A New Class of Asset Pricing Models with Lévy Processes: Theory and Applications

Chayawat Ornthanalai*

Desautels Faculty of Management, McGill University

December 20th, 2008

Abstract

We develop a new class of discrete-time asset pricing models with Lévy processes and use affine GARCH dynamics to drive the models' time variation. These models are easy to implement and can capture three important stylized facts of asset returns, which are non-normality, time-varying return volatility, and the leverage effect. In addition, this framework yields asset return dynamics that have an affine structure in their conditional transform, which leads to simple valuation of various derivatives including zero-coupon bonds and European options. We apply this newly proposed framework to various two-factor models consisting of a normal and a pure jump Lévy component. The results from joint estimation of options and returns on the market index reveal the important economic role of jumps. Models without jumps cannot reconcile the difference between market-realized returns and investors' ex-ante expectations of returns with an economically justifiable equity premium level. We find that investors demand 3 to 5% in annual excess return for bearing the market jump risk.

JEL Classification: G10; G12; G13; C51.

Keywords: Lévy process; time change; discrete-time; GARCH; affine transform; option valuation; filtering; risk premia.

*I would like to thank my co-supervisors Peter Christoffersen and Kris Jacobs for encouragement and support. This paper has benefited from the comments by Darrell Duffie, David Bates, Louis Ederington, Olaf Korn, and seminar participants at the Georgia Institute of Technology, HEC Montreal, Bank of Canada, Oxford-Man Conference, and the 2008 FMA meeting. I would like to acknowledge CIREQ and the J.W. McConnell foundation for financial support. Any remaining inadequacies are mine alone. Correspondence to: Chayawat Ornthanalai, Desautels Faculty of Management, McGill University, 1001 Sherbrooke Street West, Montreal, Canada H3A 1G5; Tel: (514) 398-4000 ext 00996; Fax: (514) 398-3876; E-mail: chayawat.ornthanalai@mail.mcgill.ca; Website: www.ornthanalai.com.

1 Introduction

This paper introduces a rich class of discrete-time asset pricing models that combine the flexibility of Lévy processes with the ease of implementation of GARCH dynamics. Our models can capture three important stylized facts of asset returns, namely the presence of jumps, time-varying dynamic of return volatility, and the leverage effect. In addition, this framework produces a large class of asset return processes that have analytical solutions for their conditional transform. The price of zero-coupon bonds is available in closed-form, and the prices of European-style derivatives can be computed using the Fourier inversion method as discussed in Heston (1993). We refer to this newly developed class of models as the Lévy GARCH. The risk neutralization of asset returns as well as the statistical method for estimating the models are also discussed in detail. We demonstrate the versatility of this newly proposed framework by estimating various two-factor return models consisting of a normal and a pure jump Lévy component. We conduct two estimation exercises using options and returns data on the S&P 500 index. The first is based on a time-series of daily returns only, while the second uses options and returns data jointly. The results from both estimation exercises confirm that infinite-activity jumps outperform the finite-activity Merton jump structure in fitting returns as well as pricing options. We use a large panel data set of options in the second estimation exercise and find that the normal and jump risks are both priced in the market.[1] More importantly, we show that models without the jump risk factor cannot jointly fit the returns and options data with an economically reasonable level of the equity premium.

Asset pricing ideally involves not only the statistical modeling of returns, but also the search for an equivalent return process under the risk-neutral measure that can be used to price derivatives. The task of a financial economist is therefore further complicated by the need to find models that can explain the behavior of asset prices under two different probability measures. A good model must, in addition, be economically supported. The divergence between the two probability measures has to be linked by an appropriate vehicle and with a return premium that is justifiable based on standard theories of risk-return trade-off.

In an attempt to model the nature of asset prices, rich and sophisticated models such as stochastic volatility, GARCH, and affine-jump diffusion models (AJD) have been developed and extensively studied.[2] Jump models have become increasingly important in the literature, and AJD models which build upon on the finite-activity compound Poisson process have been gradually adopted as the benchmark for modeling index returns.[3] However, Carr and Wu (2003b) document the existence of many small jumps that cannot be adequately modeled

[1]This is consistent with Pan (2002) who documents a large jump risk premium from time-series of short-term and near-the-money index options. In addition, we find that the normal risk factor is priced.

[2]See Hull and White (1989) and Heston (1993) for stochastic volatility models. For seminal GARCH studies, see Engle (1982) and Bollerslev (1986). For the theoretical framework of asset pricing under affine jump-diffusions, see Duffie, Pan and Singleton (2000).

[3]A short list of these studies includes Andersen, Benzoni and Lund (2002), Chernov, Gallant, Ghysels and Tauchen (2003), Eraker, Johannes and Polson (2003), Eraker (2004), and Bates (2006).

using finite-activity compound Poisson processes. In response, a new strand of literature has developed which considers more general jump structures, including infinite-activity jumps.[4] These processes, along with the Brownian motion and compound Poisson processes, belong to a larger class of stochastic processes named Lévy processes. Due to the wide range of statistical processes that fall within this classification, recent research in asset pricing has been geared towards the application of Lévy processes to modeling asset returns and pricing derivatives. This recent research includes the work of Carr, Geman, Madan and Yor (2003) and Carr and Wu (2004), who combine Lévy processes with a subordinated stochastic time change. The ensuing processes are referred to as time-changed Lévy processes. Carr and Wu (2004) show that their framework encompasses almost all of the models proposed in the option pricing literature.

The Lévy GARCH models introduced in this paper can be thought of as discrete-time counterparts of the continuous time models in Carr and Wu (2004). However, our approach for producing the equivalent mechanism of the random time change effect differs from the one used in the existing literature (see also Huang and Wu (2004), and Bakshi, Carr and Wu (2007)). Instead of evaluating Lévy processes at stochastic time points, we introduce time-varying dynamics by relying on heteroskedasticity in the parameters that govern these processes. We also expand on Carr and Wu (2004) by providing a tractable framework for the risk neutralization of asset returns, thereby enabling joint studies of their dynamics under the physical and risk-neutral measures.

We use affine GARCH dynamics to drive the time-varying dynamic for three main reasons. The first reason is the tractability of the pricing formulae. We provide a general solution to the conditional transform of asset returns that can be drawn from various combinations of Lévy processes and affine GARCH dynamics. Our result, in fact, nests all the existing affine GARCH models used in the option pricing literature.[5] The second reason is the leverage effect that is automatically built into the model. Because most affine GARCH dynamics permit asymmetric response between asset returns and their volatilities, the leverage effect can be easily incorporated into our models. Carr and Wu (2004) show that when introducing the leverage effect in their models, the conditional transform of asset returns can only be solved analytically using the well-established methods in the literature after applying a complex measure change technique. This is not the case for the Lévy GARCH models as we show that the conditional transform of asset returns can be directly solved using simple iterated expectations. Our third reason for using affine GARCH dynamics is ease of implementation. GARCH models are powerful filters that are extensively used by finance academics and practitioners.

[4]See for example the Inverse Gaussian and the Normal Inverse Gaussian models of Barndorff-Nielsen (1998), the variance-gamma model of Madan and Milne (1991), and the CGMY process of Carr, Geman, Madan and Yor (2002).

[5]This includes the affine GARCH dynamic of Heston and Nandi (2000), the Inverse Gaussian GARCH model of Christoffersen, Heston and Jacobs (2006), and the component GARCH model of Christoffersen, Jacobs, Ornthanalai and Wang (2008).

However, it must be noted that these affine GARCH dynamics differ in an important way from the standard models used in the literature. As the total return residuals are not always used in the GARCH updating process, the models do not suffer from the problem of excessively large estimates of conditional variance after large stock market movements.

The estimation of some Lévy GARCH processes may require additional filtering techniques due to the presence of latent state variables. We show that this can be easily accommodated using the particle filter (see Pitt (2002)), and that our models can be estimated using maximum likelihood. Our estimation procedure is relatively quick, and much more straightforward than the MCMC technique used in Li, Wells and Yu (2006). In addition, our maximum likelihood estimation does not require that the density of the Lévy process be known analytically, as long as these processes can be simulated using a robust algorithm. This is clearly advantageous as most of the Lévy processes do not have a closed-form density function.

Methodologically, the work presented in this paper contributes to the field of asset pricing in several ways. The first of these is the development of the Lévy GARCH framework. Our setup allows for a wide variety of asset return specifications by combining Lévy processes with affine GARCH dynamics. Our second contribution is a risk neutralization framework that is economically appealing and analytically tractable. We assume an affine structure for the equity premium which greatly facilitates the identification of the risk premia that each Lévy shocks implies on the expected excess return. In addition, the conditional equity premium dynamic can accommodate both nonlinearity and time variation as documented by Dai and Singleton (2002), and Bakshi, Carr and Wu (2008). The third contribution is the application of our framework to nonaffine GARCH dynamics. This is of paramount importance because most of the GARCH processes studied empirically are nonaffine, and hence do not admit analytical pricing transforms. We show that most of the Lévy processes considered in this paper, once transformed into the risk-neutral measure, stem from recognizable distributions. The risk-neutral dynamics of asset prices can therefore be simulated, and derivatives can be priced via Monte Carlo simulation as in Duan (1995).

This paper also contributes to the field of empirical asset pricing. We estimate and test our models on S&P 500 index options and returns using two different approaches. The first involves MLE estimation on daily returns from 1985 to 2005. To our knowledge, Bates (2008) is the only study that estimates time-changed Lévy processes of infinite-activity using only return data. Previous empirical estimations of time-changed Lévy processes such as Huang and Wu (2004) have relied solely on option prices. Li, Wells and Yu (2006) use MCMC to estimate models where returns follow a stochastic volatility process plus a Lévy jump. However, they do not estimate the case of time-changed Lévy jumps. The lack of time series estimates of time-changed Lévy processes is probably due to the inherent econometric complexity. Our discrete-time framework can therefore partially alleviate the econometric challenges that have hindered empirical research in this area.

The second estimation approach involves joint MLE estimation on weekly call options and daily returns data. We use a large data set of options and returns (1996-2005). To our knowledge, this is the most extensive joint estimation exercise ever conducted in the option pricing literature.[6] The use of an extensive data set in joint estimation allows us to precisely estimate the long-run factor risk premia in our models. In turn, this helps us answer the most important economic issue in empirical option pricing: which risk factors are priced in the market, and by how much?

We estimate four models where returns are driven by shocks from the normal and pure-jump Lévy component. Two choices of jump processes are studied. The first is the finite-activity Merton jump, and we refer to models associated with this jump as MJ-LGARCH. The second is the infinite-activity Normal Inverse Gaussian (NIG) jump. We refer to models with NIG jumps as NIG-LGARCH. In addition to two different jump types, we model each jump under two different dynamics: one which has a homoskedastic jump process, and one which has a heteroskedastic (i.e., time-changed) jump process.

Using these two different estimation approaches, we are able to draw a number of important conclusions. **(1)** Infinite-activity jumps are preferable to the standard finite-activity Merton jump structure. This statement holds true from the perspective of returns fitting as well as options pricing. **(2)** When fitting the time series of returns, and in the presence of a time-varying normal component, it is not necessary to model the infinite-activity NIG jumps as time-varying. On the other hand, it is important to have time-varying dynamics for the finite-activity Merton jump process. **(3)** The presence of jumps in option pricing models is economically important. Without the jump component, the divergence between the physical and risk-neutral measures cannot be explained using an economically justifiable level of equity premium. Our results from joint MLE show that the equity premium level implied by a model without jumps is 23% per annum. On the other hand, for models with jumps, the implied equity premium levels are about 8% and 6.3% for the MJ-LGARCH, and NIG-LGARCH respectively. **(4)** Both jump and normal risk factors are priced in the market. For the MJ-LGARCH, we find that investors demand 3% and 5% in excess annual returns for bearing the market jump and diffusive normal risks respectively. Similarly for the NIG-LGARCH, the jump risk is priced at 5% per year, while the diffusive normal risk is priced at about 1.3% per year.

The rest of this paper proceeds as follows. Section 2 discusses the construction of time-changed Lévy processes in discrete time. Section 3 introduces the model and presents the closed-form generating function for asset prices. In section 4, the risk neutralization procedure is discussed. Section 5 discusses the methodology used to estimate the Lévy GARCH models.

[6] Existing studies on the joint estimation use data covering short time periods or a small subset of the cross section of options. Chernov and Ghysels (2000) use approximately one option per day (at-the-money and shortest to maturity). Pan (2002) uses two options per day (at-the-money and short-term) from 1989 to 1996. Eraker (2004) uses up to three options per day from 1987 to 1990.

Section 6 presents the empirics, and finally Section 7 concludes.

2 Building discrete-time models based on time-changed Lévy innovations

2.1 Lévy processes

Lévy processes have been the driving force behind most asset pricing models. In order to appreciate the richness of this class of stochastic processes, consider a sample path that is right continuous with left limit of $\{X_t \in \mathbb{R} | t \geq 0\}$ and $X_0 = 0$. Under the usual probability space and filtration, if the increments $X_{t+s} - X_t$ for $s > 0$ have a stationary and independent distribution, we say that X_t has infinitely divisible distribution and is a Lévy process. This definition of a Lévy process encompasses most of the past and existing distributions used in the finance and economics literature. It is convenient to work with the generalized Fourier transform of a Lévy process because the density is not always available analytically. The transform is given by

$$F_x(\phi) = E\left[e^{\phi X_t}\right] = e^{t\Psi_x(\phi)}, \qquad \phi \in \mathcal{D}^x \subset \mathbb{C} \tag{2.1}$$

where ϕ is in the complex domain $\mathcal{D}^x \subset \mathbb{C}$ such that (2.1) is well-defined. Following Wu (2006) and Bates (2008), we will refer to $\Psi_x(\phi)$ as the cumulant exponent of X_t. Note that the characteristic function and the moment generating function are special cases of the generalized Fourier transform. For generality, we use the generalized Fourier transform in this paper by specifying ϕ in the domain $\mathcal{D}^x$ of the complex plane, and for brevity, we refer to $F_x(\phi)$ simply as the conditional transform in most parts of this paper.

The log of the generalized Fourier transform of a Lévy process X_t is linear in time, with the slope being its cumulant exponent. This fact follows from a well-known property of infinitely divisible distributions. The Lévy-Khintchine theorem states that any Lévy processes can be decomposed into a constant drift, a Brownian part, and a pure jump part. Using one version of the Lévy-Khintchine formula, we can write the cumulant exponent of a Lévy process as

$$\Psi_x(\phi) = \mu\phi + \frac{1}{2}\sigma^2\phi + \int_{\mathbb{R}_0} \left(e^{\phi x} - 1 - \phi x \mathbf{1}_{\{|x|<1\}}\right) \nu(dx). \tag{2.2}$$

The measure $\nu(dx)$ is called the Lévy measure defined on $\mathbb{R}_0$ which dictates how jumps occur.[7] The term μ represents the constant drift and σ^2 is the variance of the Brownian part. In the case of a pure jump process, there is no Brownian part, and hence $\sigma^2 = 0$. The triplet

[7] Jumps of size x in the set $\mathbb{R}_0$ (real line excluding zero) occur according to a Poisson process with intensity parameter $\int_{\mathbb{R}_0} \nu(dx)$.

$[\mu, \sigma^2, \nu(dx)]$ is often referred to as the Lévy characteristics, and completely determines the properties of the Lévy process.

2.2 Time-changed Lévy processes

Carr and Wu (2004) develop option pricing models based on time-changed Lévy process in order to address three stylized facts of asset prices. The first is that asset prices jump, leading to nonnormal return innovations. This stylized fact can be trivially captured through an appropriate choice from a plethora of Lévy processes. The second stylized fact is that return volatility is stochastic. Because Lévy processes are infinitely divisible, stochastic volatilities can be generated by evaluating the sample path of a Lévy process at a random time. Carr and Wu (2004) apply this technique to option pricing and refer to the resulting models as time-changed Lévy processes. Specifically, they model the activity rate of the stochastic time change using well-established techniques from the affine term structure literature. This results in great tractability of option pricing formula under several rich specifications for the stock price. The generalized Fourier transform for time-changed Lévy processes $\left\{\mathbf{X}_t \in \mathbb{R}^d \mid t \geq 0\right\}$ with $\mathbf{X}_0 = \mathbf{0}$ involves taking expectations over two sources of randomness,

$$F_{\mathbf{X}_t}(\Phi) = E\left[e^{\Phi' \mathbf{X}_{T_t}}\right] = E\left[E\left[e^{\Phi' \mathbf{X}_s} \mid T_t = s\right]\right] \tag{2.3}$$

The inner expectation is taken conditional on $\mathbf{X}_{T_t}$ fixed at $T_t = s$, and its solution is given by (2.2). The outer expectation then operates on all the possible values of stochastic time T_t, and its solution is taken from the bond pricing literature.

The third stylized fact in asset prices is that returns and their volatilities are correlated. This is commonly referred to as the leverage effect which simply demands that $\mathbf{X}_{T_t}$ be correlated with its Lévy subordinator. Unfortunately, when introducing correlation between stochastic time change and its Lévy innovation, the expectation in (2.3) cannot be solved through simple iterated expectations. To overcome this problem, Carr and Wu (2004) introduce an ingenious method of taking the expectation in (2.3) under a complex measure change. This new measure is free of the leverage effect, as the correlation between the time change and Lévy innovation is absorbed into the measure change. We refer interested readers to Carr and Wu (2004) for the details of this procedure.

2.3 Modeling time-changed Lévy processes in discrete-time

In discrete time, it is quite counter-intuitive to think of time units as being random. Moreover, data on asset prices are often recorded at fixed frequency. Consider a simple case of daily log returns with $\ln(S_{t+1}/S_t) = X_{t+1}$. The one-day conditional discrete-time version of the generalized transform is given by

$$F_x(\phi;t;t+1) = E\left[e^{\phi X_{t+1}} \middle| \, t\right] = e^{\Psi_x(\phi;t,t+1)}, \quad \phi \in \mathcal{D}^x \subset \mathbb{C} \tag{2.4}$$

where $\Psi_x(\phi;t,t+1)$ is the one-day conditional version of the cumulant exponent of X_{t+1}. It follows that its conditional n-day ahead cumulant exponent is given by n summations of $\Psi_x(\phi;t,t+1)$. Because the exponent of the conditional transform in (2.4) is no longer linear in time as in (2.1), the applicability of random time change is not so obvious.

In order to produce the equivalent effect that random time change has on Lévy innovations, we make use of an observation that almost all Lévy processes are time homogeneous in one of their parameters. This property is defined as follows.

Definition 1 ***The property of time homogeneity.*** *Consider a Lévy process $x_{t+1} \in \mathbb{R}$ with distributional parameters Θ. If there exists a non-empty set $h_{t+1} \subset \Theta$ such that the conditional cumulant exponent of x_{t+1} over the time interval $(t, t+1)$ is given by*

$$\Psi_x(\phi;t,t+1) = h_{t+1}\xi_x(\phi), \tag{2.5}$$

with $\xi_x(\phi)$ independent of h_{t+1}, then we say that X is time homogeneous in the parameter h_{t+1}, and $\xi_x(\phi)$ is the coefficient in the cumulant exponent.

The above definition can be extended to d-dimensional Lévy processes $\mathbf{X}_{t+1} \in \mathbb{R}^d$. In this case, we will have $\Psi_{\mathbf{X}}(\Phi;t,t+1) = \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Phi})' \, \mathbf{h}_{t+1}$ and it is easy to see that the conditional cumulant exponent of $\mathbf{X}_{t+1}$ is affine in the time-homogeneous parameters $\mathbf{h}_{t+1}$. A well-known example of a time-homogeneous Lévy process is the zero-mean normal distribution, or Brownian motion. Consider daily log returns that distributed according to $N(0,\sigma^2)$. The generalized Fourier transform is given by $\frac{1}{2}\phi^2\sigma^2$, and hence is time homogeneous in σ^2. This property implies that annual log returns (365 days) will be distributed as $Normal(0,\ 365\sigma^2)$. Under the decomposition in (2.5), the cumulant exponent of Lévy process X is now linear in h_{t+1}. We therefore see that this time-homogeneous parameter is an ideal candidate for dynamic modeling in order to mimic the effect of random time change on Lévy innovations. Thus, our approach for producing random time change effect is through heteroskedastic specifications of $\mathbf{h}_{t+1}$.

In order to model the leverage effect, which is one of the three stylized facts of asset prices, we must allow for $\mathbf{h}_{t+1}$ to be correlated with the return process. Equally important is that we want to choose specifications of $\mathbf{h}_{t+1}$ such that the generalized Fourier transform for asset returns is analytically tractable. Given this, we propose to model h_{t+1} with affine GARCH dynamics. We denote this conditional dynamic as $\mathbf{h}_{t+1} = \mathcal{G}(\mathbf{h}_t, \boldsymbol{v}_t | \, \Omega_t)$, where Ω_t is the conditioning information set from time 0 to t. We let $\boldsymbol{v}_t$ represent a vector of state variables (other than $\mathbf{h}_t$), which must be predictable one period prior to t.

The generalized Fourier transform for Lévy processes $\mathbf{X}_T$ conditional on time $t < T$ is found through a series of iterated expectations

$$\begin{aligned} F_{\mathbf{X}}(\Phi; t;\ T) &= E_t\left[E_{t+1}\left[\ldots E_{T-2}\left[E_{T-1}\left[e^{\Phi'\mathbf{X}_T}\right]\right]\ldots\right]\right] \\ &= e^{\Psi_{\mathbf{X}}(\Phi;\ t,\ T)}, \end{aligned}$$

where $E_s[\,]$ denotes the expectation taken with respect to the conditioning information set Ω_s. Accordingly, the cumulant exponent $\Psi_{\mathbf{X}}(\Phi;\ t,\ T)$ will follow a recursive updating rule that depends on the dynamic in $\mathcal{G}(\mathbf{h}_t, \boldsymbol{\upsilon}_t |\ \Omega_t)$ and the choice of Lévy innovations. As usual, we will assume the exponential affine form in asset prices using d-dimensional Lévy processes $\mathbf{X}_{t+1}$. The GARCH filtration of $\mathbf{h}_{t+1}$ is then driven by at least one conditioning innovation in $\mathbf{X}_t$. We advocate the use of an affine GARCH dynamic because it has three main advantages. The first is the tractability in the cumulant exponent which enables us to price European-style derivatives by inverting the transform. Second, the leverage effect can be trivially generated in the GARCH process as $\mathbf{h}_{t+1}$ must be correlated with at least one of the processes in $\mathbf{X}_t$. This produces automatic correlation in asset returns and volatilities. In addition, pricing derivatives will also be straightforward as we do not have to apply a complex measure change technique as in Carr and Wu (2004) in order to solve for $\Psi_{\mathbf{X}}(\Phi;\ t,\ T)$ analytically. The third advantage of using GARCH is that it is a simple filter to implement. Models that use GARCH as the filter are therefore extremely useful in the empirical asset pricing research. For future reference, we name asset pricing dynamics based on our framework introduced above as Lévy GARCH processes.

3 Lévy GARCH models

3.1 Asset returns under the physical measure

We start our analysis in the physical measure using the d-dimensional contemporaneously independent Lévy processes $\mathbf{X}_{t+1}$.

Assumption 1 *Let $\mathbb{P}$ denote the physical measure with the asset return process given by*

$$R_{t+1} = \log\left(\frac{S_{t+1}}{S_t}\right) = r_{t+1} + \boldsymbol{\lambda}'\mathbf{h}_{t+1} + \boldsymbol{\vartheta}'\mathbf{X}_{t+1} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})'\,\mathbf{h}_{t+1}, \tag{3.1}$$

where $\boldsymbol{\vartheta} \in \mathbb{R}^d$, and the elements in $\mathbf{h}_{t+1} \in \mathbb{R}^d$ are time-homogeneous parameters of $\mathbf{X}_{t+1}$.

Note that $\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})'\,\mathbf{h}_{t+1}$ is the convexity adjustment term which makes $\boldsymbol{\vartheta}'\mathbf{X}_{t+1}$ a martingale. The conditional expectation of the asset price is given by $E_t[S_{t+1}] = S_t e^{r_{t+1}+\boldsymbol{\lambda}'\mathbf{h}_{t+1}}$, where r_{t+1} is the risk-free rate applicable from t to $t+1$. We assume time-deterministic functions for interest rates and ignore dividends for notational simplicity. Thus, the rate of return in excess of the risk-free rate is equal to $\boldsymbol{\lambda}'\mathbf{h}_{t+1}$. We will refer to $\boldsymbol{\lambda}'\mathbf{h}_{t+1}$ as the conditional equity premium. This is similar to the risk return trade-off in the multifactor APT framework, with $\boldsymbol{\lambda} \in \mathbb{R}^d$ denoting the market price of risks. When $\mathbf{h}_{t+1}$ follows a GARCH dynamic, the model produces time-varying risk premia. The conditional variance of the log return is given by $Var_t(R_{t+1}) = \boldsymbol{\vartheta}' Cov_t\left(\mathbf{X}_{t+1}\right)\boldsymbol{\vartheta}$. The $Cov_t\left(\mathbf{X}_{t+1}\right)$ term is the conditional covariance matrix, which is diagonal and affine in $\mathbf{h}_{t+1}$. Consequentially, the conditional return variance will also be an affine function of $\mathbf{h}_{t+1}$. Therefore, the time-homogeneous parameters $\mathbf{h}_{t+1}$ can most usefully be thought of as the parameters that control the variance of returns.

3.2 Specifications of the Lévy process

Many different Lévy processes can be used in the Lévy GARCH framework. We summarize their properties in Table 1. All the processes considered here satisfy the property of time homogeneity. We let h_{t+1} denote the time-homogeneous parameter in all types of Lévy processes that we consider.

The Lévy triplet $[\mu,\sigma^2,\nu(dx)]$ completely determines the characteristics of a Lévy process. A simple case of the normal distribution arises when there is no jump, $\nu(dx) = 0$, and the drift part is zero, $\mu = 0$. The variance in the diffusive part becomes the time-homogeneous parameter $h_{t+1} = \sigma^2$. This type of Lévy innovation will serve as the building block for several affine GARCH dynamics. The cumulant exponent of the $N(0, h_{t+1})$ is given by

$$\Psi_{Normal}(\phi; t, t+1) = h_{t+1}\frac{\phi^2}{2}.$$

Pure jump Lévy processes are classified according to the property of their Lévy measure $\nu(dx)$. When $\int_{\mathbb{R}_0} \nu(dx) < \infty$, the Lévy process is of finite activity as there are finitely many jumps (including zero) in any finite interval. The jump process of Merton (1976) is a well-known example of a finite activity jump process. Its construction is based on the compound Poisson process where each jump is distributed as $Normal\left(\theta, \delta^2\right)$ and arrives according to a Poisson distribution with arrival rate h_{t+1}. The cumulant exponent of a Merton jump process can be derived using (2.2) or through successively applying iterated expectations on Poisson and Normal random variables. It has the following form

$$\Psi_{MJ}(\phi;\ t,\ t+1) = h_{t+1}(e^{\theta\phi+\frac{1}{2}\phi^2\delta^2} - 1).$$

The use of the Merton jump structure is the standard practice in the continuous-time jump-diffusion literature. See Duffie, Pan and Singleton (2000) for details. Applications of the Merton jump process to discrete-time GARCH processes include the work of Maheu and Mccurdy (2004), Duan, Ritchken, and Sun (2006).

Unlike finite activity jump processes, infinite activity jumps can arrive with infinite numbers in any finite time interval. In this case, the integral of the Lévy measure $\int_{\mathbb{R}_0} \nu(dx)$ no longer exists. Several of these processes have been extensively studied in the asset pricing literature. Examples include distributions that exist only on the positive half line such as the Gamma distribution, the Inverse Gaussian (IG) distribution, and the Tempered Stable (TS) distribution. These positive-supported distributions are good candidates for modelling jumps in volatility with an affine GARCH dynamic. Examples of infinite activity jump processes that exist on the real line are the Variance Gamma (VG) model of Madan and Milane (1991), the Normal Inverse Gaussian (NIG) of Barndorff-Nielsen (1998), the CGMY model of Carr et al. (2002), the log-stable (LS) model of Carr and Wu (2003a), and the Meixner process of Schoutens (2000). The VG and NIG distributions are subclasses of the family of Generalized Hyperbolic distributions studied in Prause (1999).

3.3 Affine GARCH specifications

For any one-period models with final time T, the conditional transform of asset returns can usually be derived without much difficulty. According to our Lévy GARCH framework (3.1), we can write this transform as an exponential affine function in $\mathbf{h}_T$ as $E_{T-1}\left[e^{\phi R_T}\right] = e^{\phi r_T + \mathcal{Z}(\phi)'\mathbf{h}_T}$, where $\mathcal{Z}(\phi)$ is a d-dimensional vector with a general form

$$\phi\boldsymbol{\lambda} - \phi\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}) + \boldsymbol{\xi}_{\mathbf{X}}(\phi\boldsymbol{\vartheta}).$$

The explicit expression of $\mathcal{Z}(\phi)$ will depend on the choice of Lévy innovations as $\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})$ is the coefficient in the cumulant exponent of $\mathbf{X}_{t+1}$ which we summarize in Table 1. For multiple-return periods, the expression for the conditional transform of asset returns will also depend on the dynamic of $\mathbf{h}_{t+1}$, which we model using affine GARCH processes.

Definition 2 ***Affine GARCH.*** *Consider a GARCH dynamic $\mathcal{G}(\mathbf{h}_t, \boldsymbol{v}_t | \Omega_t)$ that provides the filtration for the time series of $\mathbf{h}_{t+1} \in \mathbb{R}^d$ based on the Lévy innovations $\mathbf{X}_t \in \mathbb{R}^d$. For generality, we let m be the number of auto-regressive lags on $\mathbf{h}_{t+1}$ and n be the number of lags on the vector of state variables $\boldsymbol{v}_{t+1} \in \mathbb{R}^q$, where both $\mathbf{h}_{t+1}$ and $\boldsymbol{v}_{t+1}$ are predictable at time t. If the joint conditional transform of*

$$(\mathbf{X}_{t+1}, \mathbf{h}_{t+2}, \boldsymbol{v}_{t+2}) \text{ evaluated at } \Pi = (\Pi_x, \Pi_h, \Pi_v)$$

with Π_x and Π_h $d \times 1$ vectors, and Π_v a $q \times 1$ vector, has an affine expression in $(\mathbf{h}_{t+1}, \boldsymbol{v}_{t+1})$ and their lags according to

$$E_t\left[e^{\Pi_x'\mathbf{X}_{t+1} + \Pi_h'\mathbf{h}_{t+2} + \Pi_v'\boldsymbol{v}_{t+2}}\right] = e^{\mathcal{V}(\Pi) + \sum_{i=1}^{m} \mathcal{W}_i(\Pi)'\mathbf{h}_{t+2-i} + \sum_{j=1}^{n} \mathcal{Y}_j(\Pi)'\boldsymbol{v}_{t+2-j}} \tag{3.2}$$

with a scalar $\mathcal{V}(\Pi)$, $d \times 1$ vectors $\mathcal{W}_i(\Pi)$'s and $q \times 1$ vectors $\mathcal{Y}_j(\Pi)$'s, then $\mathcal{G}(\mathbf{h}_t, \boldsymbol{v}_t | \Omega_t)$ has an affine GARCH dynamic.

Our definition only requires that the joint conditional transform of $\left(\mathbf{X}_{t+1}^{\prime}, \mathbf{h}_{t+2}^{\prime}, \; \boldsymbol{\upsilon}_{t+2}^{\prime}\right)^{\prime}$ be exponential affine in $(\mathbf{h}_{t+1}, \boldsymbol{\upsilon}_{t+1})$ and their lags. This is sufficient as we will show that, when this condition is met, the generating function for asset prices will be exponential affine, and hence analytically tractable. For applications to asset pricing at time T conditional on period t, we are interested in the solution to the generating function of S_T

$$f(\phi; t, T) = E_t\left[S_T^{\phi}\right] = S_t^{\phi} E_t\left[e^{\phi \sum_{k=1}^{T-t} R_{t+k}}\right]$$

for $\phi \in \mathbb{R}$.[8] There exist a plethora of combinations of GARCH dynamics and Lévy innovations that can be nested into the Lévy GARCH framework. We therefore provide a general form of their generating function. In subsequent sections, we present explicit expressions for a few selected cases that are widely used in the existing literature.

Proposition 1 *Consider the asset price dynamic in (3.1), where heteroskedasticity in the model is driven by the affine GARCH process* $\mathcal{G}(\mathbf{h}_t, \boldsymbol{\upsilon}_t | \, \Omega_t)$ *on the time-homogeneous parameter* $\mathbf{h}_{t+1}$*.The solution to the generating function of the asset price at time* T*, conditional on time* t*, takes the form*

$$f(\phi; t, T) = S_t^{\phi} e^{\mathcal{A}(\phi;t,T) + \sum_{i=1}^{m} \mathcal{B}_i(\phi;t,T)^{\prime} \mathbf{h}_{t+2-i} + \sum_{j=1}^{n} \mathcal{C}_j(\phi;t,T)^{\prime} \boldsymbol{\upsilon}_{t+2-j}}, \tag{3.3}$$

where the affine coefficients can be solved through the following recursive relations

$$\begin{aligned}
\mathcal{A}(\phi; t, T) &= \phi r_{t+1} + \mathcal{A}(\phi; t+1, T) + \mathcal{V}(\Pi) \\
\mathcal{B}_1(\phi; t, T) &= \phi(\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})) + \mathcal{B}_2(\phi; t+1, T) + \mathcal{W}_1(\Pi) \\
\mathcal{B}_i(\phi; t, T) &= \mathcal{B}_{i+1}(\phi; t+1, T) + \mathcal{W}_k(\Pi) \quad \textit{for} \;\; i = 2, \ldots, m-1 \\
\mathcal{C}_j(\phi; t, T) &= \mathcal{C}_{j+1}(\phi; t+1, T) + \mathcal{Y}_j(\Pi) \qquad \textit{for} \;\; j = 1, \ldots, n-1 \\
\mathcal{B}_m(\phi; t, T) &= \mathcal{W}_m(\Pi) \quad ; \quad \mathcal{C}_n(\phi; t, T) = \mathcal{Y}_n(\Pi)
\end{aligned}$$

with

$$\Pi = (\phi\boldsymbol{\vartheta}, \mathcal{B}_1(\phi; t+1, T), \mathcal{C}_1(\phi; t+1, T)).$$

At the terminal date, these affine coefficients must satisfy the boundary conditions $\mathcal{A}(\phi; T, T) = 0$*,* $\mathcal{B}_i(\phi; T, T) = 0$ *for all* i*'s, and* $\mathcal{C}_j(\phi; T, T) = 0$ *for all* j*'s.*

Proof. See appendix A. ■

[8] Notice that the generating function of the asset price $f(\phi; t, T)$ is a product of the conditional transform of the asset return and S_t^{ϕ}.

The simplest case is when $\mathbf{h}_t$ is homoskedastic because there is no need to recursively update these coefficients. In fact, their solutions take extremely simple forms

$$\mathcal{A}(\phi;t,T) = \phi \sum_{k=1}^{T-t} r_{t+k} \; ; \quad \mathcal{B}_1(\phi;t\;,T) = \mathcal{Z}(\phi)(T-t)\,,$$

and zeros for the rest of the affine coefficients.

In this paper, we illustrate the application of Lévy GARCH by focusing on a version of (3.1) that is most representative of modern asset pricing models. The recent trend in option pricing is to model the asset's log returns with two stochastic components: a Brownian component and a pure jump component. To model the leverage effect, the Brownian part is specified to be correlated with shocks in volatilities. Under this setup, the $\mathbb{P}$-measure log return process can be written as

$$\mathbb{P}\; measure: \; R_{t+1} = r_{t+1} + \mu_z h_{z,t+1} + \mu_y h_{y,t+1} + z_{t+1} + y_{t+1}, \tag{3.4}$$

where z_{t+1} is normally distributed as $Normal\,(0, h_{z,t+1})$ and y_{t+1} is a pure jump process. For generality, we leave y_{t+1} unspecified with only the requirement that $h_{y,t+1}$ be its time-homogeneous parameter. We note that this dynamic is a special case of (3.1) with $\mathbf{X}'_{t+1} = (z_{t+1}, y_{t+1})$, and $\boldsymbol{\vartheta}' = (1,1)$. To assist with econometric identification, we let $\mu_z = \lambda_z - \xi_z(1)$ and $\mu_y = \lambda_y - \xi_y(1)$, which represent the market price of risks and the martingale compensators respectively.

3.3.1 Heteroskedasticity via a GARCH(1,1) dynamic

Huang and Wu (2004) use a similar setup as in (3.4) to study option pricing models based on time-changed Lévy processes with stochastic time change that follows the square root model of Cox, Ingersoll and Ross (1985). We can extend this setup to our Lévy GARCH process. For the dynamic of $\mathbf{h}_{t+1}$, we use the affine GARCH(1,1) of Heston and Nandi (2000), which is known to have CIR square root process as its continuous-time limit; see appendix B in Heston and Nandi (2000) for proof.

We assume two different affine GARCH(1,1) dynamics for $h_{z,t+1}$ and $h_{y,t+1}$

$$\begin{aligned} h_{z,t+1} &= w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}}(z_t - c_z h_{z,t})^2 \\ h_{y,t+1} &= w_y + b_y h_{y,t} + \frac{a_y}{h_{z,t}}(z_t - c_y h_{z,t})^2 . \end{aligned} \tag{3.5}$$

The return specification in (3.4) together with the GARCH specification (3.5) encompasses various models that have been studied in the option pricing literature. In fact, it has continuous-time limits that are equivalent to those of the SV4 model of Huang and Wu (2004). We note that the parametrization of (3.5) is slightly different from the affine GARCH(1,1) model in

Heston and Nandi (2000). This is because we use nonstandardized normal innovations as the conditioning variables in the above GARCH dynamic. There is no empirical consequence from writing the model according to (3.5) instead of using the notation in Heston and Nandi (2000). The solution to the conditional transform for the return process (3.4) and the affine GARCH(1,1) dynamic (3.5) takes the following form

$$f(\phi;\ t,\ T) = S_t^{\phi} e^{\mathcal{A}(\phi;\ t,\ T)+\mathcal{B}^z(\phi;\ t,\ T)h_{z,t+1}+\mathcal{B}^y(\phi;\ t,\ T)h_{y,t+1}}, \tag{3.6}$$

where the expressions for the affine coefficients are provided in appendix B. Notice that in the simple GARCH(1,1) case, there is no state variable $\boldsymbol{\upsilon}_{t+1}$, and hence all $\mathcal{C}_j$ coefficients disappear. In addition, we drop the i subscripts in $\mathcal{B}_i$ since there is only one autoregressive lag in $\mathbf{h}_{t+1}$. However, extensions to GARCH(m,n) can be easily derived based on our general solution.

When there is no pure jump component, $h_{y,t} = 0$, (3.4) and (3.5) reduce to the Heston-Nandi GARCH(1,1) model. The GARCH filtrations of $h_{z,t+1}$ and $h_{y,t+1}$ are based on expost knowledge of the normal shock z_t of the return R_t. This approach is quite unconventional of a GARCH specification as z_t is now a latent process, and must be filtered out. The motivation for separating the normal component from the total return innovation $z_t + y_t$ is to obtain an affine GARCH dynamic. It could be argued that the need to filter z_t somewhat undermines the purpose of a GARCH process. However, $h_{z,t+1}$ and $h_{y,t+1}$ can still be easily updated once z_t is known. The filtration of z_t is constructed using $E\left[z_t \mid \Omega_t\right]$, and particle filtering can be used for estimating the model. Particle filtering is very fast and efficient. It can handle highly nonlinear state-space forms. Moreover, its implementation naturally lends itself to model estimation based on the maximum likelihood method as shown by Pitt (2002). We discuss our proposed method for filtering and estimation in a later section.

3.3.2 Richer heteroskedastic behavior through a component model

It is known that it is far from accurate to model the dynamic of returns and volatility using a normal innovation together with a simple GARCH(1,1) process. A similar observation applies to stochastic volatility models such as Heston (1993). This issue has been addressed in the literature by either modeling the return with nonnormal innovations or by allowing for richer volatility dynamics. The latter approach includes the introduction of component factors in the volatility modeling as in Engle and Lee (1999), Alizadeh, Brandt and Diebold (2002), Bollerslev and Zhou (2002), and Chernov, Gallant, Ghysels and Tauchen (2002). In the GARCH literature, Engle and Lee (1999) model volatility using two factors representing the short-run and long-run components. Christoffersen, Jacobs, Ornthanalai and Wang (2008) develop an affine GARCH version of Engle and Lee (1999), and demonstrate its superior option pricing performance relative to the benchmark affine GARCH(1,1) process. The richer volatility dynamics enable the component model to capture patterns in long-maturity as well as in short-maturity options.

We can apply the specification of Christoffersen, Jacobs, Ornthanalai and Wang (2008) to the Lévy GARCH framework. Assuming the return process (3.4), we write the dynamic of $h_{x_i,t+1}$ for $x_i = \{z, y\}$ as

$$\begin{aligned} h_{x_i,t+1} &= \upsilon_{x_i,t+1} + b_{x_i}\left(h_{x_i,t} - \upsilon_{x_i,t}\right) + \frac{a_{x_i}}{h_{z,t}}\left(\left(z_t^2 - h_{z,t}\right) - 2c_{x_i} z_t h_{z,t}\right) \\ \upsilon_{x_i,t+1} &= w_{x_i} + \rho_{x_i}\upsilon_{x_i,t} + \frac{\varphi_{x_i}}{h_{z,t}}\left(\left(z_t^2 - h_{z,t}\right) - 2d_{x_i} z_t h_{z,t}\right). \end{aligned} \tag{3.7}$$

In this setup, the dynamic of $h_{z,t+1}$ and $h_{y,t+1}$ can most usefully be thought of as having two components: $\upsilon_{z,t+1}$ and $\upsilon_{y,t+1}$ which represent the long-run components, and $h_{z,t+1} - \upsilon_{z,t+1}$ and $h_{y,t+1} - \upsilon_{y,t+1}$ which represent the short-run components.

From the dynamic in (3.7), we see that the long-run components, $\upsilon_{z,t+1}$ and $\upsilon_{y,t+1}$, are the state variables in our affine GARCH setup. We can also observe that the affine component GARCH dynamic has one lag ($m = n = 1$) in both the time-homogeneous parameters and the state variables. Therefore, the solution to the generating function for asset price will also contain a coefficient $\mathcal{C}(\phi; t, T)$ in addition to $\mathcal{B}(\phi; t, T)$ and $\mathcal{A}(\phi; t, T)$. The generating function takes the form

$$f(\phi;\ t,\ T) = S_t^{\phi} e^{\mathcal{A}(\phi;\ t,T) + \mathcal{B}(\phi;\ t,T)'\mathbf{h}_{t+1} + \mathcal{C}(\phi;\ t,T)'\boldsymbol{\upsilon}_{t+1}}, \tag{3.8}$$

where $\mathbf{h}_{t+1} = (h_{z,t+1}, h_{y,t+1})'$, $\boldsymbol{\upsilon}_{t+1} = (\upsilon_{z,t+1}, \upsilon_{y,t+1})'$, and

$$\mathcal{B}(\phi; t, T) = \begin{pmatrix} \mathcal{B}^z(\phi; t, T) \\ \mathcal{B}^y(\phi; t, T) \end{pmatrix}, \text{ and } \mathcal{C}(\phi; t, T) = \begin{pmatrix} \mathcal{C}^z(\phi; t, T) \\ \mathcal{C}^y(\phi; t, T) \end{pmatrix}.$$

Note again that since $m = n = 1$, we have dropped the i and j subscripts in $\mathcal{B}_i$ and $\mathcal{C}_j$ for notational simplicity. The recursive relations for these affine coefficients are included in appendix B.

3.3.3 Heteroskedasticity with jumps in affine GARCH

The consensus in the literature is that jumps are important for realistic modeling of the return dynamic. However, opinions on jumps in volatility remain mixed. But increasingly, evidence from option data suggests the presence of jumps in volatility (see Eraker (2004), and Broadie, Chernov and Johannes (2007)). This feature of asset pricing can also be modeled in the Lévy GARCH framework. One simple method is to add a jump (infinite or finite-activity) process to the existing affine GARCH dynamic. For instance, consider the following return and affine

GARCH(1,1) dynamic augmented with a jump

$$\begin{aligned} R_{t+1} &= r_{t+1}+\mu_z h_{z,t+1}+\mu_y h_{y,t+1}+z_{t+1}+\eta y_{t+1} \\ h_{x_i,t+1} &= w_{x_i}+b_{x_i}h_{x_i,t}+d_{x_i}y_t+\frac{a_{x_i}}{h_{z,t}}\left(z_t-c_{x_i}h_{z,t}\right)^2 \end{aligned} \tag{3.9}$$

for $x_i = \{z, y\}$. The above model is a slight adaptation of (3.4) where z_{t+1} is normally distributed as $N(0, h_{z,t+1})$, and y_{t+1} is a pure jump process but with positive support. We limit the above model to jumps with the probability of $P(y_{t+1} < 0) = 0$ in order to ensure positivity in the dynamic of $\mathbf{h}_{t+1}$. This includes the restriction that d_z and d_y be equal or greater than zero. We also introduce the scaling coefficient $\eta \in \mathbb{R}$ for the jump innovations in the return dynamic, and hence we must have $\mu_y = \lambda_y - \xi_y(\eta)$ in (3.9). A negative value for η would imply that the jump y_{t+1} represents a downside risk in the return which is often referred to as a "crash", while a positive value of η would signify the opposite. Examples of jump processes with positive support include the Poisson jump with constant jump size θ, the inverse Gaussian distribution, the Gamma distribution, and the Meixner process. We note that Carr and Wu (2004) introduce a similar process in their time-changed Lévy framework, where they use the log-stable (LS) process to model the negative jump. Since the LS process consists of infinite number of negative jumps, we can also extend their approach to (3.9) by letting y_{t+1} be distributed as LS and restricting d_z and d_y to be negative.

Adding jumps to the affine GARCH dynamic in this manner does not destroy the affine property because the cumulant exponent of y_t is affine in its time-homogeneous parameter. For the return and GARCH model (3.9), the generating function takes the form

$$f(\phi;\ t,\ T) = S_t^{\phi} e^{\mathcal{A}(\phi;\ t,\ T)+\mathcal{B}^z(\phi;\ t,\ T)h_{z,t+1}+\mathcal{B}^y(\phi;\ t,\ T)h_{y,t+1}} \tag{3.10}$$

where the expressions for the affine coefficients are provided in appendix B. A more complex affine GARCH model with jumps is studied in Christoffersen, Heston and Jacobs (2006), who model returns and price options using an affine inverse Gaussian GARCH. Their specification can also be nested in the Lévy GARCH framework.

4 The risk-neutral measure

4.1 Change of measure

In the discrete-time framework, stock prices can jump to an infinite set of values in a single period, thus the equivalent martingale measure (EMM) is not unique. We follow the usual approach (e.g. Heston (1993)) by establishing the existence of a risk-neutral probability density such that returns on all assets' ex-dividend payout are equal to the risk-free rate.

Assumption 2 *The conditional Radon-Nikodym derivative that links the physical measure*

($\mathbb{P}$) to the risk-neutral measure ($\mathbb{Q}$) is given by

$$\frac{\frac{dQ^{\Lambda}}{dP} |\ \mathcal{F}_{t+1}}{\frac{dQ^{\Lambda}}{dP} |\ \mathcal{F}_{t}} = \exp\left(\boldsymbol{\Lambda}'\mathbf{X}_{t+1} - \boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\Lambda}\right)'\mathbf{h}_{t+1}\right), \tag{4.1}$$

where $\mathbf{X}_{t+1} \in \mathbb{R}^d$ is a vector of Lévy innovations in returns, and $\boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\Lambda}\right)'\mathbf{h}_{t+1}$ is similarly defined as in (3.1). We refer to $\boldsymbol{\Lambda} \in \mathbb{R}^d$ as the EMM coefficients that act as the wedge between the objective and risk-neutral measure.

This $\mathbb{Q}$ measure depends on the values in $\boldsymbol{\Lambda}$, which have to be solved prior to pricing derivatives. The solutions will be such that the discounted return on assets less its dividend is a martingale. Our choice of the Radon-Nikodym derivative falls within the framework of Gerber and Shiu (1994) who use the Esscher transform (1932) to price options. Applications of the Esscher transform in derivatives pricing is common, see for example Carr and Wu (2004) for an application in option pricing, and Ahn, Dai and Singleton (2007) for an application to term structure modeling.[9]

The solution for $\boldsymbol{\Lambda}$ is obtained by imposing the local martingale restriction on asset returns under the $\mathbb{Q}$ measure. That is, for any unit time interval,

$$E_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}\right)\right] = \exp\left(r_{t+1}\right).$$

Multiple solutions may exist such that the local martingale restriction holds. However, the identification of $\boldsymbol{\Lambda}$ is facilitated by the specification of our Lévy GARCH framework. We make use of the property that the cumulant exponent of our Lévy processes and total equity premium are both affine in $\mathbf{h}_{t+1}$ to derive the following proposition.

Proposition 2 *Assume an exponential affine model for asset prices as in (3.1) with the Radon-Nikodym derivative given by (4.1). For the existence of an equivalent martingale measure $\mathbb{Q}$, the EMM coefficients in $\boldsymbol{\Lambda} \in \mathbb{R}^d$ must satisfy the following set of equations*

$$\boldsymbol{\lambda} = \boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\vartheta}\right) + \boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\Lambda}\right) - \boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\vartheta} + \boldsymbol{\Lambda}\right), \tag{4.2}$$

Proof. See appendix C. ■

For $\mathbf{X}_{t+1} \in \mathbb{R}^d$, (4.2) consists of d equations. These equations are solved independently with their solutions corresponding to each value in $\boldsymbol{\Lambda}$. We note how $\boldsymbol{\lambda}$ enters into (4.2) where each equation has its own market price of risk parameter. This simplifies identification of the risk premium that is associated with each Lévy shock. For example, consider the return dynamic in (3.4). Applying the local martingale restriction, we arrive at the following two

[9] See also Buhlmann et al. (1998), Chan (1999), and Siu, Tong and Yang (2004) who economically motivate the use of the Esscher transform as the economy's pricing kernel.

equations:

$$\lambda_y = \xi_y(1) + \xi_y(\Lambda_y) - \xi_y(1+\Lambda_y) \tag{4.3}$$

$$\lambda_z = -\Lambda_z. \tag{4.4}$$

We have the same solution for Λ_z as that from the method of LRNVR of Duan (1995). In order to solve for Λ_y, more details on the specification of the jump process is required. In some cases, we cannot analytically solve for Λ_y, but from (4.3), it can be solved for numerically with ease. The market prices of risks λ_z and λ_y enter separately into the above two equations (4.4) and (4.3). This result is very useful because it allows us to isolate the two sources of risk in the economy, namely normal and jump risks.

4.2 The market price of risk

We assume that the total equity premium is affine in $\mathbf{h}_{t+1}$. According to (3.1), the total equity premium is given by

$$\gamma_{t+1} = \log \frac{E_t^{\mathbb{P}}[\exp(R_{t+1})]}{E_t^{\mathbb{Q}}[\exp(R_{t+1})]} = \sum_{i=1}^{d} \lambda_{x_i} h_{x_i,t+1}, \tag{4.5}$$

where λ_{x_i} and $h_{x_i,t+1}$ are the market price of risk and the time-homogenous parameters associated with the i^{th} Lévy innovation in $\mathbf{X}_{t+1}$. The equity premium γ_{t+1} exhibits time variation that depends on the GARCH dynamic of $\mathbf{h}_{t+1}$. This is a desirable result given the existing evidence for stochastic risk premia, see for instance Fama (1984), Fama and Bliss (1987), and Bakshi, Carr and Wu (2008). When $h_{x_i,t+1} = h_{x_i}$ is constant, the premium required to hold the i^{th} factor will also be constant. Intuitively, the required premium on an investment that is exposed to the uncertainty x_i should not change unless the nature of the risk in x_i changes (h_{x_i} changes), or there is a change in the investor's risk-taking behavior (λ_{x_i} changes).

The long-run equity premium is determined by taking unconditional expectations of (4.5). This gives $\overline{\gamma} = \sum_{i=1}^{d} \lambda_{x_i} E[h_{x_i,t+1}]$, where $E[h_{x_i,t+1}]$ will depend on the choice of GARCH dynamic for each $h_{x_i,t+1}$. The affine structure of $\overline{\gamma}$ has an interesting implication for empirical option pricing studies. Given a fixed level of the long-run equity premium, we can study how changing the weights that each shock has in the total equity premium affects option prices. This could shed light on the prevailing debate of relative significance that jump risk premia have on option pricing.

There is a direct relationship between the market prices of risks and the EMM coefficients. That is, for each set of risk and return trade-offs $\boldsymbol{\lambda}$, there is a corresponding $\boldsymbol{\Lambda}$ that determines the divergence between the physical and risk-neutral distributions of asset prices. Notice that in equation (4.2), when the risk x_i is not priced in the market $(\lambda_{x_i} = 0)$, we must have $\Lambda_{x_i} = 0$ as the solution. In a special case when none of the risks are priced, $\boldsymbol{\Lambda} = \mathbf{0}$, the distributions of the asset price under the two probability measures are identical. In the standard risk-return trade-off, an average investor would demand a reward for bearing some level of uncertainty ($\lambda_{x_i} > 0$). This aversion to risk implies a positive value of the total equity premium. Consequentially, the resulting EMM coefficient Λ_{x_i} will be negative to ensure that the discounted stock price is a $\mathbb{Q}$ martingale. In the case of term structure models where all innovations are normally distributed, the two parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\Lambda}$ are linearly related (see equation (4.4)), and their names often used interchangeably. Ahn, Dai and Singleton (2007) refer to Λ as the market price of risk parameter.

Our choice of the Radon-Nikodym derivative (4.1) leads to a pricing kernel that is identical to the one commonly used in the affine literature which takes the form

$$\ln \frac{M_t}{M_{t+1}} = u_m - \rho_m \ln \frac{S_{t+1}}{S_t} - \sum_{i=1}^{d} \rho_i x_{i,t+1}, \tag{4.6}$$

with $\rho_i = -\rho_m - \Lambda_{x_i}$ and u_m is such that $E_t [M_{t+1}/M_t] = \exp(-r_{t+1})$. As pointed out in Bates (2006), this specification nests various approaches. For instance, when $\rho_m = -\Lambda_{x_i}$ for all $i = 1, \ldots, d$, (4.6) reduces to a myopic power utility pricing kernel used in the implicit pricing kernel literature. Furthermore, if S_t is a good proxy for the overall wealth such as the market index, ρ_m is the coefficient of relative risk aversion in an economy with power utility. When the coefficient ρ_i is nonzero, it implies that there is a risk premium for the i^{th} shock in addition to the direct wealth-related effects on marginal utility captured by $\rho_m \ln (S_{t+1}/S_t)$. For instance, in our return model (3.4), nonzero values for ρ_z and ρ_y imply the presence of volatility and jump risk premia that are extraneous to the direct effects of the overall wealth. This corresponds to an economy where investors are averse to jumps (see Bates (2006)) and volatility-related shifts in the investment opportunity set.

4.3 Asset price dynamics under the risk-neutral measure

To price derivatives, we must characterize the dynamic of the asset price under the risk-neutral measure. First, we look at the effect that the change of measure (4.1) exerts on the Lévy processes $\mathbf{X}_{t+1}$.

Proposition 3 *Consider a $\mathbb{P}$-measure Lévy process $\mathbf{X}_{t+1} \in \mathbb{R}^d$ with the conditional cumulant exponent given by (2.5). Its equivalent process $\mathbf{X}^*_{t+1}$ under the risk-neutral measure, consistent with the Radon-Nikodym derivative in (4.1), is characterized by the following conditional cumulant exponent*

$$\begin{aligned} \Psi^{\mathbb{Q}}_{\mathbf{X}} (\Phi;\ t,\ t+1) &= (\boldsymbol{\xi}_{\mathbf{X}} (\Phi + \boldsymbol{\Lambda}) - \boldsymbol{\xi}_{\mathbf{X}} (\boldsymbol{\Lambda}))' \, \mathbf{h}_{t+1} \\ &= \boldsymbol{\xi}^*_{\mathbf{X}} (\Phi)' \, \mathbf{h}^*_{t+1}, \end{aligned} \tag{4.7}$$

where $\mathbf{h}^*_{t+1} \in \mathbb{R}^d$ is a vector of $\mathbb{Q}$-measure time-homogeneous parameters, and $\boldsymbol{\xi}^*_{\mathbf{X}} (\Phi)$ is the coefficient in the cumulant exponent of $\mathbf{X}^*_{t+1}$.

Proof. Apply the change of measure and take the expectation as follows

$$\begin{aligned} E_t^{\mathbb{Q}}\left[e^{\Phi'\mathbf{X}_{t+1}}\right] &= E_t\left[e^{\mathbf{\Lambda}'\mathbf{X}_{t+1}-\boldsymbol{\xi}_{\mathbf{X}}(\mathbf{\Lambda})'\mathbf{h}_{t+1}+\Phi'\mathbf{X}_{t+1}}\right] \\ &= e^{(\boldsymbol{\xi}_{\mathbf{X}}(\Phi+\mathbf{\Lambda})-\boldsymbol{\xi}_{\mathbf{X}}(\mathbf{\Lambda}))'\mathbf{h}_{t+1}} = e^{\Psi_{\mathbf{X}}^{\mathbb{Q}}(\Phi;\ t,\ t+1)}, \end{aligned}$$

and equate the exponents. ■

Another advantage of the Lévy GARCH framework is that the $\mathbb{Q}$-measure cumulant exponent of $\mathbf{X}_{t+1}^*$ will also be affine in the vector of time-homogeneous parameters. Given this, we are ready to show the risk-neutral return dynamic for the Lévy GARCH process in (3.1).

Lemma 1 ***Risk-Neutral Lévy GARCH dynamic.*** *Under the risk-neutral measure, the asset's return dynamic can be written as*

$$\log\left(\frac{S_{t+1}}{S_t}\right) = r_{t+1} - \boldsymbol{\xi}_{\mathbf{X}}^*\left(\boldsymbol{\vartheta}\right)'\mathbf{h}_{t+1}^* + \boldsymbol{\vartheta}'\mathbf{X}_{t+1}^*$$

where $\mathbf{X}_{t+1}^*$ is the risk-neutral Lévy process with the cumulant exponent (4.7).

Proof. See appendix C. ■

When using the Esscher transform to define the change of measure, the new process $\mathbf{X}_{t+1}^*$ will always be an equivalent Lévy process.[10] The conditional $\mathbb{Q}$-measure cumulant exponent of the asset return can be derived using Proposition 3, and the conditional risk-neutral density of asset returns can be computed via Fourier inversion. This is advantageous because we do not need to recognize the distribution that characterizes $\mathbf{X}_{t+1}^*$ in order to price derivatives, as long as its conditional transform under the physical measure is known.

If, under the risk-neutral measure, the affine GARCH dynamic $\mathcal{G}\left(\mathbf{h}_t, \boldsymbol{\upsilon}_t\middle|\, \Omega_t^{\mathbb{Q}}\right)$ remains affine, which according to our definition (3.2) implies that

$$E_t^{\mathbb{Q}}\left[e^{\mathbf{\Pi}'\left(\mathbf{X}_{t+1}',\mathbf{h}_{t+2}',\ \boldsymbol{\upsilon}_{t+2}'\right)'}\right] = e^{\mathcal{V}^*(\Pi)+\sum_{i=1}^m \mathcal{W}_i^*(\Pi)'\mathbf{h}_{t+2-i}+\sum_{j=1}^n \mathcal{Y}_j^*(\Pi)'\boldsymbol{\upsilon}_{t+2-j}}. \tag{4.8}$$

In this case, the generating function of asset prices is analytically known and can be immediately derived. Note that we apply star superscripts to the above affine coefficients in order to distinguish them from their physical measure counterparts. The generating function of asset prices under the risk-neutral measure at time T, conditional on the current period t, takes the same form as (3.3)

$$f^*\left(\phi; t, T\right) = S_t^{\phi} e^{\mathcal{A}^*(\phi;t,T)+\sum_{i=1}^m \mathcal{B}_i^*(\phi;t,T)'\mathbf{h}_{t+2-i}+\sum_{j=1}^n \mathcal{C}_j^*(\phi;t,T)'\boldsymbol{\upsilon}_{t+2-j}}.$$

[10] See the discussion of Delbaen, Schachermayer and Schweizer in Gerber and Shiu (1994). For the continuous-time version, see appendix A in Carr and Wu (2004).

Assuming the risk-neutral affine GARCH property in (4.8), these coefficients can be solved using the same recursive relations as in Proposition 1

$$\begin{aligned}
\mathcal{A}^*(\phi;t,T) &= \phi r_{t+1} + \mathcal{A}^*(\phi;t+1,T) + \mathcal{V}^*(\Pi) \\
\mathcal{B}_1^*(\phi;t,T) &= \phi(\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})) + \mathcal{B}_2^*(\phi;t+1,T) + \mathcal{W}_1^*(\Pi) \\
\mathcal{B}_i^*(\phi;t,T) &= \mathcal{B}_{i+1}^*(\phi;t+1,T) + \mathcal{W}_k^*(\Pi) \quad \text{for} \quad i = 2,\ldots,m-1 \\
\mathcal{C}_j^*(\phi;t,T) &= \mathcal{C}_{j+1}^*(\phi;t+1,T) + \mathcal{Y}_j^*(\Pi) \quad \text{for} \quad j = 1,\ldots,n-1 \\
\mathcal{B}_m^*(\phi;t,T) &= \mathcal{W}_m^*(\Pi) \quad ; \quad \mathcal{C}_n^*(\phi;t,T) = \mathcal{Y}_n^*(\Pi)
\end{aligned}$$

with $\Pi = (\phi\boldsymbol{\vartheta}, \mathcal{B}_1^*(\phi;t+1,T), \mathcal{C}_1^*(\phi;t+1,T))$. Although the change of measure does not guarantee that $\mathcal{G}\left(\mathbf{h}_t, \boldsymbol{v}_t \middle| \Omega_t^{\mathbb{Q}}\right)$ will be affine, the design of an affine GARCH process often ensures that the affine property will hold after the measure change. In fact, all existing affine GARCH processes in the literature are also affine in the $\mathbb{Q}$ measure, including our examples in sections 3.3.1-3.3.3.

Although the pricing of derivatives using the inverse transform method does not require knowledge of the distribution of $\mathbf{X}_{t+1}^*$, it is often useful to understand the effect that measure change has on these Lévy innovations. If the processes in $\mathbf{X}_{t+1}^*$ stem from recognizable distributions, then we can price derivatives via Monte-Carlo simulations, which leads to potential applications of the Lévy GARCH framework to nonaffine GARCH dynamics. This is of paramount importance as most of the empirically studied GARCH dynamics are nonaffine. It turns out that, under the Lévy GARCH framework, several processes once transformed into the risk-neutral measure will be distributed in the same way as they were under the physical measure. For clarity, we use the simple example of a normally distributed random variable $z_{t+1} \sim N(0, h_{z,t+1})$ to illustrate our point. Applying the change of measure according to (4.7), we see that the risk-neutral process z_{t+1}^* will have the following conditional cumulant exponent

$$\Psi_z^{\mathbb{Q}}(\phi;t,t+1) = \phi\Lambda_z h_{z,t+1} + \frac{1}{2}\phi^2 h_{z,t+1.}$$

The above is essentially the conditional cumulant exponent for a normally distributed random variable with variance $h_{z,t+1}$and mean $\Lambda_z h_{z,t+1}$. Note that, from (4.4), $\Lambda_z = -\lambda_z$ and therefore the change of measure affects z_{t+1} by shifting its mean to the left by $\lambda_z h_{z,t+1}$.

Similarly, we can apply the change of measure (4.7) to each pure jump Lévy process presented in Table 1, and try to recognize the distribution of its transformed cumulant exponent. For most of the Lévy processes presented in Table 1, their risk-neutral and physical distributions are distributed according to the same type of stochastic process, although with different parameters. For convenience, we summarize the results in Table 2. We cannot associate the risk-neutral transformed cumulant exponent of the double exponential jump (DEP), and the log-stable (LS) process with any of the well-known distributions. We therefore denote their $\mathbb{Q}$

measure distribution as "unrecognizable". However, Mo and Wu (2007) use a special case of the DEP jump, with $p = \eta_1$ and $1 - p = \eta_2$, to study risk premia in international economies. They show that this process has a recognizable risk-neutral transformed density, and we refer to this distribution, which is a special case of Kou's DEP as the DEP-MW. We summarize the effect that the change of measure has on the parameters governing each Lévy process on the right column of Table 2. Because the transformation of these parameters from $\mathbb{P}$ to $\mathbb{Q}$ is purely algebraic, and is accomplished by applying (4.7), we do not include them in this paper.

Returning to the $\mathbb{P}$-measure asset price in (3.4), we can write its equivalent risk-neutral dynamic as

$$R_{t+1} = r_{t+1} + \left(\lambda_z - \frac{1}{2}\right) h_{z,t+1} - \xi_y^* (1) h_{y,t+1}^* + z_{t+1}^* + y_{t+1}^*,$$

where $z_{t+1}^* \sim N(-\lambda_z h_{z,t+1}, h_{z,t+1})$. For generality, we keep the distribution of y_{t+1} unspecified. Therefore, little can be said about y_{t+1}^* except that its conditional cumulant exponent is given by $\xi_y^* (\phi) h_{y,t+1}^*$. The convention in the GARCH literature is to express the normal shock as a mean zero innovation. Therefore, we use the transformation $z_{t+1}^* = z_{t+1} - \lambda_z h_{z,t+1}$, and rewrite the $\mathbb{Q}$-measure asset price as

$$\mathbb{Q}\ measure: \quad R_{t+1} = r_{t+1} - \frac{1}{2} h_{z,t+1} - \xi_y^* (1) h_{y,t+1}^* + z_{t+1} + y_{t+1}^* \tag{4.9}$$

with $z_{t+1} \sim N(0, h_{z,t+1})$.

The procedure for risk neutralizing an affine GARCH process is straightforward. It involves replacing the Lévy innovations $\mathbf{X}_{t+1}$ in the GARCH updating with the risk-neutral $\mathbf{X}_{t+1}^*$. Furthermore, we can write the $\mathbb{Q}$-measure affine GARCH process using the risk-neutral transformed parameters. We note that this condition is not necessary for the derivation of the $\mathbb{Q}$-measure generating function of the asset price. It is, however, the convention in the literature. The reparametrization of an affine GARCH dynamic implies writing $\mathcal{G}\left(\mathbf{h}_t, \boldsymbol{\upsilon}_t \middle| \Omega_t^{\mathbb{Q}}\right)$ as $\mathcal{G}\left(\mathbf{h}_t^*, \boldsymbol{\upsilon}_t^* \middle| \Omega_t^{\mathbb{Q}}\right)$. The relationship between $h_{x_i,t+1}^*$ and $h_{x_i,t+1}$ in the Lévy process x_i in $\mathbf{X}_{t+1}$ is given by

$$h_{x_i,t+1}^* = \left(\frac{\boldsymbol{\xi}_{x_i}(1 + \Lambda_{x_i}) - \boldsymbol{\xi}_{x_i}(\Lambda_{x_i})}{\boldsymbol{\xi}_{x_i}^*(1)}\right) h_{x_i,t+1},$$

for $i = 1, \ldots, d$. This follows from our result in (4.7). For the reparametrization of $\boldsymbol{\upsilon}_t$ using $\boldsymbol{\upsilon}_t^*$, more information on the nature of $\boldsymbol{\upsilon}_t$ is required. However, in most cases, it is possible to write the dynamic of $\mathcal{G}\left(\mathbf{h}_t^*, \boldsymbol{\upsilon}_t^* \middle| \Omega_t^{\mathbb{Q}}\right)$ using the same structure as $\mathcal{G}(\mathbf{h}_t, \boldsymbol{\upsilon}_t | \Omega_t)$ but with the risk-neutral transformed parameters and reparametrized GARCH coefficients. This method is highly convenient as the generating function for the asset price under the risk-neutral measure $f^*(\phi; t, T) = E_t^{\mathbb{Q}}\left[S_T^{\phi}\right]$ will be of the same form as $f(\phi; t, T)$. For convenience, we include the derivation of the risk-neutral dynamic for affine GARCH models discussed in sections 3.3.1-3.3.3 in the appendix D.

4.4 Option pricing

Given that the risk-neutral generating function of asset price $f^*(\phi; t, T)$ is known, we can value European-style derivatives using the Fourier inversion method as in Heston (1993), Lewis (2000), Duffie, Pan and Singleton (2000), and Bakshi and Madan (2000). Using one version of the inverse transform, the price of a European call option at time t with expiration date T , strike price K, and current spot price S_t is given by

$$\begin{aligned} C(S_t, K, T, r) &= e^{-r(T-t)} E_t^{\mathbb{Q}}\left[(S_T - K)^+\right] \\ &= S_t\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi} f^*(i\phi+1)}{i\phi\, f^*(1)}\right] d\phi\right) \\ &\quad - e^{-r(T-t)} K\left(\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{K^{-i\phi} f^*(i\phi)}{i\phi}\right] d\phi\right). \end{aligned} \tag{4.10}$$

Note that we have assumed a constant interest rate in order to simplify the expression.

4.5 Bond pricing

Following the extant literature on affine term structure models, we assume that the one-period risk-free rate from t to $t+1$ is given by $r_{t+1} = \delta_0 + \delta_x' \mathbf{X}_{t+1}$, where $\mathbf{X}_{t+1}$ is the d-dimensional Lévy process with $\delta_x \in \mathbb{R}_+^d$. The affine structure of r_{t+1} allows us to derive the conditional transform for the interest rate dynamic and to price zero-coupon bonds using the same techniques discussed in section 3.3. For the interest rate dynamic, the $\mathbb{P}$-measure conditional transform is given by $E_t\left[e^{\phi r_{t+k}}\right] = e^{\Psi_r(\phi; t, t+1)}$, where

$$\Psi_r(\phi; t, t+1) = \phi\delta_0 + \boldsymbol{\xi}_{\mathbf{X}}(\phi\delta_x)' \mathbf{h}_{t+1}$$

is the conditional cumulant exponent. Given this, all the statistical moments of the one-period interest rate dynamic can be derived, and its transition density $P(r_{t+1}|r_t)$ can be computed via Fourier inversion.

We now discuss the implications of our model for pricing zero-coupon bonds. The time t zero-coupon bond price paying \$1 at maturity T is given by

$$B(t, T) = E_t^{\mathbb{Q}}\left[e^{-\sum_{k=1}^{T-t} r_{t+k}}\right] = E_t\left[\prod_{k=1}^{T-t} \frac{M_{t+k}}{M_{t+k-1}}\right], \tag{4.11}$$

where M_{t+k} is the pricing kernel consistent with the equilibrium pricing measure

$$\frac{M_{t+k}}{M_{t+k-1}} = \frac{U'(C_{t+k-1})}{U'(C_{t+k})}$$

for $k = 1, \ldots, T-t$, and C_t is the aggregate consumption at time t. In order to solve the expectation (4.11) in the context of the Lévy GARCH framework, the dynamic of the pricing kernel must have an affine structure. From (4.6), we see that the setup of our economy leads to an affine pricing kernel. We can therefore apply the technique in Proposition 1 and solve for the price of a zero-coupon bond. For generality, and to keep notation manageable, we assume that the one-period affine pricing kernel takes the form

$$\log \frac{M_{t+1}}{M_t} = u_m + \theta_0' \mathbf{h}_{t+1} + \theta_x' \mathbf{X}_{t+1},$$

where θ_0 and θ_x are d-dimensional vectors. The constant term u_m is such that $E_t \left[M_{t+1}/M_t \right] = \exp\left(-r_{t+1}\right)$, where the statistical dynamic for r_{t+1} is discussed above. For notational simplicity, we will assume that the affine GARCH dynamic for $\mathbf{h}_{t+1}$ satisfies the property (3.5). Using the technique shown in the proof of Proposition 1, solving (4.11) is equivalent to solving for

$$E_t \left[e^{-\phi \sum_{k=1}^{T-t} \log \frac{M_{t+k}}{M_{t+k-1}}} \right] = e^{\mathcal{A}(\phi;t,T) + \sum_{i=1}^{m} \mathcal{B}_i(\phi;t,T)' \mathbf{h}_{t+2-i} + \sum_{j=1}^{n} \mathcal{C}_i(\phi;t,T)' \boldsymbol{v}_{t+2-j}}$$

and evaluating it at $\phi = 1$. The affine coefficients are solved using the same recursive relations as in Proposition 1, but with the following slight modifications

$$\begin{aligned} \mathcal{A}\left(\phi;t,T\right) &= \phi \mu_m + \mathcal{A}\left(\phi;t+1,T\right) + \mathcal{V}\left(\Pi\right) \ , \\ \mathcal{B}_1\left(\phi;t,T\right) &= \phi \theta_0 + \mathcal{B}_2\left(\phi;t+1,T\right) + \mathcal{W}_1\left(\Pi\right), \end{aligned}$$

and $\Pi = \left(\phi\theta_x, \mathcal{B}_1\left(\phi;t+1,T\right), \mathcal{C}_1\left(\phi;t+1,T\right)\right)$.

5 Estimation method

5.1 Filtration of latent variables

GARCH models are very popular among finance academics and practitioners due to their ease of implementation. When $\mathcal{G}\left(\mathbf{h}_t, \boldsymbol{v}_t \middle| \Omega_t\right)$ is an updating scheme conditional only on $\mathbf{h}_t$ and $\mathbf{1}'\mathbf{X}_t$, there is no need to separately identify each residual in $\mathbf{X}_t$. In this case, the filtration of $\mathbf{h}_{t+1}$ is extremely straightforward. For clarity, we illustrate this using the return model in (3.4) and adapt the GARCH dynamic of $h_{z,t+1}$ in (3.5) to

$$h_{z,t+1} = w_z + b_z h_{z,t} + \frac{a_z}{h_{z,t}} \left(z_t + y_t - c_z h_{z,t}\right)^2. \tag{5.1}$$

The difference between (3.5) and (5.1) lies in the conditioning innovation where, in (5.1), the total return residual $z_t + y_t$ enters directly into the GARCH dynamic. Therefore, the filtration of $h_{z,t+1}$ in (5.1) is extremely simple and can be performed quickly. We can directly estimate the model using standard maximum likelihood. Unfortunately, the dynamic of (5.1) does not admit tractable formulae for the cumulant exponent and hence no closed-form solution for

European-style derivatives exists.

When the dynamic of $\mathcal{G}(\mathbf{h}_t, \boldsymbol{\upsilon}_t | \Omega_t)$ requires that the total return residual is separated, we need to rely on a filtering technique in addition to GARCH. Because a model of this type consists of jumps which produce significant nonlinearity, a filter that relies on a linearized state-space, such as the Kalman filter, is not appropriate. In addition, many of the well-known Lévy processes do not have closed-form density functions, which makes the estimation problem more challenging, as we cannot construct likelihood functions in the traditional way. A few methods have been proposed for estimating models based on time-changed Lévy processes. Li, Wells and Yu (2006) use Markov Chain Monte Carlo (MCMC) method to estimate models with stochastic volatilities and Lévy jumps. This is a very reliable method because the model's parameters are drawn from its exact (or close approximate of the) posterior distributions. The downside of the MCMC method is that each model requires a different, custom-tailored estimation strategy. Moreover, it is computationally intensive.

Bates (2008) and Bakshi, Carr and Wu (2008) use Fourier inversion to obtain the conditional density of returns for a time-changed CGMY Lévy process. They estimate their models by maximizing their resulting likelihood functions. When the filtration of the time change process is exact, Fourier inversion will yield the true likelihood. Nevertheless, likelihood estimation based on the Fourier inversion can be computationally intensive as each inversion involves a numerical integration on the complex plane. For the filtration of the latent time change process, Bates (2008) uses the Approximate Maximum Likelihood (AML) method introduced in Bates (2006). On the other hand, Bakshi, Carr and Wu (2007) use the unscented Kalman filter which is accurate up to the second order for any nonlinearity.

One of our objectives in the development of the Lévy GARCH models is ease of implementation. In addition, we wish to use an estimation strategy that is based on the likelihood framework because it allows for simple performance comparison between various models. We propose that the particle filtering (PF) technique is used for the filtration of the return innovation.

5.2 The particle filter

The PF is based on the Sampling Importance Resampling (SIR) algorithm where the latent state variables are sampled and resampled with predetermined weights. The algorithm provides exact filtration when the number of particles used approaches infinity. Gordon, Salmond and Smith (1993) show that PF is a convenient filter for non-linear models. Johannes, Polson and Stroud (2008) apply the PF to continuous-time jump diffusion models and examine its property in detail. Christoffersen, Jacobs and Mimouni (2007) use the PF to fit returns and option prices on various affine and nonaffine stochastic volatility models.

Consider $\mathbf{L}_t$ as a vector of latent variables. The problem at hand is to compute the expectation

$$E\left[\mathbf{L}_t \mid \Omega_t\right] = \int \mathbf{L}_t \; P\left(\mathbf{L}_t \mid \Omega_t\right) d\mathbf{L}_t, \tag{5.2}$$

where $P\left(\mathbf{L}_t \mid \Omega_t\right)$ is the filtering density. The conditioning information set $\Omega_t = (R_0, R_1, \ldots, R_t)$, includes all current and past information on the asset's returns. By Bayes' rule, the filtering density is given by

$$\begin{aligned} P\left(\mathbf{L}_t \mid \Omega_t\right) &= \frac{P\left(R_t \mid \mathbf{L}_t,\ \Omega_{t-1}\right) P\left(\mathbf{L}_t \mid \Omega_{t-1}\right)}{P\left(R_t \mid \ \Omega_{t-1}\right)} \\ &\propto P\left(R_t \mid \mathbf{L}_t,\ \Omega_{t-1}\right) P\left(\mathbf{L}_t \mid \Omega_{t-1}\right). \end{aligned} \tag{5.3}$$

Direct integration of (5.2) is difficult as the density $P\left(\mathbf{L}_t \mid \Omega_t\right)$ is not always available in analytical form. However, if we can sample $\mathbf{L}_t$ from $P\left(\mathbf{L}_t \mid \Omega_t\right)$, then we can compute (5.2) by simple averaging. This procedure is referred to as Monte Carlo integration. The PF algorithm works by approximating the filtering density such that the sampling of $\mathbf{L}_t$ is efficient. The application of the PF only requires two assumptions: (A1) the state evolution $\mathbf{L}_t$ can be simulated from transition density $P\left(\mathbf{L}_t \mid \mathbf{L}_{t-1}\right)$; and (A2) the likelihood of R_t conditional on $\mathbf{L}_{t-1}$ and R_{t-1} can be exactly evaluated.

Given (3.4) and the GARCH dynamic (3.5), our objective is to filter out z_t from the return R_t. There are two Lévy innovations here: a normal part and a pure jump part. Thus, the filtering problem for z_t is equivalent to the filtering problem for y_t. This follows from the fact that

$$E\left[z_t \mid \Omega_t\right] = R_t - r_t - \mu_z h_{z,t} - \mu_y h_{y,t} - E\left[y_t \mid \Omega_t\right].$$

Following (5.3), the filtering density for the pure jump component can be written as

$$P\left(\ y_t \mid \Omega_t\right) \propto P\left(R_t \mid y_t,\ \Omega_{t-1}\right) P\left(y_t \mid \Omega_{t-1}\right).$$

The PF algorithm involves two recursive steps. Step 1 consists of simulating N particles of y_t^i from $P\left(y_t \mid \Omega_{t-1}\right)$, and step 2 involves resampling these particles proportional to the probability weights

$$w_t^i = P\left(R_t \mid y_t^i,\ \Omega_{t-1}\right).$$

These resampled $\left\{y_t^*\right\}^i$ are now approximately distributed according to $P\left(y_t \mid \Omega_t\right)$. Given this, the expectation of y_t and z_t can be trivially computed. The use of PF to filter out the pure jump part offers two remarkable advantages. First, the density function of y_t does not have to be analytical, as long as it can be simulated efficiently. This is extremely useful because a large number of Lévy processes only have analytical expressions in their characteristic function. Several Lévy processes such LS, Meixner, and CGMY do not have analytical density function but can be simulated through a robust algorithm. In fact, the Lévy-Khintchine formula (2.2) shows that all pure jump Lévy processes can be modeled as several tiny Poisson jumps. See

Schoutens (2003) for an overview of simulation procedures. The second advantage is in the resampling stage. It is extremely simple because the resampling weights w_t^i are computed using the normal density function.

5.3 Maximum likelihood estimation

The conditional density of the returns process (3.4) is given by

$$P(R_t|\ \Omega_{t-1}) = \int P(R_t \mid y_t,\ \Omega_{t-1})\, P(y_t \mid \Omega_{t-1})\, dy_t.$$

This integration cannot always be solved analytically. We then turn to the method of Monte Carlo integration. Pitt (2002) and Gordon, Salmon and Smith (1993) show that the above integration is just the mean of non-normalized resampling weights w_t^i computed in the PF resampling stage (step 2). The model parameters can be estimated by maximizing the Maximum Likelihood Importance Sampling (MLIS) criterion function:

$$MLIS = \sum_{t=0}^{T} \log\left(\frac{1}{N}\sum_{i=1}^{N} w_t^i\right), \tag{5.4}$$

where N is the number of particles used for PF, and T is the number of return periods. It is easy to see that when N becomes increasingly large, MLIS approaches the true likelihood function. For improved efficiency, we can apply techniques such as the smooth resampling of Pitt (2002), and the auxiliary particle filter (APF) of Pitt and Shephard (1999).

6 Empirical results

6.1 Data and Methodology

We investigate various two-factor Lévy GARCH models consisting of a Normal and a Lévy jump component. We estimate and test our models using two different approaches. The first approach involves MLE estimation on daily S&P 500 index returns from 1985 to 2005. After, we risk neutralize these MLE estimates and value the index call options. The second estimation approach involves joint estimation using an extensive data set of S&P 500 index options and returns (1996-2005). In practice, the task of fitting a model to option prices alone is numerically demanding. Estimating a model based on returns and options data jointly further adds to the econometric challenge. Existing studies that implement joint estimation therefore use data covering short time periods, or spanning small subsets of the cross sections of options.[11] Our ability to conduct this extensive joint estimation clearly underlines the numerical advantage of the Lévy GARCH framework. In addition, the joint estimation allows

[11] This includes Chernov and Ghysels (2000), Pan (2002), and Eraker (2004).

us to provide a more detailed study of long-run risk premia. This is of great economic importance because these risk premia explain the differences between the physical and risk-neutral distributions of asset prices.

The return data from January 1985 to December 2005 are obtained from CRSP. Our period includes the crash of October 1987, when the index falls by almost 25 percent in a single day. We retrieve the S&P 500 index call option quotes for the period 1996-2005 from OptionMetrics and eliminate quotes that report zero trading volume. Subsequently, we apply the filters proposed by Bakshi, Cao and Chen (1997) to the data. We only keep Wednesday options with maturities of more than one week and less than a full calendar year. We choose Wednesday because it is the least likely day to be a holiday, and it is less likely to be impacted by day-of-the-week effects. For further discussion of the advantages of Wednesday data, see Dumas, Fleming and Whaley (1998). Table 3 presents descriptive statistics for the option quotes by moneyness and maturity. The shape of the volatility smirk is evident from Panel C across all maturities, with short term options exhibiting the steepest volatility smirk.

6.2 MLE based on daily returns

We estimate the models using a time series of S&P 500 daily returns from January 1985 to 2005. We use this sample because it is sufficiently long enough to obtain precise estimates, and the interval also covers the ten-year period of our option data. We model the log return dynamic under the physical measure as

$$R_{t+1} = r_{t+1} + \mu_z h_{z,t+1} + \mu_y h_{y,t+1} + z_{t+1} + y_{t+1}.$$

This equation is identical to (3.4), but we present it here again for convenience. We model y_{t+1} using two different pure jump Lévy processes: the finite-activity Merton jump (MJ), and the infinite-activity Normal Inverse Gaussian jump (NIG). The MJ process has been extensively studied in the option pricing literature in the context of jump-diffusion models.[12] The NIG process of Barndorff-Nielsen (1998) is also well-known and has been applied to price American options in Stentoft (2007) and fit returns and volatility dynamics in Forsberg and Bollerslev (2002). These two processes have very different jump structures. The number of jumps arriving over any interval is finite for MJ, while for the NIG it is infinite. We will refer to the return dynamic in (3.4) with the MJ jump as the MJ-LGARCH, and with the NIG jump as the NIG-LGARCH.

In addition to two different jump processes, we assume that the dynamic of the time-homogeneous parameter $h_{y,t+1}$ of each jump type is either constant or state-dependent. Specif-

[12] See for instance, Bates (1996, 2000), Pan (2002), Andersen, Benzoni, and Lund (2002), Eraker (2004), Chernov, Gallant, Ghysels, and Tauchen (2003), Broadie, Chernov, and Johannes (2007).

ically, we study the following cases

$$\text{LGARCH}(1): \ h_{y,t+1} = k \quad \text{and} \quad \text{LGARCH}(3): \ h_{y,t+1} = k h_{z,t+1}$$

This classification of the dynamic for $h_{y,t+1}$ is inspired by Huang and Wu (2004) who consider four different specifications which exhaust all sources of heteroskedasticity in models based on jump and diffusion shocks. For brevity, we focus on the two most common specifications in this paper. LGARCH(1) is the most common specification in the jump-diffusion literature. In fact, if we let y_{t+1} follow a MJ process, the MJ-LGARCH(1) becomes closely linked to the SVJ model in the continuous-time literature.[13] The specification of LGARCH(3) has the intuitive feature that jumps arrive at a rate proportional to the risk associated with the normal shock to returns. Therefore jumps will arrive at higher frequency in high volatility periods.

We assume a simple affine GARCH(1,1) dynamic for the variance of the normal component $h_{z,t+1}$. Following our example in section 3.3, the LGARCH(1) is a special case of (3.5) when $h_y = w_y$, $b_y = 0$, $a_y = 0$, and $c_y = 0$. Similarly, the LGARCH(3) is nested as a special case of (3.5) with $w_y = w_z k$, $b_y = b_z$, $a_y = a_z k$, and $c_y = c_z$. In addition to the four different models above (two jump processes plus two different dynamics for $h_{y,t+1}$), we also estimate the benchmark Heston-Nandi GARCH(1,1) model. We refer to this model as the HN-LGARCH which is a special case of our return process (3.4) without the jump component. In all of our estimations, we use the method of variance targeting. This method reduces the number of parameters by one, and ensures consistency in the variance level across all models. For simplicity, we also assume a constant 5% risk-free rate.

6.2.1 Discussion of the MLE estimates

Table 4 reports the results from MLE estimation of the five different models that we consider. The results are based on the MLIS method using the PF algorithm with 10,000 particles. Our most parsimonious model is the HN-LGARCH which has four parameters compared to seven or eight in the models with jumps. Note that there is one less parameter in the LGARCH(3) model because it is not possible to econometrically identify μ_z from μ_y using only information from returns. This is because the mean return for the LGARCH(3) model is given by

$$\mu_z h_{z,t+1} + \mu_y h_{y,t+1} = \left(\mu_z + \mu_y k\right) h_{z,t+1}.$$

We therefore only report the estimate of $\mu = \left(\mu_z + \mu_y k\right)$ for the LGARCH(3) models. Comparison of the log likelihood values indicates that the HN-LGARCH performs poorly relative to the other four models which contain an additional jump component. Based on the log likelihood values in Table 4, we arrive at two conclusions. The first is that there is an ad-

[13] For empirical studies of the SVJ model, see Bakshi, Cao and Chen (1997), Bates (1998), Bates (2000), Pan (2002), Andersen, Benzoni and Lund (2002), and Eraker, Johannes and Polson (2003).

ditional benefit of using infinite-activity jumps to fit the return of the S&P 500 index. This finding is consistent with Li, Wells and Yu (2006) who estimate continuous-time models with Lévy jumps. We conjecture that this increase in likelihood is due to the more flexible jump structure that is inherent in the infinite-activity jump processes.

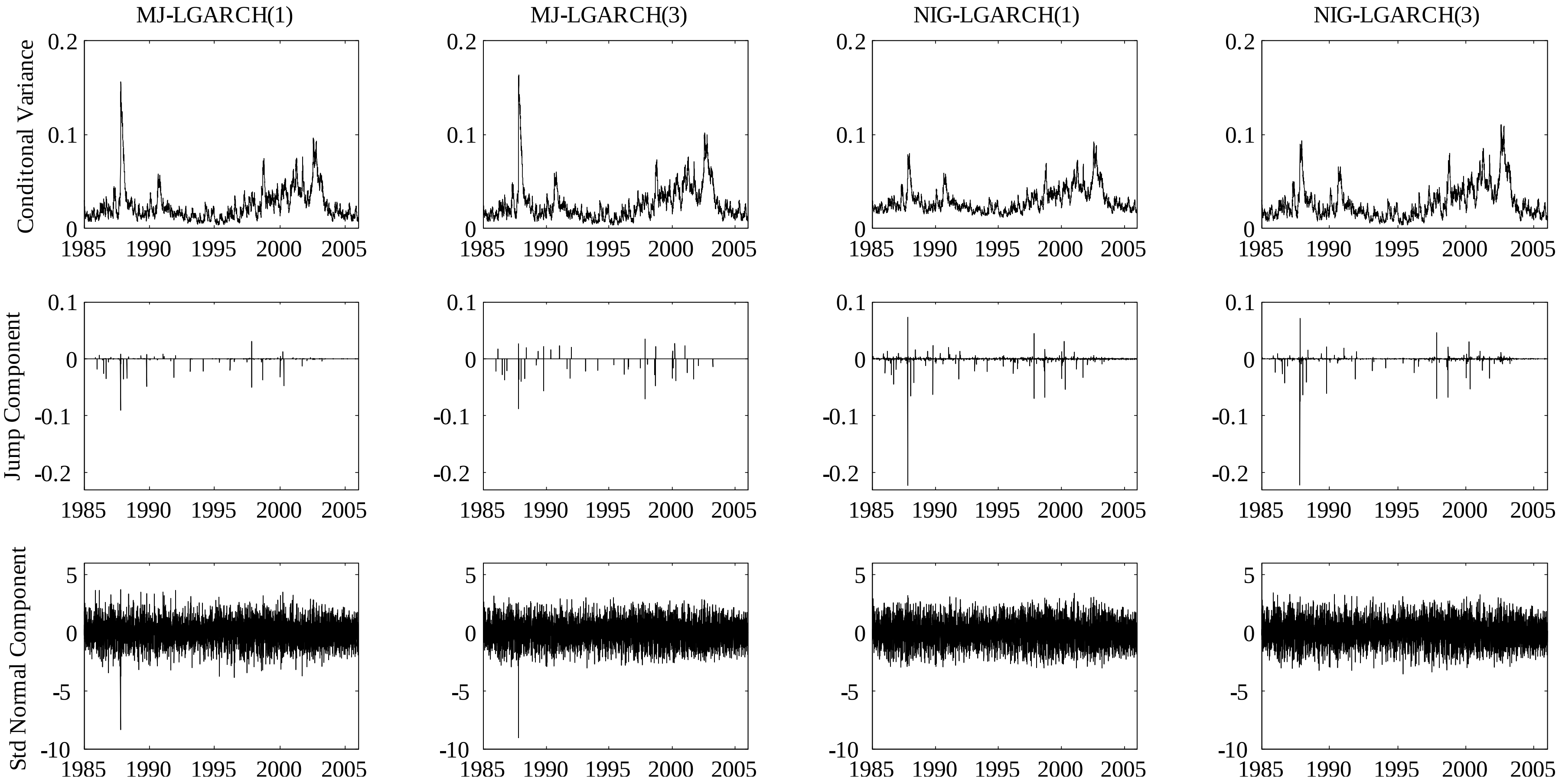
Our second conclusion concerns the importance of the time variation in the jump structure. This is equivalent to the importance of subordinating a pure jump Lévy process with a stochastic time change. Our findings indicate that, while it is important to have time-varying dynamic in the finite-activity Merton jump process, this is not necessarily true for the infinite-activity NIG jump process. This result significantly contributes to the literature on time-changed Lévy processes. To our knowledge, existing studies do not investigate the importance of "time change" in the infinite-activity jump process. Bates (2008) estimates daily returns using a time-changed CGMY process. However, his study does not look at the marginal importance of the time change effect in the infinite-activity Lévy jump processes.

We explain our second conclusion as follows. Infinite-activity jump processes are constructed based on many tiny jumps which arrive according to a stochastic Poisson intensity.[14] Therefore, the time change effect is already built into the structure of these jumps. This is especially true for the NIG process which can be constructed from evaluating Brownian processes at stochastic time intervals according to the IG process. Therefore, additional time change effects on the NIG process will be of second-order importance in the presence of stochastic volatility in the Brownian component of returns.Our finding on the importance of time-varying jump arrival rates (or jump intensities) is supportive of the work of Bates (2006), and Christoffersen, Jacobs and Ornthanalai (2008).

The results from particle filtering based on the maximum likelihood method are shown in Figure 1. We plot the conditional return variance of our four LGARCH models in the top panels. To save space, we exclude the analysis of HN-LGARCH for this figure. The middle and bottom panels show the time series of filtered jump and standardized normal components of the daily returns. The conditional variance from the NIG-LGARCH is very different from the MJ-LGARCH around the 1987 crash, as the conditional variance of the MJ-LGARCH increases significantly during this period. On the other hand, the spike in the conditional variance implied by the NIG-LGARCH during the crash of 1987 is at the same level as in the dot-com bubble collapse in 2000 and the post 9/11 period. Although this may seem strange, it must be noted that crashes are usually associated with negative skewness and large kurtosis. Hence, the evidence for crashes is not necessarily reflected in the second return moment.

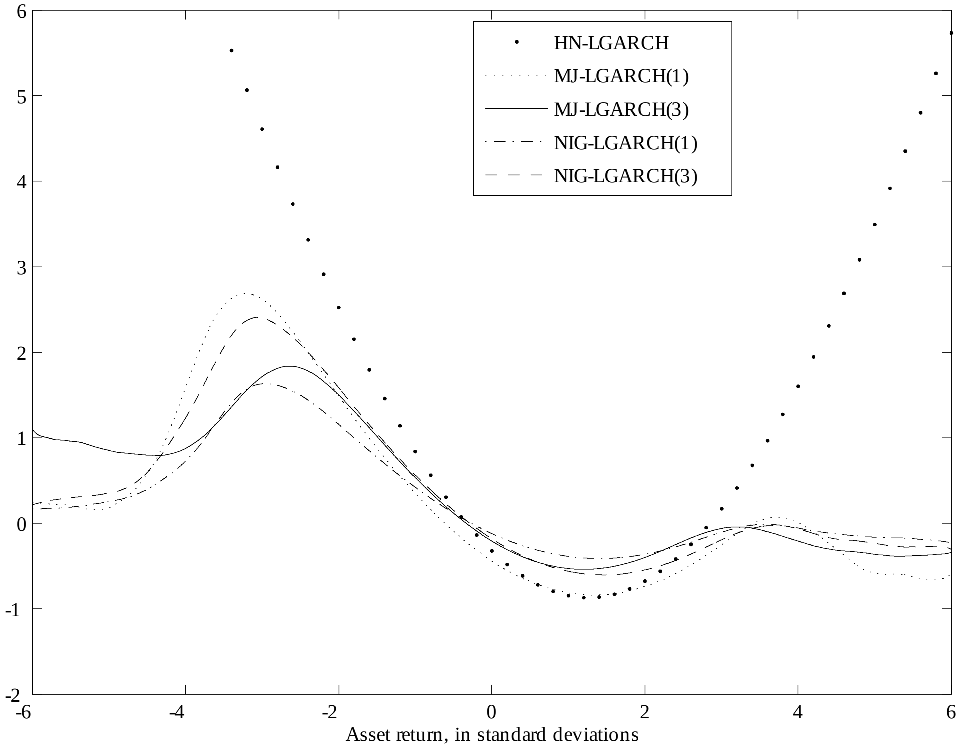
[14] This follows from the Lévy-Khintchine theorem. See equation (2.2) for more intuition.

Figure 1: Decomposition of S&P 500 returns for the MJ-LGARCH and NIG-LGARCH



Notes to figure: The top panels plot the conditional annualized variance of daily returns for the four jump models. Using the MLE estimates in Table 4, and in using the PF algorithm with 10,000 particles, we decompose the jump and the normal components of returns. The middle panels plot the time series of jumps in returns. The bottom panels plot the standardized normal component of returns. The plots in each row have identical scale for ease of comparison.

Figure 2: News impact curves for various LGARCH models



Notes to figure: We plot the news impact curve for various LGARCH models based on the parameters from Table 4. The plots illustrate, for each model, how the current period's return residuals (news impact) conditionally affect the volatility of returns on the next period. The y-axis represents the percentage change in the annualized return volatilities $\left(\sqrt{Var_t\left(R_{t+1}\right)} - \sqrt{Var_{t-1}\left(R_t\right)}\right)\sqrt{252} \times 100$. The x-axis represents standardized returns of magnitude $R_t/\sqrt{Var_{t-1}\left(R_t\right)}$. For all models, we assume that the current volatility is equal to the model's implied long-run volatility level

The benefit of using infinite-activity jump processes is also evident from the middle and bottom panels of Figure 1 where we decompose the jump and normal shocks from daily returns. We standardize the normal shocks with their conditional variance to illustrate misspecification. It is evident that the MJ-LGARCH model has difficulty modeling the crash of 1987. The NIG-LGARCH model handles this much better by letting the jump component explains almost the entire 25 percent drop in the index return.

6.2.2 The news impact curve

Bates (2006, 2008) correctly argues that standard GARCH models (i.e., HN-LGARCH) generate excessively large estimates of conditional variance after large stock market movements. Our MJ- and NIG-LGARCH models, however, do not suffer from this problem. This is because only a fraction of the return residual enters into the GARCH updating scheme. Therefore, the conditional variance predicted by our models will not necessarily be a monotonic function of the stock return movement. Figure 2 illustrates this property of our MJ- and NIG-LGARCH models using the "news impact curve", which is commonly used for interpreting the differences between volatility models. This method was introduced by Schwert (1990), and later christened by Engle and Ng (1993). All models produce news impact curves that are asymmetric, with negative returns having a larger impact on volatility updating than positive returns. This feature is often referred to as the leverage effect. The volatility level of the LGARCH models do not increase significantly after the days with moderately good news (positive returns). This property is consistent with the evidence in Chen and Ghysels (2008) who examine the news impact curve of high frequency returns data using semi-parametric MIDAS regressions. The news impact curve of the HN-LGARCH model is clearly distinct from the rest. The volatility revisions are monotonically linked to the magnitude of asset returns. Therefore, large returns, and especially the negative ones, will drive the conditional volatility up excessively high.

6.2.3 Discussion of the option pricing performance

Using the estimates from Table 4, we risk-neutralize our models using the framework described in section 4, and price S&P 500 call options from January 1996 to December 2005. For a fair comparison across all models, we assume a zero equity premium level when pricing these options. This is accomplished by setting $\lambda_z = 0$ and $\lambda_y = 0$ (see equation (4.5)). We fix the market price of risk parameters because their estimates can be imprecise when estimated using returns data.

Using a zero equity premium assumption, the risk-neutral parameters (except the market prices of risks) are equal to their MLE estimates from the physical measure. At each time period t, we compute call option prices

$$C_j = C\left(K_j, \tau_j, S_j, r_j, \Theta^{\mathbb{Q}}, h_z(t+1)\right)$$

by applying (4.10). The parameters K_j, τ_j, r_j, and S_j are the strike, maturity, risk-free rate, and the underlying index level associated with each j^{th} option. Note that in all LGARCH models considered here, each call price is also a function of the model's risk-neutral parameters $\Theta^{\mathbb{Q}}$, and the variance of the normal component $h_z(t+1)$. The daily time series of $h_z(t+1)$, for $t = 0, \ldots, T$, is filtered from an affine GARCH dynamic under the physical measure, and then matched to each j^{th} option traded on day t. We report the pricing errors under the implied volatility root-mean-squared error (IVRMSE) metric. This is an adopted standard in the option pricing literature, and the benefits of this loss function is discussed in Broadie, Chernov, and Johannes (2007). The computation of IVRMSE is as follows. We first compute the implied volatility of each call option price using the Black-Scholes formula. This gives us the Black-Scholes implied volatilities $IV\left(C_j, K_j, \tau_j, S_j, r_j\right)$. The IVRMSE is then computed as the square root of

$$\frac{1}{N}\sum_{j=1}^{N}\left(\sigma_j^{BS}-IV\left(C_j,K_j,\tau_{j,}S_{j,}r_j\right)\right)^2,$$

where σ_j^{BS} is the Black-Scholes implied volatility of the j^{th} observed call price, and $N=21,718$ is the total number of option contracts used in the analysis.

We report option IVRMSE next to the log likelihood value in Table 4. For the HN-LGARCH, we report the annualized IVRMSE (%), while for the other four models, we report the ratios of their IVRMSE relative to the HN-LGARCH model. The results show that MJ-LGARCH models slightly improve on the benchmark HN-LGARCH. Nevertheless, MJ-LGARCH(3) performs significantly better than MJ-LGARCH(1). These findings support the work of Christoffersen, Jacobs and Ornthanalai (2008) who show that state-dependent jump intensity is important for option pricing. They also find that jumps of the Merton type cannot significantly improve option pricing performance unless sizeable jump risk premia are present, which is consistent with our results here because we set $\lambda_y=0$.

The NIG-LGARCH models with infinite-activity jumps perform particularly well at pricing options. The improvement of 12 to 15 percent is quite remarkable considering that these parameters' estimates are not fitted to option data. Similar to the result from the Merton jump models, we see that LGARCH(3) is preferable to LGARCH(1) from an option pricing perspective. Therefore, a time-varying dynamic (or stochastic time change) in the infinite-activity jump process seems beneficial for the pricing of derivatives. This is not surprising as the conditional skewness and kurtosis play significant roles in generating the smirk effect observed in the implied volatility curve. Therefore, models with richer specification in the higher moments such as the infinite-activity jumps may be more preferable than the finite-activity jump processes.

We provide additional evidence on the pricing performance of various LGARCH models in Table 5 where IVRMSEs are compared across three dimensions: moneyness, maturity, and the VIX index level. For the HN-LGARCH model, we report the annualized IVRMSE(%) while for the others, we report their ratio IVRMSEs relative to the HN-LGARCH. Table 5 shows that option pricing improvement of the NIG-LGARCH model relative to the HN-LGARCH is robust across moneyness, maturity, and the market variance (VIX level). The evidence is quite strong, especially for the NIG-LGARCH(3) model which outperforms the HN-LGARCH model by more than 10 percent at all levels of moneyness and maturity.

6.3 Joint MLE based on options and returns

A good option pricing model must come with an economically justifiable assumption on the pricing kernel. Besides fitting options data, the models must produce daily returns that are consistent with the dynamic of the underlying asset under the physical measure. Unfortunately, most empirical option pricing studies are primarily interested in minimizing option pricing errors. Consequently, much is known about the risk-neutral dynamic of the asset price, while little is known about how these underlying risk factors are priced in the model. In order to recover risk premia, we employ an objective function that allows us to anchor our parameters in both the physical and the risk-neutral measures. We estimate our parameters using the joint MLE of weekly index options on each Wednesday from 1996 to 2005, as well as daily S&P 500 index returns for 1995-2005. We use one year of daily returns prior to the option sample in order to initialize the path of the volatility filter. The total log likelihood, which is our objective function, can be written as

$$LogLkhood = L^{\mathbb{P}}_{returns} + L^{\mathbb{Q}}_{options}. \tag{6.1}$$

The construction of the log likelihood from daily index returns $L^{\mathbb{P}}_{returns}$ is given by equation (5.4), and is identical to the previous section. For the log likelihood of options $L^{\mathbb{Q}}_{options}$, we assume the following data generating process

$$\sigma_j^{BS} = IV\left(C_j, K_j, \tau_j, S_j, r_j\right) + \varepsilon_j, \tag{6.2}$$

with $\varepsilon_j \sim Normal\left(0, \sigma_\varepsilon^2\right)$. Because the residuals of the implied volatility pricing error are assumed to be normally distributed, the construction of $L^{\mathbb{Q}}_{options}$ is straightforward. We estimate our models by maximizing the joint log likelihood (6.1) with respect to the structural parameters in $\Theta^{\mathbb{P}}$. The structural parameters in the joint MLE of MJ-LGARCH and NIG-LGARCH models are

$$\Theta^{\mathbb{P}}_{MJ-LGARCH} = \{\lambda_z, \lambda_y, b, a, c, \theta, \delta, k\} \text{ and } \Theta^{\mathbb{P}}_{NIG-LGARCH} = \{\lambda_z, \lambda_y, b, a, c, \alpha, \beta, k\}.$$

It is important to note that all parameters which are required for the GARCH filtration and the pricing of options are identified in $\Theta^{\mathbb{P}}$. This is because the risk-neutral parameters are endogenously determined by the market price of risk parameters λ_z and λ_y. In addition, these market prices of risk show up in the equity premium equation (4.5), which determines the gross rate of realized return in the physical measure. Our joint estimation therefore imposes consistency between the two probability measures. Using the results in Table 2, and the risk-neutral reparametrization of the affine GARCH(1,1) dynamic in the appendix D, we can determine the risk-neutral parameters for our models. For convenience, we explicitly show how these parameters are endogenously linked to their physical measure parameters in appendix E.

We focus our empirical analysis on LGARCH(3) models. We exclude the joint MLE of LGARCH(1) models because our from the previous section show that a time-varying dynamic in the jump component is important for option pricing. Joint MLE on options and returns is, therefore, conducted on three models: MJ-LGARCH, NIG-LGARCH, and the benchmark

HN-LGARCH.

6.3.1 Discussion on the estimates from joint MLE

Table 6 reports the results from joint MLE of the three models considered in this section. First, we underline the consistency and the stability of the structural parameters that govern the GARCH dynamic and the jump innovation by comparing our estimates in Table 6 to the results in Table 4. Interestingly, the differences are minimal. In theory, these two sets of parameters should not differ as they describe the dynamic of the same underlying asset under the physical measure.

All parameters are statistically significant, including the market price of risk parameters. For models with jumps, we can identify the market price of jump risk separately from the market price of risk associated with the normal component of returns. Table 6 also reports the long-run total equity premium implied by each model in annualized percentage terms. The premia are computed by taking unconditional expectations of equation (4.5). This results in

$$\overline{\gamma} = \lambda_z \sigma_z^2 + \lambda_y \sigma_y^2,$$

where $\sigma_z^2 = E\left[h_{z,t+1}\right] = (w+a)/\left(1-b-ac^2\right)$ is the model's long-run implied variance of the normal component of returns. Similarly, $\sigma_y^2 = E\left[h_{y,t+1}\right]$, and it follows that $\sigma_y^2 = k\sigma_z^2$, because of the LGARCH(3) specification. In the MJ-LGARCH and NIG-LGARCH models, the equity premium is composed of two risk factors. We interpret $\lambda_z \sigma_z^2$ as the long-run normal risk premium, and $\lambda_y \sigma_y^2$ as the long-run jump risk premium.

Judging the performance of our three models based on the log likelihood values shows that the NIG-LGARCH model is far superior to the MJ-LGARCH model, and also to the HN-LGARCH model. This evidence supports the finding of Huang and Wu (2004) and Carr and Wu (2003a) who find that models with infinite-activity jump structures are better at fitting option prices than models with finite-activity Merton jumps. The superior option pricing performance of the NIG-LGARCH model is also reflected in the IVRMSE shown in Table 6. Somewhat surprisingly, the MJ-LGARCH model only outperforms the HN-LGARCH model by 3% using the IVRMSE metric. A similar finding is reported in Eraker (2004), who conducts joint estimation using the MCMC technique on S&P 500 options and returns data from January 1987 to December 1990. He finds that models with Merton jumps do not improve on the Heston (1993) stochastic volatility model, based on option pricing errors. Bates (2000) arrives at a similar conclusion. Utilizing cross sections of options from 1988 to 1993, he imposes consistency between the underlying returns and option prices. He reports a 2% improvement in option pricing from jump models relative to Heston (1993). Our finding regarding the option pricing performance of jump models differs from that in Broadie, Chernov and Johannes (2007). They find that jump models do improve on the Heston (1993) model by up to 50% based on the IVRMSE metric. However, we note that their approach is different from ours, because they do not estimate options and returns jointly. In addition, they recover the variance path by minimizing the IVRMSE, instead of inferring it from the underlying asset return dynamic. Hence, they do not impose the full consistency between the two probability measures in their estimation.

6.3.2 The economic role of jumps in option pricing models

If standard Merton-type jumps cannot significantly reduce option pricing errors, then what is their role from an economic standpoint? To answer this, we study the equity premium level implied by each LGARCH model. From Table 6, we see that the equity premium implied by the HN-LGARCH model is about 23%, which is unrealistically large. The literature on the estimate of equity premium is too extensive to cite in full here. However, most estimates in the literature vary between 3% and 10%. It is therefore difficult to justify the equity premium level implied by the HN-LGARCH model.

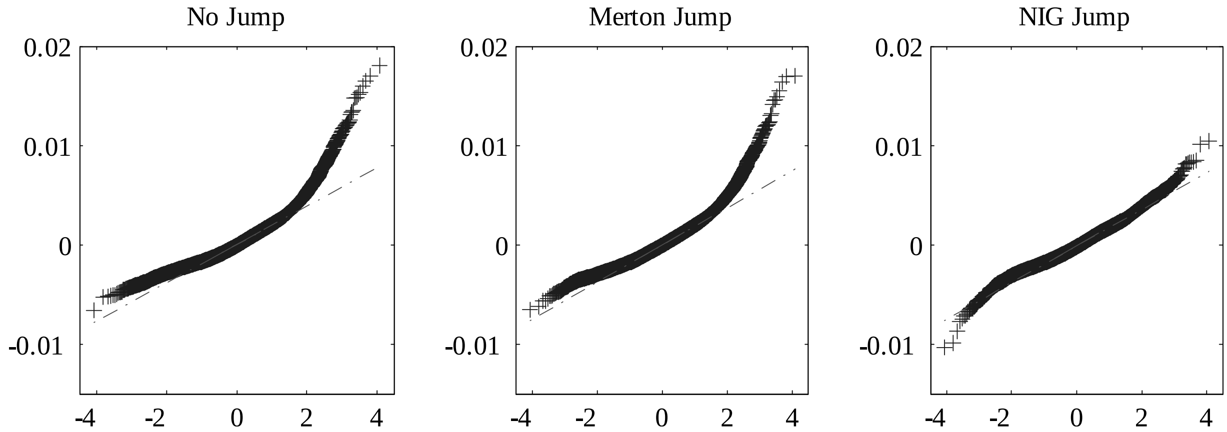
The equity premium levels implied by the MJ- and NIG-LGARCH models are 8% and 6.3% respectively. These values are much more economically plausible, and also fall within the range of estimates reported in the literature. Our results therefore illustrate the importance of jump risk factors in option pricing models. Without the jump component, the divergence between the two probability measures cannot be explained with an economically justifiable equity premium level. Bates (2000) also points out that stochastic volatility models require extreme parameters that are implausible given the time series of option prices. On the other hand, jump-diffusion models imply more plausible estimates of the volatility process parameters. Our joint MLE results therefore support the conclusion in Bates (2000). However, we are the first to formally link the differences between physical and risk-neutral measure parameters to the implied equity premium.

Our joint estimation of jump models also provides additional evidence regarding the factor risk premia. Existing studies diverge on estimates of jump and diffusive volatility risk premia. We note that the normal risk premium in our model is closely linked to the volatility risk premium in the continous-time literature in a subtle way. This is because the affine GARCH(1,1) dynamic assumes that normal innovations in returns and GARCH variance dynamic are perfectly correlated. Our estimate is roughly consistent with Pan (2002), who documents return risk premia of 3.5% and 5.5% for the jump and diffusive risks, respectively. Nevertheless, her estimate of the diffusive risk premium is not statistically significant from zero. Eraker (2004) reports significant jump and marginally significant volatility risk premia only in the most complex jump specifications that he considered.[15] Nevertheless, his estimates of the risk premia are small, and it is difficult to interpret them economically in term of the expected return risk

[15] The most complex specification considered in Eraker (2004) is the SVSCJ model. It is closely related to the MJ-LGARCH(3) specification in this paper with the addition of correlated jumps in volatilities and returns.

premia. Because our data set covers such a long time period and spans rich cross sections of options, we are able to obtain more precise estimates of the jump and normal risk premia. Depending on the model, we find that investors demand approximately 3% (MJ-LGARCH) to 5% (NIG-LGARCH) on average per year, in excess return, for bearing the market jump risk. To infer the magnitude of the normal risk premium, we subtract the jump risk premium from the total equity premium in each model. This yields annual return premium for bearing the market normal risk of approximately 5% (MJ-LGARCH) and 1.3% (NIG-LGARCH).

Figure 3: QQ-plots of implied volatility residuals from joint options and returns MLE

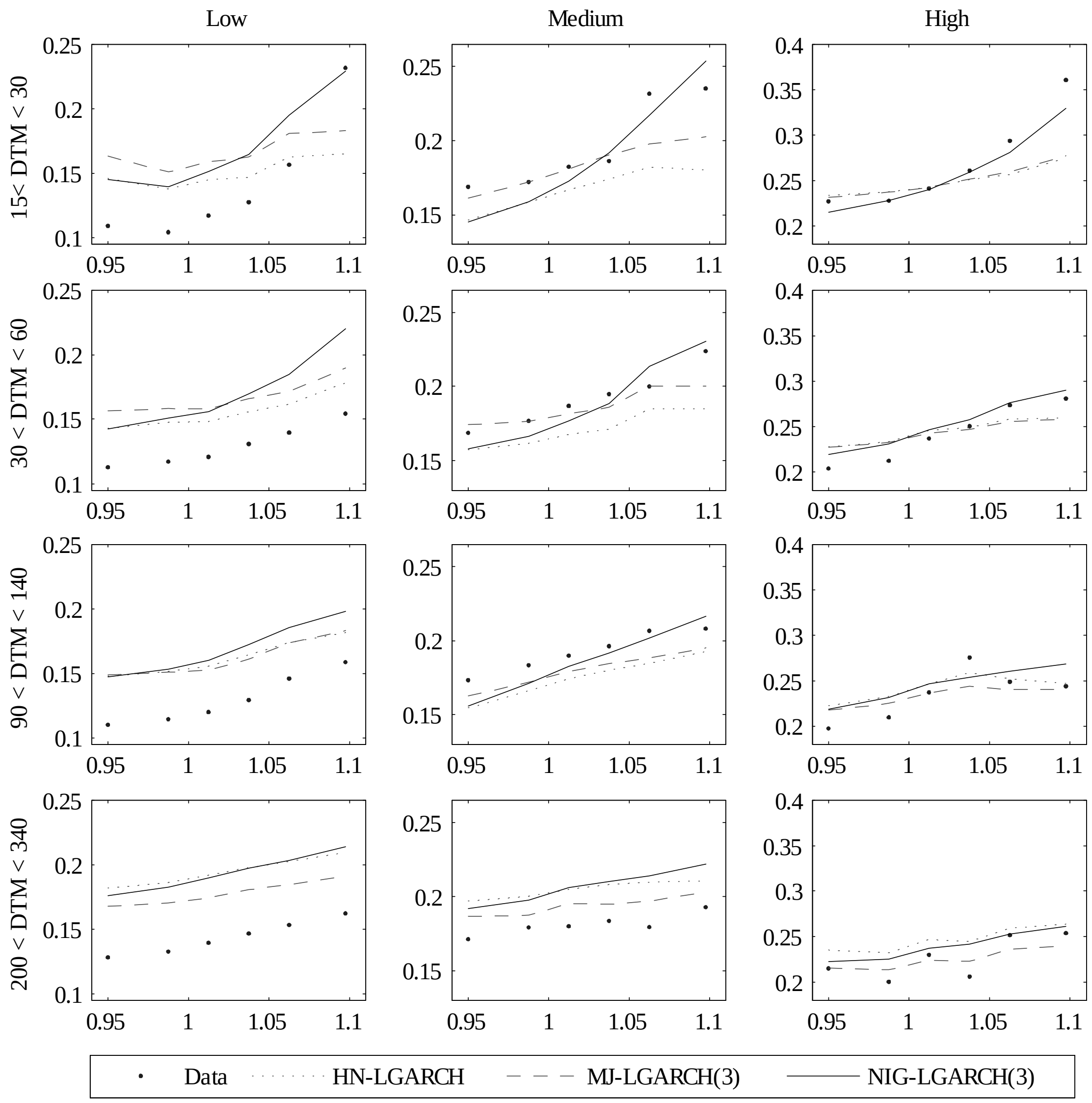


Notes to figure: We use the implied volatility residuals of each model computed using estimates in Table 6. We plot these implied volatility residuals against the normal quantiles (QQ-plot), which is represented on the x-axis. From left to right, the panel shows the QQ-plot for the implied volatility residuals of the HN-LGARCH model, the MJ-LGARCH(3) model, and the NIG-LGARCH(3) model, respectively. The dash-dot line in each plot represents the case of a normal density.

6.3.3 Further analysis on the option pricing performance

Figure 3 shows the conventional QQ-plots of the implied volatility residuals from joint options and returns MLE. These plots show that the data generating process in (6.2) is misspecified. The misspecification in Figure 3 is the least severe for the NIG-LGARCH model. The bottom rows of Table 6 present the statistical moments of the implied volatility residuals for each model. The violation of the nonnormality assumption is the most severe for the HN-LGARCH, with the MJ-LGARCH model a close second. The weakness of the standard Merton jump model is clearly seen in Figure 3 as it fails to capture the right tail of the distribution of the implied volatility's residuals.

Figure 4: Implied volatility smirks for various maturities and models



Notes to Figure: We use Black-Scholes implied volatilities computed from the joint MLE and plot implied volatility smirks of the data and the models for three different periods: a low volatility period, a medium volatility period and a high volatility period. Low volatility period is between 2005/01/01 to 2005/04/31. Medium volatility period is between 1997/03/01 to 1997/06/31, when the average VIX level is 19.98. High volatility period is between 2001/06/01 to 2001/10/31. The moneyness is on the horizontal axis and each row of panels corresponds to a different maturity.

We next look at the ability of each model to produce the well-known smirk effect. Figure 4 plots, for each of the models, the implied volatility smirks across moneyness for four different maturity buckets: 15-30, 30-60, 90-140, and 200-340 days to maturity. To study the importance of different volatility regimes, we present our analysis for three different volatility periods. The first sample period is from January 1, 2005 until April 30, 2005. The average value of the VIX over this period is 13.22%, and we refer to it as a low volatility period. The second sample period is from March 1, 1997 until June 30, 1997. This is when the average VIX value is about 19.98%, and we deem it a medium volatility period. The third sample period is from June 1, 2001 until October 30, 2001, which is a high volatility period. The average value of the VIX index is 26.11% in this period. The main conclusion from Figure 4 is that the NIG-LGARCH model produces smirks with the most realistic slopes. All models, however, produce roughly the same level of implied volatility. Overall, Figure 4 confirms the superiority of the infinite-activity jump process (NIG) over the finite-activity Merton jump process for the purpose of option valuation

7 Conclusions

In this paper, we develop a rich class of discrete-time models for asset pricing. We use affine GARCH dynamics to drive the heteroskedasticity in our models, and we show that our entire framework produces affine conditional transforms of asset returns. We choose affine GARCH dynamics to model heteroskedasticity in asset returns because they admit tractable formulae for many securities of interest, they can capture the leverage effect, and they are relatively simple to implement. Due to the tractability of the conditional transform, the price of zero-coupon bonds is known analytically, and European-style derivatives can be priced via Fourier inversion.

In addition to the theoretical development of this framework, we also suggest a systematic approach for estimating these models based on the maximum likelihood methodology. Our discrete-time framework produce models that converge to time-changed Lévy processes in the continous-time limits. However, our models are more simple to implement than their continous-time counterparts due to the aid of GARCH dynamics. This suggests various possible directions of future empirical research, such as conducting a large-scale specification analysis of Lévy GARCH models using returns data similar to Bates (2008), and implementing the Lévy GARCH framework to model credit default swaps.

Using a comprehensive data set of options and returns in the joint MLE, we uncover the important economic role of jumps. We find that, without a jump component, the divergence between the physical and risk-neutral probability measures cannot be explained with an economically justifiable equity premium level. Finally, we note that future research in which models are estimated from options and returns data jointly could lead to a better understanding of how risk factors are priced in the market. This is of great significance because risk premia are the economic fundamentals in the pricing of all financial assets.

Appendix

7.1 A: Proof of proposition 1

The solution to the generating function is achieved by exploiting the affine structure of our setup. We start by assuming that the generating function of asset price at any future period $t+k$, conditional on time t, has the following exponential affine form

$$\begin{aligned} f(\phi;t,t+k) &= E_t\left[S_{t+k}^{\phi}\right] = S_t^{\phi} E_t\left[e^{\phi\sum_{i=1}^{k} R_{t+i}}\right] \\ &= S_t^{\phi} e^{\mathcal{A}(\phi;t,t+k)+\sum_{i=1}^{m}\mathcal{B}_i(\phi;t,t+k)'\mathbf{h}_{t+2-i}+\sum_{j=1}^{n}\mathcal{C}_j(\phi;t,t+k)'\boldsymbol{v}_{t+2-j}}, \end{aligned} \tag{7.1}$$

where $\mathcal{A}(\phi;t,t+k)$ is a scalar, $\mathcal{B}_i(\phi;t,t+k)$'s are $d\times 1$ vectors, and $\mathcal{C}_j(\phi;t,t+k)$'s are $q\times 1$ vectors. For asset price with terminal period T, we are interested in finding the expressions for the affine coefficients

$$\mathcal{A}(\phi;t,T) \quad ; \quad \mathcal{B}_i(\phi;t,T)\text{'s} \quad ; \quad \mathcal{C}_j(\phi;t,T)\text{'s}$$

which solve the conditional generating function $f(\phi;t,T) = E_t\left[S_T^{\phi}\right]$. We proceed by using the property of iterated expectation

$$\begin{aligned} f(\phi;t,T) &= E_t\left[E_{t+1}\left[S_T^{\phi}\right]\right] = E_t\left[f(\phi;t+1,T)\right] \\ &= E_t\left[S_{t+1}^{\phi} e^{\mathcal{A}(\phi;t+1,T)+\sum_{i=1}^{m}\mathcal{B}_i(\phi;t+1,T)'\mathbf{h}_{t+3-i}+\sum_{j=1}^{n}\mathcal{C}_j(\phi;t+1,T)'\boldsymbol{v}_{t+3-j}}\right] \\ &= S_t^{\phi} E_t\left[e^{\phi R_{t+1}+\mathcal{A}(\phi;t+1,T)+\sum_{i=1}^{m}\mathcal{B}_i(\phi;t+1,T)'\mathbf{h}_{t+3-i}+\sum_{j=1}^{n}\mathcal{C}_j(\phi;t+1,T)'\boldsymbol{v}_{t+3-j}}\right] \end{aligned}$$

where, in the second line above, we have substituted the assumed solution to $f(\phi;t+1,T)$ according to (7.1). We also substitute in the dynamic of R_{t+1} from (3.1) into the above equation. After simplification and removing the nonstochastic component from the expectation operator, we have the following expression for $f(\phi;t,T)/S_t^{\phi}$

$$\begin{aligned} & e^{\phi r_{t+1}+\phi(\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))'\mathbf{h}_{t+1}+\mathcal{A}(\phi;t+1,T)+\sum_{i=2}^{m}\mathcal{B}_i(\phi;t+1,T)'\mathbf{h}_{t+3-i}+\sum_{j=2}^{n}\mathcal{C}_j(\phi;t+1,T)'\boldsymbol{v}_{t+3-j}} \times \\ & E_t\left[e^{\phi\boldsymbol{\vartheta}'\mathbf{X}_{t+1}+\mathcal{B}_1(\phi;t+1,T)'\mathbf{h}_{t+2}+\mathcal{C}_1(\phi;t+1,T)'\boldsymbol{v}_{t+2}}\right]. \end{aligned}$$

In order to solve the conditional expectation above, we make use of the affine GARCH definition in (3.2) and note that

$$\Pi = \left(\phi\boldsymbol{\vartheta}, \mathcal{B}_1(\phi;t+1,T), \mathcal{C}_1(\phi;t+1,T)\right).$$

This allows us to write the expression $f(\phi;t,T)/S_t^{\phi}$ in the exponential affine form with its exponent given by

$$\phi r_{t+1}+\phi\left(\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}\left(\boldsymbol{\vartheta}\right)\right)'\mathbf{h}_{t+1}+\mathcal{A}\left(\phi;t+1,T\right)+\sum_{i=2}^{m}\mathcal{B}_i\left(\phi;t+1,T\right)'\mathbf{h}_{t+3-i}+ \sum_{j=2}^{n}\mathcal{C}_j\left(\phi;t+1,T\right)'\boldsymbol{\upsilon}_{t+3-j}+\mathcal{V}\left(\Pi\right)+\sum_{i=1}^{m}\mathcal{W}_i\left(\Pi\right)'\mathbf{h}_{t+2-i}+\sum_{j=1}^{n}\mathcal{Y}_j\left(\Pi\right)'\boldsymbol{\upsilon}_{t+2-j}. \tag{7.2}$$

Using the assumed affine structure (7.1), the expression in the exponent of $f\left(\phi;t,T\right)/S_t^{\phi}$ can also be written as

$$\mathcal{A}\left(\phi;t,T\right)+\sum_{i=1}^{m}\mathcal{B}_i\left(\phi;t,T\right)'\mathbf{h}_{t+2-i}+\sum_{j=1}^{n}\mathcal{C}_i\left(\phi;t,T\right)'\boldsymbol{\upsilon}_{t+2-j}. \tag{7.3}$$

The expressions for $\mathcal{A}\left(\phi;t,T\right)$, $\mathcal{B}_i\left(\phi;t,T\right)$, and $\mathcal{C}_j\left(\phi;t,T\right)$ can now be solved by matching the coefficients of $\mathbf{h}_{t+2-i}$ and $\boldsymbol{\upsilon}_{t+2-j}$, between equations (7.2) and (7.3), for $i=1,\ldots,m$ and $j=1,\ldots,n$. This procedure will yield recursive relations for the affine coefficients as shown in Proposition 1. The boundary conditions for the affine coefficients are derived using the fact that $E_T\left[S_T^{\phi}\right]=S_T^{\phi}$.

7.2 B: Expressions for the affine coefficients

Heteroskedasticity via a GARCH(1,1) dynamic

The solutions for the affine coefficients in (3.6) are given by the following set of recursive relations

$$\begin{aligned}
\mathcal{A}\left(\phi;\ t,\ T\right) &= \phi r_{t+1}+\mathcal{A}\left(\phi;t+1,T\right)+\mathcal{B}^z\left(\phi;t+1,T\right)w_z+\mathcal{B}^y\left(\phi;t+1,\ T\right)w_y \\
&\quad -\frac{1}{2}\log\left(1-2\mathcal{B}^z\left(\phi;t+1,T\right)a_z-2\mathcal{B}^y\left(\phi;t+1,T\right)a_y\right) \\
\mathcal{B}^z\left(\phi;\ t,\ T\right) &= \phi\mu_z+\mathcal{B}^z\left(\phi;\ t+1,\ T\right)\left(b_z+a_z c_z^2\right)+\mathcal{B}^y\left(\phi;\ t+1,\ T\right)a_y c_y^2+ \\
&\quad \frac{\left(\phi-2\mathcal{B}^z\left(\phi;\ t+1,\ T\right)a_z c_z-2\mathcal{B}^y\left(\phi;\ t+1,\ T\right)a_y c_y\right)^2}{2(1-2\mathcal{B}^z\left(\phi;\ t+1,\ T\right)a_z-2\mathcal{B}^y\left(\phi;\ t+1,\ T\right)a_y)} \\
\mathcal{B}^y\left(\phi;\ t,\ T\right) &= b_y\mathcal{B}^y\left(\phi;\ t+1,\ T\right)+\phi\mu_y+\xi_y\left(\phi\right),
\end{aligned}$$

where the parameters are defined according to section (3.3.1) and $\xi_y\left(\phi\right)$ is the coefficient in the cumulant exponent of the jump process evaluated at ϕ.

Heteroskedasticity via a component GARCH dynamic

Assuming the return dynamic (3.6) together with the component GARCH model (3.7), the coefficients in the generating function (3.8) can be solved through the following recursive relations

$$\begin{aligned}
\mathcal{A}(\phi;t,T) &= \phi r_{t+1} + \mathcal{A}(\phi;t+1,T) \\
&- \sum_{x_i=\{z,y\}} \left(\mathcal{B}^{x_i}(\phi;t+1,T)\, a_{x_i} + \mathcal{C}^{x_i}(\phi;t+1,T)\left(w_{x_i} - \varphi_{x_i}\right)\right) \\
&- \frac{1}{2}\log\left(1 - \sum_{x_i=\{z,y\}} 2\mathcal{B}^{x_i}(\phi;t+1,T)\, a_{x_i} - \sum_{x_i=\{z,y\}} 2\mathcal{C}^{x_i}(\phi;t+1,T)\,\varphi_{x_i}\right) \\
\mathcal{B}^{z}(\phi;t,T) &= \phi\mu_z + \mathcal{B}^{z}(\phi;t+1,T)\, b_z - \\
&\frac{\left(\phi - 2\sum_{x_i=\{z,y\}}\left(\mathcal{B}^{x_i}(\phi;t+1,T)\, a_{x_i} c_{x_i} - \mathcal{C}^{x_i}(\phi;t+1,T)\,\varphi_{x_i} d_{x_i}\right)\right)^2}{2\left(1 - 2\sum_{x_i=\{z,y\}}\left(\mathcal{B}^{x_i}(\phi;t+1,T)\, a_{x_i} - \mathcal{C}^{x_i}(\phi;t+1,T)\,\varphi_{x_i}\right)\right)} \\
\mathcal{B}^{y}(\phi;t,T) &= \mathcal{B}^{y}(\phi;t+1,T)\, b_y + \phi\mu_y + \xi_y(\phi) \\
\mathcal{C}^{z}(\phi;t,T) &= \mathcal{B}^{z}(\phi;t+1,T)\,(b_z - 1) + \mathcal{C}^{z}(\phi;t+1,T)\,\rho_z \\
\mathcal{C}^{y}(\phi;t,T) &= \mathcal{B}^{y}(\phi;t+1,T)\,(b_y - 1) + \mathcal{C}^{y}(\phi;t+1,T)\,\rho_y.
\end{aligned}$$

Heteroskedasticity with jumps in a GARCH(1,1) dynamic

The solutions to the affine coefficients in (3.10) are identical to the simple GARCH(1,1) case for $\mathcal{A}(\phi;t,T)$ and $\mathcal{B}^{z}(\phi;t,T)$. However, the recursive relation for $\mathcal{B}^{y}(\phi;t,T)$ is different and is given by

$$\mathcal{B}^{y}(\phi;t,T) = b_y \mathcal{B}^{y}(\phi;t+1,T) + \phi\mu_y + \xi_y(\Xi),$$

where $\xi_y(\Xi)$ is the coefficient in the cumulant exponent for the jump innovation (with positive support) evaluated at

$$\Xi = \eta\phi + \mathcal{B}^{z}(\phi;t+1,T)\, d_z + \mathcal{B}^{y}(\phi;t+1,T)\, d_y.$$

C: Return dynamics under the risk-neutral measure for Lévy GARCH models

Proof of Proposition 2

The solution for $\mathbf{\Lambda}$ is solved by imposing a local martingale restriction $E_t^{\mathbb{Q}}\left[e^{R_{t+1}}\right] = e^{r_{t+1}}$. After applying the change of measure according to the Radon-Nikodym derivative (4.1), we have

$$\begin{aligned}
E_t^{\mathbb{Q}}\left[e^{R_{t+1}}\right] &= E_t\left[\exp^{\mathbf{\Lambda}'\mathbf{X}_{t+1} - \boldsymbol{\xi}_{\mathbf{X}}(\mathbf{\Lambda})'\mathbf{h}_{t+1} + R_{t+1}}\right] \\
&= E_t\left[e^{r_{t+1} + \mathbf{\Lambda}'\mathbf{X}_{t+1} - \boldsymbol{\xi}_{\mathbf{X}}(\mathbf{\Lambda})'\mathbf{h}_{t+1} + (\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))'\mathbf{h}_{t+1} + \boldsymbol{\vartheta}'\mathbf{X}_{t+1}}\right] \\
&= e^{r_{t+1} + \ (\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}) - \boldsymbol{\xi}_{\mathbf{X}}(\mathbf{\Lambda}))'\mathbf{h}_{t+1}} E_t\left[e^{(\boldsymbol{\vartheta} + \mathbf{\Lambda})'\mathbf{X}_{t+1}}\right].
\end{aligned}$$

Note that when taking the expectation, the exponent of the last line above has to equate to r_{t+1}. This leads us to the following result

$$(\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}) - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Lambda}))' \mathbf{h}_{t+1} + \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta} + \boldsymbol{\Lambda})' \mathbf{h}_{t+1} = 0,$$

which sufficiently proves the result in Proposition 2.

Proof of Lemma 1

We start by directly applying the Radon-Nikodym derivative (4.1) to the asset price dynamic

$$\begin{aligned} E_t^{\mathbb{Q}}\left[\frac{S_{t+1}}{S_t}\right] &= E_t\left[e^{r_{t+1}+\boldsymbol{\Lambda}'\mathbf{X}_{t+1}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Lambda})'\mathbf{h}_{t+1}+(\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))'\mathbf{h}_{t+1}+\boldsymbol{\vartheta}'\mathbf{X}_{t+1}}\right] \\ &= e^{r_{t+1}+\ (\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Lambda}))'\mathbf{h}_{t+1}} E_t\left[e^{(\boldsymbol{\vartheta}+\boldsymbol{\Lambda})'\mathbf{X}_{t+1}}\right] \\ &= e^{r_{t+1}+\ (\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))'\mathbf{h}_{t+1}+(\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}+\boldsymbol{\Lambda})-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Lambda}))'\mathbf{h}_{t+1}} \\ &= e^{r_{t+1}+(\boldsymbol{\lambda}-\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))'\mathbf{h}_{t+1}+\Psi_{\mathbf{X}}^{\mathbb{Q}}(\boldsymbol{\vartheta};\ t,\ t+1)}. \end{aligned} \tag{7.4}$$

Note that we have applied the result in Proposition (3) to the last line of the above equations, where

$$\Psi_{\mathbf{X}}^{\mathbb{Q}}(\boldsymbol{\vartheta}; t, t+1) = (\boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta} + \boldsymbol{\Lambda}) - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\Lambda}))' \mathbf{h}_{t+1} = \boldsymbol{\xi}_{\mathbf{X}}^{*}(\boldsymbol{\vartheta}) \mathbf{h}_{t+1}^{*},$$

which is the conditional cumulant exponent of the risk-neutral Lévy processes $\mathbf{X}_{t+1}^{*}$ evaluated at $\boldsymbol{\vartheta}$. The asset return under the $\mathbb{Q}-$measure can now be equivalently written as

$$\log\left(\frac{S_{t+1}}{S_t}\right)^{*} = r_{t+1} + (\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta}))' \mathbf{h}_{t+1} + \mathbf{X}_{t+1}^{*}.$$

We are left to prove that $(\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})) \mathbf{h}_{t+1}$ is equal to $-\boldsymbol{\xi}_{\mathbf{X}}^{*}(\boldsymbol{\vartheta}) \mathbf{h}_{t+1}^{*}$. To do this, we use the fact that under the equivalent martingale measure $\mathbb{Q}$, all assets have returns equal to the risk-free rate. Therefore, it follows from (7.4) that

$$(\boldsymbol{\lambda} - \boldsymbol{\xi}_{\mathbf{X}}(\boldsymbol{\vartheta})) \mathbf{h}_{t+1} = -\Psi_{\mathbf{X}}^{\mathbb{Q}}(\boldsymbol{\vartheta}; t, t+1) = -\boldsymbol{\xi}_{\mathbf{X}}^{*}(\boldsymbol{\vartheta}) \mathbf{h}_{t+1}^{*}$$

which completes the proof.

D: Risk-neutral reparametrization of affine GARCH dynamics

Risk-neutral reparametrization of a GARCH(1,1) dynamic

The procedure for reparametrizing the GARCH dynamics is systematic. For a GARCH(1,1) dynamic (3.5), we first write its dynamic using the risk-neutral innovations z_t^* and y_t^*

$$h_{x_i,t+1} = w_{x_i} + b_{x_i} h_{x_i,t} + \frac{a_{x_i}}{h_{z,t}} \left(z_t^* - c_{x_i} h_{z,t}\right)^2 \tag{7.5}$$

for $x_i = \{z, y\}$. Because z_t^* has a mean of $-\lambda_z h_{z,t}$, we substitute $z_t^* = z_t - \lambda_z h_{z,t}$ into the above equation. In addition, we recall that from (4.7), we have the following transformation

$$h_{x_i,t+1}^* = \left(\frac{\boldsymbol{\xi}_{x_i}\left(1 + \Lambda_{x_i}\right) - \boldsymbol{\xi}_{x_i}\left(\Lambda_{x_i}\right)}{\boldsymbol{\xi}_{x_i}^*\left(1\right)} \right) h_{x_i,t+1} = \Pi_{x_i} h_{x_i,t+1.}$$

It is important to note that, for the normal innovation, $\Pi_z = 1$ and thus $h_{z,t+1}^* = h_{z,t+1}$. Because we leave y_{t+1} unspecified, we do not know the explicit form of $\Pi_{y.}$ Substituting all this into (7.5), and simplifying we have

$$h_{x_i,t+1}^* = w_{x_i}^* + b_{x_i} h_{x_i,t} + \frac{a_{x_i}^*}{h_{z,t}} \left(z_t - c_{x_i}^* h_{z,t}\right)^2$$

for $x_i = \{z, y\}$ with the following reparametrizations of the GARCH parameters

$$c_{x_i}^* = c_{x_i} + \lambda_z \qquad w_{x_i}^* = w_{x_i} \Pi_{x_i} \qquad a_{x_i}^* = a_{x_i} \Pi_{x_i}.$$

Risk-neutral reparametrization of a component GARCH dynamic

Replacing z_t with z_t^* in (3.7), gives

$$\begin{aligned} h_{x_i,t+1} &= \upsilon_{x_i,t+1} + b_{x_i}\left(h_{x_i,t} - \upsilon_{x_i,t}\right) + \frac{a_{x_i}}{h_{z,t}} \left(\left(\left(z_t^*\right)^2 - h_{z,t}\right) - 2c_{x_i} z_t^* h_{z,t}\right) \\ \upsilon_{x_i,t+1} &= w_{x_i} + \rho_{x_i} \upsilon_{x_i,t} + \frac{\varphi_{x_i}}{h_{z,t}} \left(\left(\left(z_t^*\right)^2 - h_{z,t}\right) - 2d_{x_i} z_t^* h_{z,t}\right) \end{aligned}$$

for $x_i = \{z,\ y\}$. Next, we substitute $z_t^* = z_t - \lambda_z h_{z,t}$, $h_{x_i,t}^* = \Pi_{x_i} h_{x_i,t}$, and $\upsilon_{x_i,t+1}^* = \Pi_{x_i} \upsilon_{x_i,t+1}$ into the above, then simplify. It turns out that the component GARCH model under $\mathbb{Q}$ measure can be written as

$$\begin{aligned} h_{x_i,t+1}^* &= \upsilon_{x_i,t+1}^* + b_{x_i}\left(h_{x_i,t}^* - \upsilon_{x_i,t}^*\right) + e_{x_i}^* h_{z,t} + \frac{a_{x_i}^*}{h_{z,t}} \left(\left(z_t^2 - h_{z,t}\right) - 2c_{x_i}^* z_t h_{z,t}\right) \\ \upsilon_{x_i,t+1}^* &= w_{x_i}^* + \rho_{x_i} \upsilon_{x_i,t}^* + g_{x_i}^* h_{z,t} + \frac{\varphi_{x_i}^*}{h_{z,t}} \left(\left(z_t^2 - h_{z,t}\right) - 2d_{x_i}^* z_t h_{z,t}\right) \end{aligned}$$

with the following reparametrizations of the GARCH parameters

$$w_{x_i}^* = w_{x_i}\Pi_{x_i} \qquad a_{x_i}^* = a_{x_i}^*\Pi_{x_i} \qquad c_{x_i}^* = (c_{x_i} + \lambda_z) \qquad d_{x_i}^* = (d_{x_i} + \lambda_z)$$

$$g_{x_i}^* = \Pi_{x_i}\varphi_{x_i}\lambda_z\left(2d_{x_i} + \lambda_z\right) \qquad \varphi_i^* = \Pi_{x_i}\varphi \qquad e_{x_i}^* = a_{x_i}\lambda_z\Pi_{x_i}\left(2c_{x_i} + \lambda_z\right).$$

Note that the dynamic for $h_{z,t+1}^*$ can be further simplified because $\Pi_z = 1$.

Risk-neutral reparametrization of a GARCH(1,1) dynamic with jumps

Replacing $\{z_t, y_t\}$ with $\{z_t^*, y_t^*\}$ and following the systematic procedure as shown in the two cases above, the $\mathbb{Q}$-measure dynamic of a GARCH(1,1) with jumps in (3.9) can be written as

$$h_{x_i,t+1}^* = w_{x_i}^* + b_{x_i}h_{x_i,t} + d_{x_i}^*y_t^* + \frac{a_{x_i}^*}{h_{z,t}}\left(z_t - c_{x_i}^*h_{z,t}\right)^2$$

for $x_i = \{z, y\}$, with the following reparametrizations of the GARCH parameters

$$w_{x_i}^* = w_{x_i}\Pi_{x_i} \qquad a_{x_i}^* = a_{x_i}^*\Pi_{x_i} \qquad c_{x_i}^* = (c_{x_i} + \lambda_z) \qquad d_{x_i}^* = d_{x_i}\Pi_{x_i}.$$

E: Risk neutralization for the MJ- and NIG-LGARCH models

We give explicit expressions for the returns and GARCH dynamics under the risk-neutral measure of the two LGARCH models studied in this paper. Note that the market prices of risks disappear under the risk-neutral measure $(\lambda_z = \lambda_y = 0)$. This result follows from Lemma 1. Equation (4.4) shows that the EMM coefficient for the normal risk factor is $\Lambda_z = -\lambda_z$. However, the solution for the EMM coefficient of the jump component Λ_y, will be depend on the choice of the jump structure.

For MJ-LGARCH, applying (4.3) with the Merton jump structure gives

$$0 = \lambda_y - \left(e^{\frac{\delta^2}{2}+\theta} - 1\right) - e^{\frac{\Lambda_y^2\delta^2}{2}+\Lambda_y\theta}\left(1 - e^{\left(\frac{1}{2}+\Lambda_y\right)\delta^2+\theta}\right).$$

The above equation cannot be solved analytically. However, the solution for Λ_y can be solved numerically. After applying the measure change, the risk-neutral parameters of a Merton jump y_{t+1}^* is linked to its parameters under the physical measure by

$$h_{y,t+1}^* = h_{y,t+1}e^{\frac{\Lambda_y^2\delta^2}{2}+\Lambda_y\theta}, \text{ and } \theta^* = \theta + \Lambda_y\delta^2.$$

The above result is directly taken from Table 2. For the LGARCH(1), recall that we have $h_{y,t+1} = k$, and for the LGARCH(3), we have $h_{y,t+1} = kh_{z,t+1}$, this implies that $k^* = k\exp\left(\frac{\Lambda_y^2\delta^2}{2} + \Lambda_y\theta\right)$.

Next we solve for Λ_y in the NIG-LGARCH model. Applying (4.3) with the NIG jump structure gives the following relation

$$\Lambda_y = -\frac{1}{2} - \beta - \frac{\sqrt{-\mu_y^2\left(1+\mu_y^2\right)\left(1-4\alpha^2+\mu_y^2\right)}}{2\left(1+\mu_y^2\right)},$$

where

$$\mu_y = \lambda_y - \left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+1)^2}\right).$$

In order to see how the risk-neutral jump parameters are linked to their physical measure counterparts, we again refer to Table 2. This shows us that $\beta^* = \beta + \Lambda_y$, while the rest of the parameters remain unchanged.

Finally, using (7.5), the risk-neutral affine GARCH(1,1) dynamic that drives the heteroskedasticity of $h_{z,t+1}$ in both the MJ- and NIG-LGARCH models becomes

$$h_{z,t+1} = w_z + bh_{z,t} + \frac{a_z}{h_{z,t}}\left(z_t - c_z^* h_{z,t}\right)^2.$$

The only GARCH parameter that changes due to the risk neutralization is $c_z^* = c_z + \lambda_z$.

References

[1] Alizadeh, S., Brandt, M., Diebold, F., 2002. Ranged-based estimation of stochastic volatility models, Journal of Finance 57, 1047-1091.

[2] Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. Journal of Finance 57, 1239-1284.

[3] Ahn, L., Dai, Q., Singleton, K., 2006. Discrete-time dynamic term structure models with generalized market prices of risks. Unpublished working paper, University of North Carolina at Chapel Hill.

[4] Bakshi, G., Cao, C., Chen, Z., 1997. Empirical performance of alternative option pricing models. Journal of Finance 52, 2003-2049.

[5] Bakshi, G., Carr, P., Wu, L., 2008. Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies. Journal of Financial Economics 87, 132-156.

[6] Bakshi, G., Madan, D., 2000. Spanning and derivative-security valuation. Journal of Financial Economics 55, 205-238.

[7] Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. Review of Financial Studies 16, 101-143.

[8] Bakshi, G., Cao, C., 2004. Disentangling the contribution of return-jumps and volatility-jumps: insights from individual equity options. Unpublished working paper, University of Maryland.

[9] Bates, D., 1996. Jumps and stochastic volatility: exchange rate processes implicit in Deutsche mark options. Review of Financial Studies 9, 69-107.

[10] Bates, D., 2000. Post-87 crash fears in S&P 500 futures options. Journal of Econometrics 94, 181-238.

[11] Bates, D., 2006. Maximum likelihood estimation of latent affine processes. Review of Financial Studies 19, 909-965.

[12] Bates, D., 2008. Estimation and filtration of time-changed Lévy processes. Unpublished working paper, University of Iowa.

[13] Barndorff-Nielsen, O.E., 1998. Processes of normal inverse Gaussian type. Finance and Stochastics 2, 41-68.

[14] Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-659.

[15] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327.

[16] Bollerslev, T., Forsberg, L., 2002. Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): the GARCH-NIG model. Journal of Applied Econometrics, 535-548.

[17] Bollerslev, T., Zhou, H., 2002. Estimating stochastic volatility using conditional moments of integrated volatility. Journal of Econometrics 92, 75-99.

[18] Brennan, M., 1979. The pricing of contingent claims in discrete-time models. Journal of Finance 34, 53-68.

[19] Broadie, M., Chernov, M., Johannes, M., 2007. Model specification and risk premia: the evidence from the futures options. Journal of Finance 62. 1453-1490.

[20] Buhlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A., 1998. On Esscher transform in discrete finance models. ASTIN Bullentin.

[21] Carr, P., Geman, H., Madan, D., Yor, M., 2002. The fine structure of asset returns: an empirical investigation. Journal of Business 75, 305-332.

[22] Carr, P., Geman, H., Madan, D., Yor, M., 2003. Stochastic volatility for Lévy processes. Mathematical Finance 13, 345-382.

[23] Carr, P., Wu, L., 2003a. The finite moment log stable process and option pricing. Journal of Finance 58, 753-778.

[24] Carr, P., Wu, L., 2003b. What type of process underlies options? A simple robust test. Journal of Finance 57, 2581-2610.

[25] Carr, P., Wu, L., 2004. Time-changed Lévy processes and option pricing. Journal of Financial Economics 17, 113–141.

[26] Chan, C., 1999. Pricing contingent claims on stocks driven by Lévy processes. Annals of Applied Probability 9, 504-528.

[27] Chen, X., Ghysels, E., 2008. News- good or bad - and its impact on volatility predictions over multiple horizons. Unpublished working paper, University of North Carolina at Chapel Hill.

[28] Chernov, M., Ghysels, E., 2000. A study towards a unified approach to the joint estimation of objective and risk-neutral measures for the purpose of option valuation. Journal of Financial Economics 56, 407-458.

[29] Chernov, M., Gallant, R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225-257.

[30] Christoffersen, P., Heston, S., Jacobs, K., 2006. Option valuation with conditional skewness. Journal of Econometrics 131, 253-284.

[31] Christoffersen, P., Jacobs, K., Mimouni, K., 2007. Models for S&P 500 dynamics: evidence from realized volatility, daily returns, and option prices . Unpublished working paper, McGill University.

[32] Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with short run and long run volatility components. Journal of Financial Economics, forthcoming.

[33] Christoffersen, P., Jacobs, K., Ornthanalai, C., 2008. Exploring time-varying jump intensities: evidence from the S&P 500 returns and options. Unpublished working paper, McGill University.

[34] Cox, J.C., Ingersoll, J.E., Ross, S.R., 1985. A theory of the term structure of interest rates. Econometrica 53, 385-408.

[35] Dai, Q., Singleton, K., 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics 63, 415-441.

[36] Dennis, P., Mayhew S., 2002. Risk-neutral skewness: evidence from stock options. Journal of Financial and Quantitative Analysis 37, 471-493, 2002.

[37] Duan, J.-C., 1995. The GARCH option pricing model. Mathematical Finance 5, 13-32.

[38] Duan, J.-C., Ritchken, P., Sun, Z., 2006. Approximating GARCH-jump models, jump-diffusion processes, and option pricing. Mathematical Finance 16, 21-52.

[39] Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343-1376.

[40] Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica 50, 987-1008.

[41] Engle, R., Lee, G., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R., White, H. (Eds), Cointegration, causality, and forecasting: a festschrift in honor of Clive W.J. Granger. Oxford University Press, New York, 475-497.

[42] Engle, R., Ng, V., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749-1778.

[43] Esscher, F., 1932. On the probability function in the collective theory of risk. Skandinavisk Aktuarietidskrift, 15:175-195.

[44] Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. Journal of Finance 59, 1367-1403.

[45] Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. Journal of Finance 58, 1269-1300.

[46] Fama, E.F., Bliss, R.R., 1987. The information in long-maturity forward rates. American Economic Review 77 (4), 680–692.

[47] Gerber, H., Shiu, E., 1994. Option pricing by Esscher transforms. Transactions of Society of Actuaries, 46.

[48] Gordon, N., Salmond, D., Smith, A.F., 2003. Novel approach to nonlinear/nonlinear-Gaussian Bayesian state estimation. Proceedings of the Institute of Electrical Engineers 140, 107-113.

[49] Johannes, M., Polson, N., Stroud, J., 2008. Learning about jumps and stochastic volatility: filtering stochastic differential equations with jumps. Review of Financial Studies, forthcoming.

[50] Hentschel, L. 1995. All in the family: nesting symmetric and asymmetric GARCH models. Journal of Financial Economics 39, 71-104.

[51] Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6, 327-343.

[52] Heston, S., Nandi, S., 2000. A closed-form GARCH option pricing model. Review of Financial Studies 13, 585-626.

[53] Hull, J., White, A., 1987. The pricing of options on assets with stochastic volatility. Journal of Finance 42, 281-300.

[54] Huang, J.-Z., Wu, L., 2004. Specification analysis of option pricing models based on time-changed Lévy processes. Journal of Finance 59, 1405–1439.

[55] Kou, S.G., 2002. A jump diffusion model for option pricing. Management Science 48, 1086-1101.

[56] Lewis, A., 2000. Option valuation under stochastic volatility. Finance Press, California.

[57] Li, H., Wells, M., Yu, C., 2006. A Bayesian analysis of return dynamics with Lévy jumps. Review of Financial Studies, forthcoming.

[58] Madan, D.B., Seneta, E., 1990. The variance gamma (V.G.) model for share market returns. Journal of Business, 63, 511-524.

[59] Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. Journal of Finance 59, 755–793.

[60] Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125-144.

[61] Mo, H., Wu, L., 2007. International capital asset pricing: evidence from options. Journal of Empirical Finance 14, 465-498.

[62] Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63, 3–50.

[63] Pitt, M., Shephard, N., 1999. Filtering via simulation based on auxiliary particle filters. Journal of American Statistical Association 94, 590-599.

[64] Prause, K., 1999. The generalized hyperbolic model: estimation, financial derivatives, and risk measures. PhD dissertation, University of Freiburg.

[65] Ritchken, P., Trevor, R., 1999. Pricing options under generalized GARCH and stochastic volatility processes. Journal of Finance 54, 377-402.

[66] Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. Bell Journal of Economics 7, 407-425.

[67] Siu, T., Tong, H., Yang, H., 2004. On pricing derivatives under GARCH models: a dynamic Gerber-Shiu approach. North American Actuarial Journal 8 (2004), 17-31.

[68] Schoutens, W., 2000. Stochastic processes and orthogonal polynomials. Lecture Notes in Statistics 146: Springer.

[69] Schoutens, W., 2001. The Meixner process in finance. EURANDOM report 2001-2002.

[70] Schoutens, W., 2003. Lévy processes in finance. Wiley Series in Probability and Statistics: Wiley.

[71] Stentoft, L., 2007. American option pricing using GARCH models and the normal inverse Gaussian distribution. Unpublished working paper, HEC Montréal.

[72] Schwert, G., W., 1990. Stock volatility and the crash of 87. Review of Financial Studies 3, 77-102.

[73] Tweedie, M.C.K., 1984. An index which distinquishes between some important exponential families. In Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International Conference, 579-604.

Table 1: The Lévy measures, cumulant exponents, and references for selected pure jump Lévy processes

Pure Lévy jump components	**Lévy measures** $\nu(dx)$	**Cumulant Exponent** $\Psi(\phi; h_{t+1}) = h_{t+1}\,\xi_y(\phi)$	**Reference**
Finite-activity jump process			
Poisson Jump (PJ)- fixed jump size	$h_{t+1}1_{(x=\theta)}$	$h_{t+1}\left(e^{\theta\phi}-1\right)$	
Merton Jump (MJ)	$\frac{h_{t+1}}{\sqrt{2\pi\delta^2}}e^{-\frac{(x-\theta)^2}{2\delta^2}}$	$h_{t+1}\left(e^{\theta\phi+\frac{1}{2}\phi^2\delta^2}-1\right)$	Merton (1976)
Double Exponential Jump (DEP)	$\begin{cases} h_{t+1}p\eta_1 e^{-\eta_1 x} & x \geq 0 \\ h_{t+1}\eta_2(1-p)e^{\eta_2 x} & x < 0 \end{cases}$	$h_{t+1}\left(\frac{\eta_1 p}{(\eta_1-\phi)}+\frac{\eta_2(1-p)}{(\eta_2+\phi)}-1\right)$	Kou (2002) Mo and Wu (2007)
Infinite-activity jump process			
Inverse Gaussian (IG)	$\frac{h_{t+1}x^{-3/2}}{\sqrt{2\pi}}e^{-\frac{1}{2}b^2 x}1_{(x\geq 0)}$	$h_{t+1}\left(b-\sqrt{2\phi+b^2}\right)$	Christoffersen, Heston, and Jacobs (2006)
Normal Inverse Gaussian (NIG)	$\frac{h_{t+1}e^{\beta x}K_1(\alpha\|x\|)}{\pi\|x\|}$	$h_{t+1}\left(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+\phi)^2}\right)$	Barndorff-Nielsen (1998)
Gamma	$\frac{h_{t+1}}{x}e^{-bx}1_{(x>0)}$	-$h_{t+1}\log\left(1-\frac{\phi}{b}\right)$	
Variance Gamma (VG)	$\frac{\mu_\pm^2}{\upsilon_\pm}\frac{\exp\left(-\frac{\mu_\pm}{\upsilon_\pm}\|x\|\right)}{\|x\|}$; $\mu_\pm=\frac{h_{t+1}\left(\sqrt{\theta^2+2b}\pm\theta\right)}{2b}$ $\upsilon_\pm=\frac{\mu_\pm^2}{h_{t+1}}$	$-h_{t+1}\log\left(1-\frac{\theta\phi+\frac{1}{2}\phi^2}{b}\right)$	Madan and Seneta (1990) Madan and Milne (1998)
Tempered stable (TS)	$\frac{\gamma}{\Gamma(1-\gamma)}h_{t+1}2^\gamma x^{-\gamma-1}\exp\left(-\frac{x}{2}b^{1/\gamma}\right)1_{\{x\geq 0\}}$	$h_{t+1}\left(b-\left(b^{1/\gamma}-2\phi\right)^\gamma\right)$	Tweedie(1984)
Log-stable (LS)	$c_-\|x\|^{-\alpha-1}dx$; where $c_-=\frac{-h_{t+1}\sec\left(\frac{\pi\alpha}{2}\right)}{\Gamma(-\alpha)}$	$-h_{t+1}\phi^\alpha\sec\left(\frac{\pi\alpha}{2}\right)$	Carr and Wu (2003a)
CGMY	$\begin{cases} C\exp(-Mx)x^{-1-Y} & x > 0 \\ C\exp(Gx)(-x)^{-1-Y} & x < 0 \end{cases}$; $C=h_{t+1}$	$h_{t+1}\Gamma(-Y)\left((M-\phi)^Y-M^Y+(G+\phi)^Y-G^Y\right)$	Carr, Geman, Madan, and Yor (2002)
Meixner	$h_{t+1}\frac{\exp(\beta x/\alpha)}{x\sinh(\pi x/\alpha)}$	$2h_{t+1}\log\left[\cos\left(\frac{\beta}{2}\right)\text{sech}\left(\frac{i(\alpha\phi+\beta)}{2}\right)\right]$	Schoutens (2000, 2001)

Table 2: Summary of measure change from $\mathbb{P}$ to $\mathbb{Q}$ for selected pure jump Lévy processes

$\mathbb{P}$ measure distribution	$\mathbb{Q}$ measure distribution	Parameters' transformation
Finite-activity jump process		
PJ: $\Theta^{\mathbb{P}}=\{h_{t+1},\theta\}$	PJ: $\Theta^{\mathbb{Q}}=\{h_{t+1},b\}$	$h^*_{t+1}=h_{t+1}e^{\theta\Lambda}$
MJ: $\Theta^{\mathbb{P}}=\left\{h_{t+1},\theta,\delta^2\right\}$	MJ: $\Theta^{\mathbb{Q}}=\left\{h^*_{t+1},\theta^*,\delta^2\right\}$	$h^*_{t+1}=h_{t+1}e^{\left(\theta\Lambda+\frac{\Lambda^2\delta^2}{2}\right)}$; $\theta^*=\theta+\Lambda\delta^2$
DEP: $\Theta^{\mathbb{P}}=\{h_{t+1},p,\eta_1,\eta_2\}$	unrecognizable	
DEP-MW: $\Theta^{\mathbb{P}}=\{h_{t+1},\eta_1,\eta_2\}$	DEP-MW: $\Theta^{\mathbb{Q}}=\{h_{t+1},\eta^*_1,\eta^*_2\}$	$\eta^*_1=\eta_1-\Lambda$; $\eta^*_2=\eta_2-\Lambda$
Infinite-activity jump process		
IG: $\Theta^{\mathbb{P}}=\{h_{t+1},b\}$	IG: $\Theta^{\mathbb{Q}}=\{h_{t+1},b^*\}$	$b^*=\sqrt{b+2\Lambda}$
NIG: $\Theta^{\mathbb{P}}=\{h_{t+1},\alpha,\beta\}$	NIG: $\Theta^{\mathbb{Q}}=\{h_{t+1},\alpha,\beta^*\}$	$\beta^*=\beta+\Lambda$
Gamma: $\Theta^{\mathbb{P}}=\{h_{t+1},b\}$	Gamma: $\Theta^{\mathbb{Q}}=\{h_{t+1},b^*\}$	$b^*=b-\Lambda$
VG: $\Theta^{\mathbb{P}}=\{h_{t+1},\theta,b\}$	VG: $\Theta^{\mathbb{Q}}=\{h_{t+1},\theta^*,b^*\}$	$\theta^*=\theta+\Lambda$; $b^*=b-\frac{1}{2}\Lambda\left(2\theta+\Lambda\right)$
TS: $\Theta^{\mathbb{P}}=\{h_{t+1},,\gamma,b\}$	TS: $\Theta^{\mathbb{Q}}=\{h_{t+1},\gamma,b^*\}$	$b^*=\left(b^{1/\gamma}-2\Lambda\right)^{\gamma}$
LS: $\Theta^{\mathbb{P}}=\{h_{t+1},\alpha,\sigma\}$	unrecognizable	
CGMY: $\Theta^{\mathbb{P}}=\{h_{t+1},G,M,Y\}$	CGMY: $\Theta^{\mathbb{Q}}=\{h_{t+1},G^*,M^*,Y\}$	$G^*=G-\Lambda$; $M^*=M-\Lambda$
Meixner: $\Theta^{\mathbb{P}}=\{h_{t+1},\alpha,\beta\}$	Meixner: $\Theta^{\mathbb{Q}}=\{h_{t+1},\alpha,\beta^*\}$	$\beta^*=\beta+\alpha\Lambda$

Notes to Table: The table shows the resulting risk-neutral measure distributions from the change of measure through the Esscher transform for selected pure jump Lévy innovations that we present in Table 1. DEP-MW refers to a restricted version of the double-exponential jump (DEP) with $p=\eta_1$ and $1-p=\eta_2$, which is studied in Mo and Wu (2007). The parameter Λ refers to the EMM coefficient for each of the Lévy innovation, which can be solved from (4.2). Most of the Lévy innovations retain their own distribution after applying the change of measure, but with differences in the parameters; we apply star superscripts to these risk-neutral measure parameters that differ from their physical measure counterparts. Note that we cannot associate the risk-neutral transformed cumulant exponent of the DEP and LS processes with any of the well-known distributions. We therefore denote their $\mathbb{Q}$ measure distributions as "unrecognizable".

Table 3: S&P 500 index call option data (1996-2005)

	DTM<20	20<DTM<80	80<DTM<180	DTM>180	All
Panel A. Number of call option contracts					
S/K<0.975	123	1,841	2,078	2,293	6,416
0.975<S/K<1.00	554	2,557	1,076	645	4,851
1.00<S/K<1.025	867	2,282	717	366	4,236
1.025<S/K<1.05	571	1,337	413	191	2,516
1.05<S/K<1.075	257	839	263	139	1,501
1.075<S/K	298	1,190	466	237	2,198
All	2,670	10,046	5,013	3,871	21,718
Panel B. Average call price					
S/K<0.975	5.39	13.96	26.29	43.56	28.95
0.975<S/K<1.00	11.82	24.12	44.31	77.13	34.58
1.00<S/K<1.025	23.86	36.25	60.76	92.19	42.77
1.025<S/K<1.05	43.30	55.37	79.43	110.79	60.90
1.05<S/K<1.075	66.65	76.40	99.07	127.03	83.53
1.075<S/K	111.08	120.90	135.19	169.19	127.98
All	38.52	45.00	53.41	67.76	50.40
Panel C. Average implied volatility from call options					
S/K<0.975	0.2075	0.1876	0.1875	0.1831	0.1863
0.975<S/K<1.00	0.1768	0.1768	0.1831	0.1865	0.1796
1.00<S/K<1.025	0.1785	0.1813	0.1948	0.1955	0.1842
1.025<S/K<1.05	0.2034	0.1983	0.2040	0.2041	0.2009
1.05<S/K<1.075	0.2554	0.2187	0.2122	0.2056	0.2227
1.075<S/K	0.3561	0.2691	0.2379	0.2266	0.2695
All	0.2120	0.1971	0.1950	0.1893	0.1970

Notes to Table: We use European call options on the S&P 500 index. The data are obtained from OptionMetrics. The prices are taken from nonzero trading volume quotes on each Wednesday during the January 1, 1996 to December 31, 2005 period. The moneyness and maturity filters used by Bakshi, Cao and Chen (1997) are applied to the data. The implied volatilities are calculated using the Black-Scholes formula.

Table 4: MLE Estimates of Levy GARCH models on S&P 500 returns: 1985-2005

									Avg Annul Var	
No jump									**Avg Annul Var**	
		μ	b	a	c				Normal	
HN-LGARCH		6.07E-01	9.09E-01	4.52E-06	1.16E+02				0.0289	
		(7.56E-01)	(7.55E-03)	(2.74E-07)	(1.08E+01)					
LogLkhood :	**17354**		Option IVRMSE (%) :		**5.44**					
Finite-activity Merton jump									**Avg Annul Var**	
	μ_z	μ_y	b	a	c	θ	δ	k	Normal	Jump
MJ-LGARCH (1)	1.04E+00	2.31E-04	9.42E-01	2.90E-06	1.25E+02	-6.78E-03	2.57E-02	4.62E-03	0.0280	0.0008
	(1.04E-02)	(4.59E-06)	(4.78E-04)	(1.24E-07)	(3.00E+00)	(8.29E-05)	(7.15E-04)	(1.63E-06)		
LogLkhood :	**17503**		Option IVRMSE ratio :		**0.98**					
		μ	b	a	c	θ	δ	k		
MJ-LGARCH (3)		4.15E+00	9.49E-01	2.58E-06	1.22E+02	-2.83E-03	1.72E-02	5.98E+02	0.0244	0.0044
		(3.00E-03)	(3.08E-05)	(1.37E-08)	(3.99E-01)	(1.73E-06)	(5.02E-05)	(2.77E+00)		
LogLkhood :	**17519**		Option IVRMSE ratio :		**0.94**					
Infinite-activity Normal Inverse Gaussian jump									**Avg Annul Var**	
	μ_z	μ_y	b	a	c	α	β	k	Normal	Jump
NIG-LGARCH (1)	-5.01E-01	1.11E+00	9.39E-01	2.02E-06	1.55E+02	1.54E+01	-9.11E+00	3.14E-04	0.0190	0.0098
	(1.41E-03)	(2.97E-03)	(6.69E-05)	(2.73E-08)	(1.67E+00)	(7.51E-04)	(1.03E-02)	(5.55E-07)		
LogLkhood :	**17547**		Option IVRMSE ratio :		**0.88**					
		μ	b	a	c	α	β	k		
NIG-LGARCH (3)		1.40E+00	9.41E-01	2.00E-06	1.55E+02	1.53E+01	-9.04E+00	2.71E+00	0.0216	0.0073
		(2.61E-03)	(1.08E-04)	(3.65E-08)	(1.20E+00)	(4.33E-03)	(3.70E-02)	(5.73E-03)		
LogLkhood :	**17545**		Option IVRMSE ratio :		**0.85**					

Notes to Table: We apply MLE to the daily return series of the S&P 500 index from January 1985 to December 2005. HN-LGARCH refers to the Heston-Nandi GARCH model which has no jump component. We estimate two types of jump innovations, finite-activity Merton Jump (MJ) and infinite-activity Normal Inverse Gaussian (NIG). For each type of Lévy jump, we consider the LGARCH(1) model where the jump distribution is constant (homoskedastic), and LGARCH(3) where the jump distribution is heteroskedastic and state-dependent. Reported under these estimates are standard errors computed using the outer product of the gradients. We use these MLE estimates to price options and compute their IVRMSEs. We report % IVRMSE for the HN-LGARCH case, and for the jump models we report IVRMSE ratios relative to the HN-LGARCH. The two right columns refer to the annualized long-run variance implied from the MLE estimation. For jump models, we report the magnitude of the return variances according to their two sources: the normal and the jump components. All models are estimated using variance targeting.

Table 5. IVRMSE (% and ratios) of LGARCH models by moneyness, maturity, and VIX level

Panel A: Sorting by moneyness

	HN-LGARCH	MJ-LGARCH(1)	MJ-LGARCH(3)	NIG-LGARCH(1)	NIG-LGARCH(3)
Moneyness	IVRMSE(%)	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio
S/K<0.975	4.479	0.905	0.843	0.894	0.858
0.975<S/K<1.00	4.187	0.949	0.900	0.937	0.896
1.00<S/K<1.025	4.523	0.965	0.921	0.938	0.895
1.025<S/K<1.05	5.508	0.996	0.956	0.907	0.888
1.05<S/K<1.075	6.748	1.013	0.980	0.857	0.846
1.075<S/K	9.510	1.015	0.997	0.797	0.783
All	5.444	0.976	0.938	0.875	0.848

Panel B: Sorting by maturity

	HN-LGARCH	MJ-LGARCH(1)	MJ-LGARCH(3)	NIG-LGARCH(1)	NIG-LGARCH(3)
Maturity	IVRMSE(%)	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio
DTM<20	7.261	1.010	0.984	0.831	0.826
20<DTM<80	5.331	0.993	0.953	0.880	0.853
80<DTM<180	4.937	0.932	0.888	0.887	0.849
DTM>180	4.886	0.924	0.887	0.910	0.865
All	5.444	0.976	0.938	0.875	0.848

Panel C: Sorting by VIX level

	HN-LGARCH	MJ-LGARCH(1)	MJ-LGARCH(3)	NIG-LGARCH(1)	NIG-LGARCH(3)
Maturity	IVRMSE(%)	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio	IVRMSE Ratio
VIX<16	3.329	0.909	0.923	0.852	0.796
16<VIX<22	4.112	1.077	1.034	0.821	0.952
22<VIX<26	5.635	0.993	0.946	0.859	0.869
26<VIX<30	7.126	0.960	0.915	0.888	0.825
30<VIX	9.676	0.916	0.880	0.917	0.783
All	5.444	0.976	0.938	0.875	0.848

Notes to Table: We use the MLE estimates from Table 3 to compute the implied volatility root mean squared error (IVRMSE) for various moneyness, maturity, and VIX level bins. HN-LGARCH refers to the Heston-Nandi GARCH model which has no jump component. MJ-LGARCH refers to the jump models that rely on the finite-activity Merton jump process, while NIG-LGARCH refers to the jump models that rely on the infinite-activity Normal Inverse Gaussian jump process. We denote the models that have homoskedastic jump specification with an LGARCH(1) extension. Similarly, we denote the models that have heteroskedastic and state-dependent jump specification with an LGARCH(3) extension. The IVRMSE is reported in percentage levels for the HN-LGARCH model. For the MJ-LGARCH and NIG-LGARCH models, we report the IVRMSE ratios relative to the HN-LGARCH model.

Table 6: Joint option and returns MLE estimates of selected LGARCH models on the S&P 500 index

	HN-LGARCH		**MJ-LGARCH(3)**		**NIG-LGARCH(3)**	
Parameters	MLE estimates	Std error	MLE estimates	Std error	MLE estimates	Std error
λz	7.36E+00	(1.83E-02)	2.61E+00	(7.23E-04)	5.73E-01	(5.25E-04)
λy			2.092E-03	(3.09E-07)	7.72E-01	(6.61E-05)
b	9.33E-01	(4.80E-05)	9.48E-01	(1.31E-05)	9.40E-01	(1.35E-05)
a	3.68E-06	(5.33E-09)	2.77E-06	(5.35E-09)	2.02E-06	(3.62E-09)
c	1.21E+02	(3.60E-02)	1.23E+02	(9.15E-02)	1.55E+02	(1.46E-01)
θ or α			-2.83E-03	(5.87E-07)	1.60E+01	(5.16E-04)
δ or β			2.85E-02	(2.39E-05)	-9.60E+00	(3.45E-03)
k			8.00E+02	(1.08E+00)	2.76E+00	(9.57E-04)
GARCH Persistence	0.9869		0.9902		0.9887	
Avg Total Annual EP (%)	**22.90**		**8.04**		**6.30**	
Avg Annual Jump RP (%)			**3.14**		**4.96**	
Avg Annual Normal RP (%)			**4.90**		**1.34**	
IVRMSE (%)	**3.97**		**3.85**		**3.40**	
IV residual BIAS (%)	0.49		0.36		-0.05	
IV residual skewness	1.30		1.14		0.36	
IV residual kurtosis	7.09		6.56		3.63	
Log Likelihood	**112207**		**112880**		**115599**	

Notes to Table: We apply MLE to jointly estimate options and returns data. The data set consists of daily returns of the S&P 500 index from Jan 1995 to Dec 2005, and Wednesday call options from Jan 1996 to Dec 2005. We report only the results for jump models that have heteroskedastic and state-dependent jump specification. These models are denoted with an LGARCH(3) extension. Reported beside each estimate is the standard error computed using the outer product of the gradients. Reported are the physical estimates. Risk-neutral estimates can be computed from these values. “The Avg Total Annual EP” refers to the long-run equity premium implied by these estimates. “The Avg Annual Jump RP” and "The Avg Normal RP" refer to the long-run jump and normal risk premia implied by these estimates. We also report the option IVRMSE computed based on these optimal values. The "IV residual skewness" is the skewness computed from the option pricing implied volatility residual. The "IV residual kurtosis" is defined in a similar manner.