produced by
the Department of Economics at
Queen Mary, University of London**

**Copyright © 2000 Kyriakos M. Chourdakis
All rights reserved.**

**Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580
Email: j.conner@qmw.ac.uk
Website: www.econ.qmw.ac.uk/papers/wp.htm**