Journal of Financial Econometrics Advance Access published July 18, 2010

Journal of Financial Econometrics, 2010, 1–31

Inference in Infinite Superpositions of Non-Gaussian Ornstein–Uhlenbeck Processes Using Bayesian Nonparametic Methods

J. E. GRIFFIN
University of Kent

ABSTRACT
This paper describes a Bayesian nonparametric approach to volatility estimation. Volatility is assumed to follow a superposition of an infinite number of Ornstein–Uhlenbeck processes driven by a compound Poisson process with a parametric or nonparametric jump size distribution. This model allows a wide range of possible dependencies and marginal distributions for volatility. The properties of the model and prior specification are discussed, and a Markov chain Monte Carlo algorithm for inference is described. The model is fitted to daily returns of four indices: the Standard and Poors 500, the NASDAQ 100, the FTSE 100, and the Nikkei 225. (*JEL*: C11, C14, C22)

KEYWORDS: Dirichlet process, Stochastic volatility, Stock indices, Markov chain Monte Carlo, Pólya tree

This paper is concerned with the modeling of financial data such as stock prices or stock indices with stochastic volatility models. Initially, it is assumed that the price process $y(t)$ is defined by the stochastic differential equation (SDE)

$$\mathrm{d}y(t) = (\mu + \beta\sigma^2(t))\,\mathrm{d}t + \sigma^2(t)\,\mathrm{d}B(t) \tag{1}$$

where $B(t)$ is a Brownian motion, $\sigma^2(t)$ is a stochastic process representing the instantaneous volatility, μ is the riskless rate of returns, and β is a risk premium. Many specifications for the volatility process in continuous time have been discussed in the literature. This paper will concentrate on the class of non-Gaussian

Address correspondence to Jim Griffin, School of Mathematics, Statistics, and Actuarial Science, University of Kent, Canterbury, CT2 7NF, or e-mail: J.E.Griffin-28@kent.ac.uk.

doi: 10.1093/jjfinec/nbq027

Ornstein–Uhlenbeck processes (Barndorff-Nielsen and Shephard 2001) which will be referred to simply as OU processes and are defined by the SDE

$$d\sigma^2(t) = -\lambda\,\sigma^2(t) + dz(\lambda t)$$

where z is a non-Gaussian Lévy process and $\lambda > 0$ is a decay parameter. This defines a volatility process which is mean reverting, evolves by jumps which are discounted exponentially at rate λ, and has an exponential autocorrelation function $\rho(t) = \exp\{-\lambda t\}$. If the underlying process $z(\cdot)$ is interpreted as the arrival of information to the market, then the model assumes that the effect of a piece of information on volatility is discounted exponentially over time. An attractive aspect of these processes is the wide range of possible marginal distributions for $\sigma^2(t)$ which can be defined by the choice of the Lévy measure for $z(\cdot)$.

In practice, observations of returns $r_i = y(i\Delta) - y((i-1)\Delta)$ are made for some time period $\Delta > 0$ and $\sigma^2(t)$ is, usually, chosen to follow some parametric family. Inference must be made about λ and the parameters of the distribution of $\sigma^2(t)$. Barndorff-Nielsen and Shephard (2001, 2002) discuss fitting these models using realized volatility. More recent work in this direction is reviewed by Woerner (2007). Inference using characteristic functions is discussed by Valdivieso, Shoutens, and Tuerlinckx (2009) and Taufer, Leonenko, and Bee (2009). Bayesian inference using Markov chain Monte Carlo (MCMC) methods have been developed by Roberts, Papaspiliopoulos, and Dellaportas (2004), Griffin and Steel (2006), Gander and Stephens (2007a,b), and Frühwirth-Schnatter and Sögner (2009). Mostly, in the Bayesian literature, a Gamma distribution is assumed for $\sigma^2(t)$ which implies that z is a compound Poisson process. This allows exact inference to be made using MCMC methods. For general marginal distributions of $\sigma^2(t)$, $z(\cdot)$ will often be an infinite activity Lévy process for a chosen marginal distribution and $z(\cdot)$ can only be simulated using a truncation method (Cont and Tankov 2003). MCMC for general marginal distributions are discussed by Gander and Stephens (2007b). There has been no work on nonparametric inference, where we do not assume a parametric form for the marginal distribution of $\sigma^2(t)$, for these models with returns data. However, work in this direction is described by Jongbloed, Van Der Meulen, and Van Der Vaart (2005) who develop a nonparametric method of estimating λ and the Lévy density of z using a cumulant-based method applied to data drawn directly from the OU process. One aim of this paper is to develop Bayesian nonparametric approaches to the estimation of the distribution of $\sigma^2(t)$ when only the returns are observed.

Many authors have noted that the form of autocorrelation defined by the OU process is unsuitable for financial data such as stock prices or stock indices. The autocorrelation function of squared daily returns often decays quickly over a few days but decays very slowly at longer lags. This suggests either long memory in volatility or long-range dependence similar to long memory which could not be captured by a function that decays exponentially. Gander and Stephens (2007a) describe alterations to the form of the volatility equation that include jumps but lead to more general dependence, including long memory. Alternatively, more general dependence can be modeled using a superposition of OU processes which expresses the volatility process as

$$\sigma^2(t) = \sum_{i=1}^{K} \sigma_i^2(t) \tag{2}$$

where $\sigma_1^2(t), \sigma_2^2(t), \ldots, \sigma_K^2(t)$ are mutually independent and $\sigma_i^2(t)$ is an OU process with parameter λ_i and Lévy process z_i. The superposition allows fairly flexible dependence structures for the volatility, even for relatively small values of K. For example, Griffin and Steel (2006) show that a superposition of two processes fits fairly well for the returns of the Standard and Poors 500 index and describe a method for approximating posterior probabilities of K (using a method of Newton and Raftery 1994). The superposition can be viewed as a multifactor model (LeBaron 2001, Alizadeh, Brandt, and Diebold 2002, Chernov et al. 2003, Molina, Han, and Fouque 2008) which represents the volatility process as a sum of a small number of component processes with simple dynamics. Each component can be interpreted according to its strength of dependence over time. For example, there may be a slowly varying component which represents the effect of long-run macroinformation and a second more-fast-moving process which represents the effect of short-run information. Under certain conditions, the sum in Equation (2) can be extended to an integral. The probabilistic background is developed by Barndorff-Nielsen (2001) and Barndorff-Nielsen and Leonenko (2005), and an MCMC method for Bayesian inference is discussed by Griffin and Steel (2010). Unlike the superposition in Equation (2), these continuous superpositions can be constructed so that $\sigma^2(t)$ has long memory.

In general, it is difficult to find a suitable value of K in Equation (2). A similar problem arises in mixture modeling where a density is represented as a weighted sum of component distributions (such as normal distributions). It is often difficult to choose the number of component distributions. Recently, there has been interest in using Bayesian nonparametric methods (such as Dirichlet process mixtures) which avoid choosing a specific number of component distributions by assuming an infinite number of components of which only a finite number have nonneglible weight. These nonparametric models behave in a similar way to mixture models with a finite number of components but inference (using MCMC methods) is often simpler. A variation on this nonparametric approach will be used in this paper to avoid the problem of choosing K in Equation (2). A review of Bayesian nonparametric methods is given by Müller and Quintana (2004).

The model for the price process in Equation (1) allows for stochastic volatility but does not include other important features such as a leverage effect (which is the negative correlation of the returns process and the volatility process) or jumps in returns. These can be accommodated using the following SDE

$$\mathrm{d}y(t) = \left\{\mu + \beta\sigma^2(t) + \rho(z(t) - \bar{z}(t)) + J(t)\right\} \mathrm{d}t + \sigma^2(t)\,\mathrm{d}B(t) \tag{3}$$

where $\bar{z}_t = \mathrm{E}[z(t)]$ and $J(t)$ is a compound Poisson process with intensity λ_J and jumps $\mathrm{N}(0, \sigma_J^2)$. Here, ρ is a measure of the leverage effect (as described by Barndorff-Nielsen and Shephard 2001). The parameter ρ measures the effect of a jump in volatility on the mean of the returns (negative ρ would imply negative correlation between return and volatility processes). This is different to standard modeling of leverage which negatively correlated the return process with the volatility process (rather than the jumps of the volatility process as here). If we interpret the jumps of the volatility process as arrivals of information and ρ is negative, then this model implies that the arrival of bad news leads to higher volatility and, on average, a negative return. The $J(t)$ process represents jumps in returns as introduced by Eraker, Johannes, and Polson (2003).

The contribution of this paper is the development of Bayesian methods for a general OU-type model where the sum in Equation (2) is infinite and the marginal distribution of $\sigma^2(t)$ is either known or unknown. This defines a flexible model with a wide range of possible dependencies and marginal distributions of returns. This is unlike previous Bayesian nonparametric work in volatility estimation that has concentrated on the distribution of returns conditional on a parametric model for volatility (Jensen and Maheu 2010, Kalli, Walker, and Damien 2009).

The paper is organized as follows: Section 1 reviews properties of OU processes and superpositions of OU processes, Section 2 discusses prior specification for a general OU-type model, Section 3 describes an MCMC algorithm to make inference in these models, Section 4 describes applications of the method to four stock indices (Standard & Poors 500, NASDAQ 100, FTSE 100, and Nikkei 225), and Section 5 is a discussion.

1 OU PROCESSES

The use of OU processes to model volatility was first discussed by Barndorff-Nielsen and Shephard (2001). It is assumed that the instantaneous volatility $\sigma^2(t)$ is defined by the SDE

$$\mathrm{d}\sigma^2(t) = -\lambda\sigma^2(t) + \mathrm{d}z(\lambda t)$$

where z is a non-Gaussian (pure jump) Lévy process, which is called the background driving Lévy process (BDLP). The solution of the SDE is

$$\sigma^2(t) = \exp\{-\lambda t\}\sigma^2(0) + \int_0^t \exp\{-\lambda(t-s)\}\, \mathrm{d}z(\lambda s). \tag{4}$$

It follows that $\sigma^2(t)$ is strictly stationary and evolves through jumps which decay exponentially. The range of possible marginal distributions is large since an OU process of the form (4) can be constructed for any self-decomposable distribution. If the BDLP is a compound Poisson process with jump intensity ν and jump distribution F_J, the OU process can be represented in terms of a marked point process,

$$\sigma^2(t) = \sum_{i=1}^{\infty} I(\tau_i < t) \exp\{-\lambda(t-\tau_i)\} J_i \tag{5}$$

where $\tau_1, \tau_2, \tau_3, \ldots$ are the points of a Poisson process with intensity $\lambda\nu$ and $J_1, J_2, J_3, \ldots$ are marks for which $J_i \overset{i.i.d.}{\sim} F_J$. For example, an OU process with a Gamma marginal distribution with shape parameter ν and mean ν/γ is given by choosing F_J to be an exponential distribution with mean $1/\gamma$.

The marginal distribution of $\sigma^2(t)$ can be related to the BDLP $z(t)$ in the following ways. Let z have Lévy density w and u be the Lévy density of $\sigma^2(t)$ then

$$w(x) = -u(x) - xu'(x).$$

This is a useful tool for modeling since it allows the BDLP to be found for a chosen marginal distribution of $\sigma^2(t)$. Similarly, u can be derived from w using the following result due to Barndorff-Nielsen and Shephard (2001).

Lemma 1. (BNS) *Let z be a Lévy process with positive increments and cumulant function*

$$\log(E[\exp\{-\theta z(1)\}]) = -\int_{0^+}^{\infty} (1 - \exp\{-\theta x\}) W(\mathrm{d}x)$$

and assume that

$$\int_1^{\infty} \log(x)\, W(\mathrm{d}x) < \infty.$$

Suppose moreover, for simplicity, that the Lévy measure W has a differentiable density w and define the function u on $\mathbb{R}^+$ by

$$u(x) = \int_1^\infty w(\tau x)\,\mathrm{d}\tau.$$

Then u is the Lévy density of a random variable x of the form

$$x = \int_0^\infty \exp\{-s\}\,\mathrm{d}z(s)$$

and the specification

$$x(t) = \int_{-\infty}^t \exp\{-\lambda(t-s)\}\,\mathrm{d}z(s)$$

determines a stationary process $\{x(t)\}_{t\in\mathbb{R}}$ with z as its BDLP.

The cumulants of the two processes are also linked. Let the cumulant-generating function of $\sigma^2(t)$ be $\kappa'(\theta) = \log(E[\exp\{-\theta\sigma^2(t)\}])$ and the cumulant-generating function of $z(1)$ be $\kappa(\theta) = \log(E[\exp\{-\theta z(1)\}])$. Barndorff-Nielsen (2001) shows that

$$\kappa'(\theta) = \int_0^\infty \kappa(\theta\exp\{-s\})\,\mathrm{d}s.$$

and

$$\kappa(\theta) = \theta\frac{\kappa'(\theta)}{\mathrm{d}\theta}.$$

It follows that the m-th cumulant of z, κ_m, is related to the m-th cumulant of $\sigma^2(t)$, κ'_m, be the equation $\kappa_m = m\kappa'_m$.

The autocorrelation function of $\sigma^2(t)$ is given by

$$\mathrm{Corr}(\sigma^2(t), \sigma^2(t+s)) = \exp\{-\lambda s\}$$

which does not depend on the Lévy density of z (and so the marginal distribution of $\sigma^2(t)$) but does depend on λ. Its form is not suitable for asset return or stock indices (Barndorff-Nielsen and Shephard 2001) since the autocorrelation tends to have a rapid initial decay followed by a slower decay at longer lags, which has been confirmed by many subsequent applications (Gander and Stephens 2007a,b, Frühwirth-Schnatter and Sögner 2009). A more flexible dependence structure can be created using superpositions of OU processes, which were first studied by Barndorff-Nielsen (2001). In general, we can write

$$\sigma^2(t) = \sum_{i=1}^K \sigma_i^2(t) \tag{6}$$

where $\sigma_1^2(t), \sigma_2^2(t), \ldots \sigma_K^2(t)$ are mutually independent and $\sigma_i^2(t)$ is an OU process with decay parameter λ_i and BDLP $z_i(t)$. If $z_i(t)$ has the Lévy density $\nu_i \eta(\cdot)$ for all i, then it follows that the autocorrelation function has the form

$$\rho(s) = \text{Corr}(\sigma^2(t), \sigma^2(t+s)) = \sum_{i=1}^{K} \frac{\nu_i}{\sum_{i=1}^{K} \nu_i} \exp\{-\lambda_i s\}.$$

This is a flexible form and can represent a wide range of autocorrelation functions, even for small values of K.

2 A BAYESIAN NONPARAMETRIC MODEL

In this section, a Bayesian nonparametric model for volatility is constructed. This allows dependence to be modeled nonparametrically by specifying the superposition of an infinite number of OU processes and the marginal distribution to be modeled through the BDLP of each process. Attention will be restricted to OU processes driven by a compound Poisson process. All algorithms for Bayesian inference in this class of models use truncation methods to approximate the actual BDLP by a compound Poisson process. Therefore, it seems sensible to construct models with compound Poisson processes as the BDLP. It is well known that any Lévy process can be constructed as the limit of a compound Poisson process. This leads to the following definition:

Definition 1. Let $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3, \ldots)$ be an infinite sequence of nonnegative numbers for which $0 < \sum_{i=1}^{\infty} \nu_i < \infty$, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \ldots)$ be an infinite sequence of positive numbers and F_J be a distribution with support on $\mathbb{R}^+$. Suppose that

$$\sigma^2(t) = \sum_{i=1}^{\infty} \sigma_i^2(t)$$

where $\sigma_i^2(t)$ is an OU process whose BDLP has Lévy density $\nu_i \lambda_i F_J$. Then $\sigma^2(t)$ follows an Inf-Sup OU with masses $\boldsymbol{\nu}$, decays $\boldsymbol{\lambda}$, and shot distribution F_J. This is written Inf-Sup OU$(\boldsymbol{\nu}, \boldsymbol{\lambda}, F_J)$.

The existence of $\sigma^2(t)$ is guaranteed by Theorem 3.1 of Barndorff-Nielsen (2001). The definition implies that the Lévy density of each BDLP has finite integral, but this result extends to the more general case where the Lévy density of each BDLP is $\nu_i \lambda_i \eta(x)$ and $\eta(x)$ is the Lévy density of a BDLP for some OU process (which possibly has an infinite integral). The parameter ν_i controls the average contribution of the i-th component, and λ_i controls the dependence of the i-th component in the superposition. It is convenient to think of $(\nu_1, \lambda_1), (\nu_2, \lambda_2), \ldots$ in terms of the random measure $G_\lambda = \sum_{i=1}^{\infty} \nu_i \delta_{\lambda_i}$ and the random probability measure $F_\lambda = G_\lambda / G_\lambda([0,\infty])$ defined through normalization. Then F_λ can be interpreted as a mixing measure over the decays for the components of $\sigma^2(t)$ and the autocorrelation function is

$$\rho(s) = \sum_{i=1}^{\infty} w_i \exp\{-\lambda_i s\} = \int \exp\{-\lambda s\}\, \mathrm{d}F_\lambda(\lambda).$$

where $w_i = \frac{\nu_i}{\sum_{j=1}^{\infty} \nu_j}$. The discrete superposition in Equation (2) arises from taking $\nu_i = 0$ for $i > K$ and F_J to be a parametric distribution. Lemma 1 shows that the marginal distribution of $\sigma^2(t)$ has Lévy density

$$u(x) = I_J(x) \sum_{i=1}^{\infty} \nu_i = I_J(x) G_\lambda([0,\infty]) \tag{7}$$

where $I_J(x) = \int_1^\infty F_J(\tau x)\, \mathrm{d}\tau$. The choice of a compound Poisson process for z simplifies the calculation of the cumulant-generating function

$$\begin{aligned}\kappa(\theta) &= -\int_{0^+}^{\infty} (1 - \exp\{-\theta x\}) W(\mathrm{d}x) \\ &= -\sum_{i=1}^{\infty} \nu_i \int_{0^+}^{\infty} (1 - \exp\{-\theta x\}) F_J(\mathrm{d}x) \\ &= \sum_{i=1}^{\infty} \nu_i (\exp\{\kappa_J(\theta)\} - 1)\end{aligned}$$

where $\kappa_J(\theta)$ is the cumulant-generating function of F_J. This implies that the cumulant-generating function of $\sigma^2(t)$, $\kappa'(\theta)$, is

$$\kappa'(\theta) = \int_0^\infty \kappa(\theta \exp\{-s\})\, \mathrm{d}s = \sum_{i=1}^{\infty} \nu_i \int_0^\infty (\exp\{\kappa_J(\theta \exp\{-s\})\} - 1)\, \mathrm{d}s.$$

More straightforwardly, the cumulant of $\sigma^2(t)$ can be expressed as $\kappa'_m = \kappa_m / m$ and so

$$\kappa'_1 = \kappa_{J,1} = \mu_J \tag{8}$$

and

$$\kappa'_2 = \frac{1}{2}\left(\kappa_{J,2} + \kappa_{J,1}^2\right) = \frac{1}{2}(\sigma_J^2 + \mu_J^2) \tag{9}$$

where μ_J and σ_J^2 are the mean and variance of the shot distribution F_J, respectively.

This model has a BDLP which is a compound Poisson process and so the process can be expressed in terms of a marked Poisson process, to generalize Equation (5),

$$\sigma^2(t) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} I(\tau_{ij} < t) J_{ij} \exp\{-\lambda_i(t-\tau_{ij})\} \tag{10}$$

where $\tau_{i1}, \tau_{i2}, \tau_{i3}, \ldots$ follow a Poisson process with intensity $\nu_i\lambda_i$ and $J_{ij} \overset{i.i.d.}{\sim} F_J$.

The model is very general. Bayesian analysis of this model involves placing a prior on $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$ and either F_J in the nonparametric case or the parameters of F_J in the parametric case. For example, if F_J follows an exponential distribution with mean $1/\gamma$, then the marginal distribution of $\sigma^2(t)$ will be Gamma-distributed with shape $\sum_{i=1}^{\infty}\nu_i$ and mean $\sum_{i=1}^{\infty}\nu_i/\gamma$.

2.1 Prior for $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$

The prior for $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$ is given by assuming that $\boldsymbol{\nu}$ is the jumps of a Gamma process which has the Lévy density

$$\phi(\nu) = M\nu^{-1}\exp\{-\nu\}$$

where $M > 0$ and $\lambda_1, \lambda_2, \lambda_3, \ldots$ are i.i.d. from the distribution H_λ. It follows that F_λ follows a Dirichlet process (Ferguson 1973) with total mass $M_\lambda H_\lambda$. The weight of the i-th component, $w_i = \frac{\nu_i}{\sum_{j=1}^{\infty}\nu_j}$, measures the contribution of the i-th component of the superposition to the overall process. Although there are an infinite number of components, only a finite number will provide a nonneglible contribution to the superposition. Standard properties of the Dirichlet process imply that the number of such components is controlled by M_λ. Smaller values of M_λ are associated with a smaller number of nonnegligible components and a less even spread of weights. As M_λ becomes larger, F_λ will increasingly resemble the distribution H_λ. The parameter M_λ can be chosen to control the distribution of the number of nonneglible weights or a prior distribution could be chosen for M_λ and its value inferred from data.

The approach assumes that there are an infinite number of components in the superposition which avoids the need to choose a value of K in a finite superposition. It is potentially more realistic if we assume that there will be many components which give a small contribution to the volatility. The MCMC methods, in particular the key result in Equation (12), developed for inference in this paper could be extended to the case where the prior distribution for $\boldsymbol{\nu}$ is the jumps of a subordinator for which the Lévy density $\int_0^\infty \phi(x)\,\mathrm{d}x = \infty$. Then, F_λ follows a normalized random measure with independent increments (James, Lijoi, and Prünster 2009). This defines a much wider class of priors for ν such as the generalized gamma process (Brix 1999), whose normalized version is the normalized generalized gamma process (Lijoi, Mena, and Prünster 2007). Our aim is density estimation of F_λ, and it has been found in more standard density estimation that a Dirichlet process with an unknown M_λ is sufficiently flexible.

2.2 Prior for F_J

The distribution F_J is given a finite Pólya tree prior (Hanson 2006). Pólya tree priors were introduced by Ferguson (1974) and later revived by Lavine (1992, 1994) and Maudlin, Sudderth, and Williams (1992). Unlike other nonparametric priors, it is possible to place probability mass 1 on absolutely continuous distributions with a Pólya tree. The prior is defined by an infinite sequence of partitions of the support of the distribution. Lavine (1992) defines the sequence to be a dyadic partition and uses a centering distribution H_J to define the end points of the partition. Let $e_j(k) = \epsilon_1 \ldots \epsilon_j$ be the j-fold binary representation of $k-1$ then $B(e_j(k)) = (H_J^{-1}((k-1)2^{-j}), H_J^{-1}(k2^{-j}))$ for $k = 1, 2, \ldots, 2^j$. These sets, $B(e_j(1)), B(e_j(2)), \ldots, B(e_j(2^j))$, divide the support of H_J for every level j and $B(\epsilon_1 \ldots \epsilon_j) = B(\epsilon_1 \ldots \epsilon_j 0) \cup B(\epsilon_1 \ldots \epsilon_j 1)$. Therefore, the j-th level is formed by dividing each element of the partition at the $(j-1)$-th level into two parts with equal probability under H_J. A random distribution is constructed by allowing $p_{e_j(k)0} = p(X \in B(e_j(k)0) | X \in B(e_j(k)))$ which is the probability that the random variable is in one of the children given that it's in a parent and $p_{e_j(k)1} = 1 - p_{e_j(k)1}$. Under the parametric distribution, H_J, this probability is always $\frac{1}{2}$ due to the construction of the partition and so this probability in the nonparametric model is defined to be

$$p_{e_j(k)0} \sim \mathrm{Be}(a_j, a_j)$$

where $\mathrm{Be}(a, b)$ represents a beta distribution with parameters a and b. Since the mean of $p_{e_j(k)0} = \frac{1}{2}$, the nonparametric model is centered over the parametric model in the sense that $\mathrm{E}[F_J(B)] = H_J(B)$. Then

$$p(B(\epsilon_1 \ldots \epsilon_j)) = \prod_{k=1}^{j} p_{\epsilon_1 \ldots \epsilon_k}.$$

In practice, we choose a finite number of levels J for the tree and assume that f_J follows h_J within each set in the J-th partition so that

$$f_J(x) = h_J(x) 2^J \sum_{i=1}^{2^J} \pi_i I(x \in H_J^{-1}((i-1)/2^J), H_J^{-1}(i/2^J)) \tag{11}$$

where $\pi_i = p(B(e_J(i)))$ is the probability that an observation falls into i-th element of the partition at the J-th level.

This defines a prior for F_J. We are also interested in the prior that is induced on the marginal distribution of $\sigma^2(t)$. Clearly, $\mathrm{E}_{F_J}[\log(1+x)]$ is finite if $\mathrm{E}_H[\log(1+x)]$ is finite. Using Equation (7), the Lévy density of $\sigma^2(t)$ is given by

$$u(x) = \sum_{i=1}^{\infty} \nu_i \sum_{i=1}^{M} \pi_i 2^M \frac{1}{x} \left(\exp\{-\gamma \max\{x, t_{i-1}\} - \exp\{-\gamma \max\{x, t_i\}\}\right)$$

which follows from

$$\begin{aligned} u(x) &= \int_1^{\infty} w(\tau x)\,\mathrm{d}\tau \\ &= \sum_{i=1}^{\infty} \nu_i \int_1^{\infty} f_J(\tau x)\,\mathrm{d}\tau \\ &= \sum_{i=1}^{\infty} \nu_i 2^M \sum_{i=1}^{2^M} \pi_i \int_1^{\infty} h_J(\tau x) I(\tau x \in H^{-1}((i-1)/2^M), H^{-1}(i/2^M))\,\mathrm{d}\tau \\ &= \sum_{i=1}^{\infty} \nu_i 2^M \frac{1}{x} \sum_{i=1}^{2^M} \pi_i \int_x^{\infty} h_J(z) I(z \in H_J^{-1}((i-1)/2^M), H_J^{-1}(i/2^M))\,\mathrm{d}z \\ &= \sum_{i=1}^{\infty} \nu_i 2^M \frac{1}{x} \sum_{i=1}^{2^M} \pi_i H_J(\max\{x, i/2^M\}) - H_J(\max\{x, (i-1)/2^M\}) \end{aligned}$$

Equations (8) and (9) can be used to show that the mean of $\sigma^2(t)$ is

$$\kappa'_{J,1} = \mu_J = \frac{\alpha}{\gamma} 2^J \sum_{i=1}^{2^J} \pi_i \left[\Gamma(\alpha+1, \gamma H_J^{-1}(i/2^J)) - \Gamma(\alpha+1, \gamma H_J^{-1}((i-1)/2^J))\right]$$

and the variance of $\sigma^2(t)$ is

$$\kappa'_{J,2} = \frac{\alpha(\alpha+1)}{\gamma^2} 2^J \sum_{i=1}^{2^J} \pi_i \left[\Gamma(\alpha+2, \gamma H_J^{-1}(i/2^J)) - \Gamma(\alpha+2, \gamma H_J^{-1}((i-1)/2^J))\right].$$

2.3 Further Specification of the Bayesian Model

The Bayesian model is fully specified by including priors for the parameters and hyperparameters of the model. The parameters μ, β, ρ are given a $\mathrm{N}(0, 100I)$ prior which can be considered uninformative for the usual values of these parameters. Several priors have been proposed for M_λ. The most popular choice is $M_\lambda \sim \mathrm{Ga}(a, b)$ (Escobar and West 1995) which will be used in this paper. The hyperparameter $a = 1$ and the mean $a/b = 3$ which give a prior mass on a reasonable range of values. In parametric case when F_J is an exponential distribution with mean $1/\gamma$, then γ is given a vague prior $\mathrm{Ga}(0.001, 0.001)$ as in Griffin and Steel (2010). Similarly, if F_J is given a Pólya tree prior, then the scale γ of the exponential centering distribution is given the same prior distribution. A poor choice of the centering distribution H_λ (such as a vague prior) can lead to poor inference about F_λ and so H_λ is given an informative choice. Previous analyses using a single OU process suggest a value of λ around 0.01. The introduction of extra components in a superposition usually has larger values of λ. The distribution $\text{Ga}(0.4, 20)$ has a mean of 0.08, and an upper 0.1% point of the distribution is 0.25. This puts mass 0.01 but also allows processes which much shorter-term behavior. The results presented in Section 5 were robust to reasonable changes in H_λ.

Following Eraker, Johannes, and Polson (2003), the model for jumps in returns is changed slightly since λ_J is typically small so that the probability of observing more than one jump in any Δ period is very small. This suggests an alternative model where the number of jumps in observation period j is m_j and $p(m_j = 0) = 1 - \kappa_J$ and $p(N_j = 1) = \kappa_J$ where $\kappa_J \approx \exp\{-\lambda_J \Delta\}$. The priors for σ_J^2 and κ_J are those suggested by Eraker, Johannes, and Polson (2003), $\sigma_J^2 \sim \text{IG}(5, 20)$ and $\kappa_J \sim \text{Be}(2, 40)$.

3 COMPUTATIONAL METHODS

The observed data are the returns over a period of length Δ, $r_1, r_2, \ldots, r_T$ where $r_t = \log(y(t\Delta)) - \log y((t-1)\Delta)$. The model in Equation (3) implies that $r_1, r_2, \ldots, r_T$ are mutually independent with

$$r_t \sim \text{N}\left(\mu\Delta + \beta\sigma_n^2 + \rho z_n + m_t \eta_t, \sigma_n^2\right)$$

where $\sigma_n^2 = \int_{(n-1)\Delta}^{n\Delta} \sigma^2(t)\text{d}t$, $z_n = \sum_{i=1}^{\infty} \int \text{d}z(\lambda_i n\Delta) - \int \text{d}z(\lambda_i(n-1)\Delta) - \text{E}[z(\lambda_i \Delta)]$, $\eta_t \sim \text{N}(0, \sigma_J^2)$, $m_t = 1$ with probability κ_J and $m_t = 0$ with probability $1 - \kappa_J$. Efficient MCMC samplers for stochastic volatility models can be constructed by treating the volatilities $\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2$ as missing data and including their values in a Gibbs sampler. In these models, it is simpler to use the representation as marked Poisson process in Equation (10). Then J and τ are treated as missing values.

The process is doubly infinite since there is an infinite number of Poisson processes which will each contain an infinite number of points. If the BDLP is chosen to lead to a particular marginal distribution for $\sigma_i^2(t)$, then the OU process can be expressed in terms of $\sigma_i^2(0)$ and a finite number of jumps. The nonparametric model is defined through its BDLP and so we will follow Gander and Stephens (2007a) and Griffin and Steel (2010) by restricting the times τ_{ij} to the region (B, T) for all i where $T = n\Delta$ and B is chosen to avoid a large truncation error. This defines an approximate process

$$\tilde{\sigma}^2(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{N_i} \text{I}(\tau_{ij} < t) \exp\{-\lambda_i(t - \tau_{ij})\} J_{ij}$$

where N_i is the number of points in a Poisson process on (B, T) with intensity $\nu_i \lambda_i$ whose points are $\tau_{i1}, \tau_{i2}, \ldots, \tau_{iN_i}$.

This removes one infinite sum in the form of $\sigma^2(t)$, but there is still an infinite number of Poisson processes which cannot all be updated in the Gibbs sampler. However, a finite set of processes can be found which contains all Poisson processes with one or more points in (B, T). Suppose that $\nu_1 > \nu_2 > \nu_3 > \ldots$ then there is a D for which $\sum_{i=D+1}^{\infty} N_i = 0$ and $N_D > 0$. Therefore, $\tilde{\sigma}_i^2(t)$ can be calculated exactly as

$$\tilde{\sigma}_i^2(t) = \sum_{i=1}^{D} \sum_{j=1}^{N_i} \mathrm{I}(\tau_{ij} < t) \exp\{-\lambda_i (t - \tau_{ij})\} J_{ij}.$$

The probability mass function of D, conditional on $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$, is

$$p(D = k) = \begin{cases} \exp\{-(T-B)\sum_{i=1}^{\infty} \nu_i \lambda_i\} & D = 0 \\ (1 - \exp\{-(T-B)\lambda_D \nu_D\}) \exp\{-(T-B)\sum_{i=D+1}^{\infty} \lambda_i \nu_i\} & \text{otherwise} \end{cases}$$

and that the probability of $N_1, N_2, \ldots, N_D, D$ is

$$\prod_{i=1}^{D} \frac{\exp\{-(T-B)\nu_i\lambda_i\}((T-B)\nu_i\lambda_i)^{N_i}}{N_i!} \exp\left\{-(T-B)\sum_{i=D+1}^{\infty} \nu_i \lambda_i\right\}$$

where $N_D > 0$. This probability involves the infinite sequence $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$ and so cannot be calculated. However, conditioning only on $\nu_1, \nu_2, \ldots, \nu_D$ and $\lambda_1, \lambda_2, \ldots, \lambda_D$ and integrating over $\nu_{D+1}, \nu_{D+2}, \ldots$ and $\lambda_{D+1}, \lambda_{D+2}, \ldots$ leads to the following expression for the probability of $N_1, N_2, \ldots, N_D, D$

$$\prod_{i=1}^{D} \frac{\exp\{-(T-B)\nu_i\lambda_i\}((T-B)\nu_i\lambda_i)^{N_i}}{N_i!} \mathrm{E}\left[\exp\left\{-(T-B)\sum_{i=D+1}^{\infty} \nu_i \lambda_i\right\}\right].$$

The expectation is taken over the jumps smaller than ν_D which is a Lévy process and can be calculated using the results of Regazzini, Guglielmi, and Di Nunno (2002) and Regazzini, Lijoi, and Prünster (2003) who show that

$$\mathrm{E}\left[\exp\left\{-(T-B)\sum_{i=D+1}^{\infty} \nu_i \lambda_i\right\}\right] = \exp\left\{-M_\lambda \int_0^{\nu_D} (1 - \psi_{H_\lambda}((T-B)y)) y^{-1} \exp\{-y\}\, dy\right\} \quad (12)$$

where $\psi_{H_\lambda}(y)$ is the moment-generating function of H_λ, which will often be known. The integral is univariate and can be calculated using standard numerical integration techniques.

The ordered sequence of jumps of a Gamma process ν can be simulated using the representation described by Ferguson and Klass (1972). Let $W^{+}(x) = \int_x^\infty y^{-1} \exp\{-x\}\, \mathrm{d}x$ and define W^{-1} to be the inverse of W^{+}. Then

$$\nu_i = W^{-1}(u_i)$$

where $u_1, u_2, u_3, \ldots$ is a Poisson process with intensity M (this methods can be simply extended to other completely random measures).

3.1 MCMC Algorithm

MCMC methods for OU process and superpositions of OU processes are discussed by Roberts, Papaspiliopoulos, and Dellaportas (2004), Griffin and Steel (2006), and Gander and Stephens (2007a,b). The main problem with inference is the difficulty of updating the parameters λ and ν that appear in the Poisson process. Both Roberts, Papaspiliopoulos, and Dellaportas (2004) and Griffin and Steel (2006) describe methods that jointly update the Poisson process with these parameters that lead to better mixing algorithm than a Gibbs sampler without joint updates. The model will be fitted using $\tilde{\sigma}_1^2, \tilde{\sigma}_2^2, \ldots, \tilde{\sigma}_T^2$ rather than $\sigma_1^2, \sigma_2^2, \ldots, \sigma_T^2$. It is useful to note that

$$p(r|J, \tau, \nu, \lambda, \gamma) = \prod_{t=1}^{T} \mathrm{N}\left(r_t | \mu\Delta + \beta\tilde{\sigma}_t^2 + \rho z_t + m_t\eta_t, \tilde{\sigma}_t^2\right)$$

where $\tilde{\sigma}_n^2 = \tilde{\sigma}^{\star 2}(n\Delta) - \tilde{\sigma}^{\star 2}((n-1)\Delta)$ and

$$\tilde{\sigma}^{\star 2}(n\Delta) = \sum_{i=1}^{D} \frac{1}{\lambda_i} \sum_{j=1}^{N_i} I(\tau_{ij} < n\Delta) J_{ij} \left(1 - \exp\{-\lambda_i(n\Delta - \tau_{ij})\}\right).$$

If a Pólya tree prior is given to F_J, then the density is given by Equation (11) and

$$\mu_J = \frac{1}{\gamma} 2^J \sum_{i=1}^{2^J} (\exp\{-\gamma H_J^{-1}((i-1)/2^J)\} - \exp\{-\gamma H_J^{-1}(i/2^J)\}).$$

A value from F_J can be simply simulated using the inversion method. The weights in the Pólya tree can be initialized from their prior distribution so that $p(e_j(k)0) \sim \mathrm{Be}(M_j, M_j)$ for $1 \leq k \leq 2^{j-1}$ and $1 \leq j \leq J$.

Updating *J*. The jump sizes $\{J_{ij}\}_{1 \leq i \leq D, 1 \leq j \leq N_i}$ can be updated one by one using a Metropolis–Hastings random walk. Suppose that we update J_{ij}, then a new value J'_{ij} is proposed for which $\log J'_{ij} = \log J_{ij} + \epsilon$ where $\epsilon \sim \mathrm{N}(0, \sigma_J^2)$. The proposed value is accepted with probability

$$\min\left\{1, \frac{p(r|J', \tau, \nu, \lambda, \gamma) J'_{ij} f_J(J'_{ij})}{p(r|J, \tau, \nu, \lambda, \gamma) J_{ij} f_J(J_{ij})}\right\}$$

The value of σ_J^2 is chosen so that the acceptance rate is between 0.2 and 0.3.

Updating $\boldsymbol{\tau}$. The jump times $\{\tau_{ij}\}_{1 \le i \le D, 1 \le j \le N_i}$ can be updated one by one using a Metropolis–Hastings random walk. Suppose that we update τ_{ij}, then a new value τ'_{ij} is proposed by $\tau'_{ij} = \tau_{ij} + \epsilon$ where $\epsilon \sim \mathrm{N}(0, \sigma_\tau^2)$. The proposed value is rejected if $\tau'_{ij} < B$ or $\tau'_{ij} > T$. Otherwise, the new jumps is accepted with probability

$$\min\left\{1, \frac{p(r|J, \tau', \nu, \lambda, \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)}\right\}.$$

The value of σ_τ^2 is chosen so that the acceptance rate is between 0.2 and 0.3.

Updating $N_1, N_2, \ldots, N_D$. The value of N_i is updated using a reversible jump Metropolis–Hastings step. A value i is chosen uniformly from the set $\{1, 2, \ldots, D\}$ and a move Add or Delete is chosen at random with probability $\frac{1}{2}$. In the Add move, a new jump with size J_{i,N_i+1} is proposed from F_J and a jump time τ_{i,N_i+1} is proposed uniformly on $[B, T]$; then the new process is accepted with probability

$$\min\left\{1, \frac{p(r|J', \tau', \nu, \lambda, \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)} \frac{\nu_i \lambda_i (T - B)}{N_i + 1}\right\}.$$

In the Delete move, a jump is chosen uniformly in $\{1, 2, \ldots, N_i\}$ and it is proposed to be deleted from the i-th process. The move is rejected if $i = D$ and $N_D = 1$. Otherwise, this move is accepted with probability

$$\min\left\{1, \frac{p(r|J', \tau', \nu, \lambda, \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)} \frac{N_i}{\nu_i \lambda_i (T - B)}\right\}.$$

Updating $\boldsymbol{\lambda}$. We update $\lambda_1, \lambda_2, \ldots, \lambda_m$ separately using the method of Roberts, Papaspiliopoulos, and Dellaportas (2004). Suppose that λ_j is updated. A new value λ'_j is proposed by perturbing λ_j, $\log \lambda'_j = \log \lambda_j + \epsilon$ where $\epsilon \sim \mathrm{N}(0, \sigma_\lambda^2)$. If $\lambda'_j > \lambda_j$, then $K' \sim \mathrm{Pn}(\nu_i(\lambda'_i - \lambda_i)(T - B))$ new jumps are simulated where $\tau'_1, \tau'_2, \ldots . \tau'_{K'_j}$ are uniformly distributed on $[B, T]$ and $J'_1, J'_2, \ldots, J'_{K'_j} \overset{i.i.d.}{\sim} F_J$. If $\lambda'_j < \lambda_j$, the Poisson process of jumps is thinned by deleting the i-th jumps with probability $1 - \frac{\lambda'}{\lambda}$. In both cases, the proposed value accepted with probability is

$$\min\left\{1, \frac{p(r|J', \tau', \nu, \lambda', \gamma)\lambda_j' h_\lambda(\lambda_j')}{p(r|J, \tau, \nu, \lambda, \gamma)\lambda_j h_\lambda(\lambda_j)}\right\}.$$

The value of σ_λ^2 is chosen so that the acceptance rate is between 0.2 and 0.3.

Updating ν. We update $\nu_1, \nu_2, \ldots, \nu_m$ separately. The parameter ν_j is updated in the following way. If $N_j = 0$, then $nu_j' = \nu_j + \epsilon$ where $\epsilon \sim \mathrm{N}(0, \sigma_\nu^2)$. Let K_{min} be the index of the smallest value of $\nu_1, \ldots, \nu_{j-1}, \nu_j, \nu_{j+1}, \ldots, \nu_D$ and K'_{min} the index of the smallest value of $\nu_1, \ldots, \nu_{j-1}, \nu_j', \nu_{j+1}, \ldots, \nu_D$; if $N_{k_{min}} = 0$, then the move is rejected. Otherwise, the move is accepted with probability

$$\min\left\{1, \frac{p(r|J, \tau, \nu', \lambda, \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)} \exp\{-(\nu_j' - \nu_j)(1 + (T - B)\lambda_j)\alpha^\star\}\right\}$$

where

$$\alpha^\star = \begin{cases} -M_\lambda I^\star(\nu_{K_{min}}, \nu_{K'_{min}}) - M_\lambda E_1(\nu_{K'_{min}}) + M_\lambda E_1(\nu_{K_{min}}) & \text{if } \nu_{K'_{min}} > \nu_{K_{min}} \\ 0 & \text{if } \nu_{K'_{min}} = \nu_{K_{min}} \\ -M_\lambda I^\star(\nu_{K'_{min}}, \nu_{K_{min}}) - M_\lambda E_1(\nu_{K'_{min}}) + M_\lambda E_1(\nu_{K_{min}}) & \text{if } \nu_{K'_{min}} < \nu_{K_{min}} \end{cases}$$

and

$$I^\star(\nu_1, \nu_2) = \int_{\nu_1}^{\nu_2} (1 - \phi_{H_\lambda}((T - B)y)y^{-1} \exp\{-y\}\, dy$$

and $E_1(x) = \int_x^\infty y^{-1} \exp\{-y\}\, dy$. The value of σ_ν^2 is chosen so that the acceptance rate is between 0.2 and 0.3.

If $N_j \neq 0$, then ν_j' is proposed from $\mathrm{Ga}(N_j, 1 + (T - B)\lambda_j)$ which is accepted with probability

$$\min\left\{1, \frac{p(r|J, \tau, \nu', \lambda, \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)}\alpha^\star\right\}.$$

Updating γ. The parameter γ can be updated using a Metropolis–Hastings independence sampler where a new value γ' is proposed where $\gamma' \sim \mathrm{Ga}\left(\gamma_0 + \sum_{i=1}^D N_i, \gamma_1 + \sum_{i=1}^D \sum_{j=1}^{N_i} J_{ij}\right)$. The proposed value is accepted with probability

$$\min\left\{1, \frac{p(r|J, \tau, \nu, \lambda, \gamma')f_J(J|\gamma')}{p(r|J, \tau, \nu, \lambda, \gamma')f_J(J|\gamma)}\right\}$$

Updating μ, β, and ρ. These parameters can be updated jointly by sampling from the full conditional distribution

$$(\mu, \beta, \rho)^T \sim \mathrm{N}\left((X^T\Sigma^{-1}X + \Lambda^{(0)})^{-1}X^T\Sigma^{-1}r, (X^T\Sigma^{-1}X + \Lambda^{(0)})^{-1}\right)$$

where Σ is a $T \times T$ diagonal matrix with nonzero elements

$$\Sigma_{ii} = \sigma_i^2, \qquad i = 1, 2, \ldots, T,$$

$\Lambda^{(0)}$ is a 3×3 diagonal matrix with nonzero elements

$$\Lambda_{11}^{(0)} = \frac{1}{\sigma_\mu^2}, \quad \Lambda_{22}^{(0)} = \frac{1}{\sigma_\beta^2}, \Lambda_{33}^{(0)} = \frac{1}{\sigma_\rho^2}$$

and X is $T \times 3$-dimensional matrix with elements

$$X_{i1} = \Delta, \qquad X_{i2} = \tilde{\sigma}_i^2, X_{i3} = \sum_{j=1}^{D}\sum_{k=1}^{N_j} J_{jk}\mathrm{I}((i-1)\Delta < \tau_{jk} < i\Delta) - \sum_{j=1}^{D}\mathrm{E}[z(\lambda_j\Delta)]$$

Updating D. The parameter D is updated using a reversible jump Metropolis–Hastings steep. We propose to Add a new component to the superposition ($D' = D + 1$) with probability $1/2$ or to Delete the component in the superposition ($D' = D - 1$). If we choose to Add, then we form ν' by generating $\nu' = W^{-1}\left(W^{+}(\nu_D) + x\right)$ where x is simulated from an exponential distribution with mean $1/M_\lambda$ and generate a value λ' from H_λ. The proposed values will be N', τ', J', ν', and λ' which are formed as follows

$$N_i' = N_i, 1 \leq i \leq D-1, N_D' = 0, N_{D+1}' = N_D$$

$$\nu_i' = \nu_i, 1 \leq i \leq D, \nu_{D+1}' = \nu'$$

$$\lambda_i' = \lambda_i, 1 \leq i \leq D-1, \lambda_D' = \lambda', \lambda_{D+1}' = \lambda_D$$

$$J_i' = J_i, 1 \leq i \leq D-1, J_D' =, J_{D+1}' = J_D$$

$$\tau_i' = \tau_i, 1 \leq i \leq D-1, \tau_D' =, \tau_{D+1}' = \tau_D.$$

The proposed new cluster is accepted with probability

$$\min\left\{1, \frac{p(r|J', \tau', \nu', \lambda', \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)}\exp\{-c_1(T-B)\}\left(\frac{\nu'\lambda_D}{\nu_D\lambda'}\right)^{N_D}\exp\{I_1\}\right\}$$

where

$$c_1 = \nu' \lambda_D + \nu_D \lambda' - \nu_D \lambda_D$$

and

$$I_1 = M_\lambda \int_{\nu'}^{\nu_D} \phi_{H_\lambda}((T-B)x)x^{-1} \exp\{-x\}\, \mathrm{d}x.$$

If we choose to delete a component so that $D' = D - 1$, the move is rejected if $N_{D-1} \neq 0$. Otherwise,

$$N'_i = N_i, 1 \leq i \leq D-2, N'_{D-1} = N_D$$

$$\nu'_i = \nu_i, 1 \leq i \leq D-1$$

$$\lambda'_i = \lambda_i, 1 \leq i \leq D-2, \lambda'_{D-1} = \lambda_D$$

$$J'_i = J_i, 1 \leq i \leq D-2, J'_{D-1} = J_D$$

$$\tau'_i = \tau_i, 1 \leq i \leq D-2, \tau'_{D-1} = \tau_D.$$

The proposal is accepted with probability

$$\min\left\{1,, \frac{p(r|J', \tau', \nu', \lambda', \gamma)}{p(r|J, \tau, \nu, \lambda, \gamma)} \exp\{c_2(T-B)\} \left(\frac{\nu_{D-1}\lambda_{D-1}}{\nu_D \lambda_D}\right)^{N_{D-1}} \exp\{-I_2\}\right\}$$

where

$$c_2 = \nu_D \lambda_D + \nu_{D-1}\lambda_{D-1} - \nu_{D-1}\lambda_D$$

and

$$I_2 = M_\lambda \int_{\nu_D}^{\nu_{D-1}} \phi_{H_\lambda}((T-B)x)x^{-1} \exp\{-x\}\, \mathrm{d}x$$

Updating $m_1, m_2, \ldots, m_T$. The full conditional distribution of m_i is

$$p(m_i = 0) = (1-\kappa_J)(\sigma_i^2)^{-1/2} \exp\left\{-\frac{(r_i - \mu - \beta\sigma_i^2 - \rho(z_i - \mathrm{E}[z_i]))^2}{2\sigma_i^2}\right\}$$

and

$$p(m_i = 1) = \kappa_J(\sigma_i^2 + \sigma_J^2)^{-1/2} \exp\left\{-\frac{(r_i - \mu - \beta\sigma_i^2 - \rho(z_i - \mathrm{E}[z_i]))^2}{2(\sigma_i^2 + \sigma_J^2)}\right\}.$$

Updating κ_J. The full conditional distribution of κ_J is $\mathrm{Be}(2 + \sum_{i=1}^T m_i, 40 + T - \sum_{i=1}^T m_i)$.

Updating σ_J^2. Firstly, if $m_i > 0$, simulate $\eta_i \sim \mathrm{N}\left(\frac{(r_i-\mu-\beta\sigma_i^2-\rho(z_i-\mathrm{E}[z_i]))^2/\sigma_i^2}{1/\sigma_i^2+1/\sigma_J^2}, \frac{1}{1/\sigma_i^2+1/\sigma_J^2}\right)$, then the full conditional distribution of σ_J^2 is $\mathrm{IG}(5 + \frac{1}{2}\sum_{i=1}^T m_i, 20 + \frac{1}{2}\sum_{\{i|N_i=1\}} \eta_i^2)$.

Updating the weights in the Pólya tree. The value of $p(e_j(k)0) \sim \mathrm{Be}(a_j + n(B(e_j(k)0)), a_j + n(B(e_j(k)1)))$ for $1 \leq k \leq 2^{j-1}$ and $1 \leq j \leq J$. where $n(A)$ is the number of jumps $J_{m,l}$ falling into set A for $1 \leq m \leq D$ and $1 \leq l \leq N_k$. Once these are generate, only $\pi_1, \pi_2, \ldots, \pi_{2^J}$ needs to be stored for which

$$\pi_j = \prod_{k=1}^{J} p_{\epsilon_1 \ldots \epsilon_k}$$

where the $\epsilon_1 \ldots \epsilon_J$ is the J-fold binary representation of $(j-1)$.

4 ANALYSIS OF STOCK INDEX DATA

The models developed in this paper were applied to four stock indices. The Standard and Poors 500 (S&P 500) for the New York and NASDAQ stock exchanges from April 28, 1989, to February 26, 2009, the NASDAQ 100 for the NASDAQ stock exchange from April 28, 1989, to February 26, 2009, the FTSE 100 for the London stock exchange from May 17, 1988, to February 26, 2009, and the Nikkei 225 for the Tokyo Stock Exchange from October 4, 1988, to February 26, 2009. This represented a range of indices: two indices involved the American economy, one involved the UK economy, and one referred to the Japanese economies. Each data set contained 5000 observations which should have sufficient information for useful nonparametric inference. Two specific models were considered. Firstly, an infinite superposition of OU processes where $\sigma^2(t)$ had a Gamma distribution (which will be referred to as the Inf-Sup OU Gamma model) and an infinite superposition of OU processes where the shot distribution F_J was given a nonparametric prior (which will be referred to as the Inf-Sup NP OU model). In the Inf-Sup NP OU model, the number of levels in the Pólya tree, J, was set to be 7, which leads to a J-th level partition with 128 bins and substantial flexibility. The parameter setting

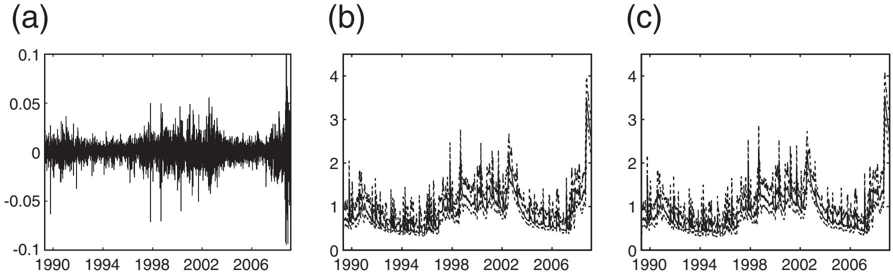

Figure 1 The S&P 500 index between April 28, 1989, and February 26, 2009: (a) the returns, (b) the posterior median of volatility (solid line) with 95% credible interval (dashed line) using the Inf-Sup Gamma OU model, and (c) the posterior median of volatility (solid line) with 95% credible interval (dashed line) using the Inf-Sup NP OU model.

$a_j = 5/2^j$ guarantees that F_J is an absolutely continuous distribution if the Pólya tree were allowed an infinite number of levels. The MCMC algorithms were run for 300,000 iterations with a thinning of 30.

4.1 Standard and Poors 500

The S&P 500 index has been well-studied using stochastic volatility models. Here, the returns of the index from April 28, 1989, until February 26, 2009, were analyzed. The data are shown in Figure 1a and indicate periods of relatively high volatility between 1997 and 2004 and from 2008 onwards. These were well-captured by the posterior volatility estimates shown in panels (b) from the Inf-Sup Gamma OU model and (c) from the Inf-Sup NP OU model. The volatility estimates are virtually indistinguishable for the two models.

Posterior estimates for various parameters of the model are shown in Table 1. The expectation of $\sigma^2(t)$ was estimated to be slightly smaller for Inf-Sup NP OU and the standard deviation to be slightly larger. Other parameters were estimated to be very similar for the two models. The leverage effect ρ was negative and the 95% credible interval was far from zero. The parameter M_λ had a posterior median of 2.44 for Inf-Sup Gamma OU and 2.78 for Inf-Sup NP OU which suggests that the measure F_λ is dominated by a small number of jumps with nonneglible mass. Both models show that there are a small number of jumps in returns (the posterior median of κ_J is 0.0058 which implies that the average number of jumps in the series is 29).

Figure 2 shows inference about the two main components of the models: the measure F_λ and the jump distribution F_J. Panel (b) shows the posterior mean of F_J which indicates that there is only a small difference between the Inf-Sup Gamma OU and Inf-Sup NP OU models. Panel (a) shows the posterior mean of F_λ, the distribution derived by normalizing F_λ shown on the log scale. The estimate for the Inf-Sup Gamma OU model has a clear trimodal shape with the modes occurring

Table 1 The S&P 500 index between 28/4/89 and 26/2/09: posterior median of the parameters with 95% credible interval shown in brackets.

	Inf-Sup Gamma OU	Inf-Sup NP OU
$E[\sigma_n^2]$	1.37 (0.81, 2.75)	1.26 (0.77, 2.43)
$SD[\sigma_n^2]$	0.86 (0.44, 2.15)	1.02 (0.46, 2.53)
$E[J]$	0.63 (0.47, 0.92)	0.53 (0.32, 0.79)
$E[\lambda]$	0.023 (0.011, 0.044)	0.026 (0.012, 0.048)
μ	0.041 (0.0093, 0.0725)	0.044 (0.0111, 0.0782)
β	−0.022 (−0.061, 0.0189)	−0.022 (−0.062, 0.0154)
ρ	−1.46 (−2.02, −1.03)	−1.44 (−1.94, −0.93)
M_λ	2.44 (0.85, 5.04)	2.78 (0.84, 6.18)
λ_J	0.0058 (0.0007, 0.0253)	0.0065 (0.0009, 0.0277)
σ_J^2	2.65 (1.36, 6.54)	2.66 (1.36, 6.25)

at roughly −7.5, −5, and −2.5 or 0.00053, 0.0067, and 0.082 for λ. One way to interpret these values is the half-life of the effect of a jump in volatility. The effect of the j-th jump in the i-th component is $J_{ij}\exp\{-\lambda_i(t-\tau_{ij})\}$ for $t > \tau_{ij}$. This effect is halved at time $t = \tau_{ij} + \frac{\log 2}{\lambda_i}$ and so the half-life is $\frac{\log 2}{\lambda_i}$. The half-lives for three modes are 1253.2, 102.9, and 8.4. The Inf-Sup NP OU model showed a similar distribution but the low-frequency mode was less distinct (although the mass around −7.5 was similar for the two models).

The three modes found in these fits suggested decomposing the volatility into three components that represent the mass around the three modal values. Figure 3 shows the decomposition of the volatility according to the value of λ for each component in the infinite superposition. The "low"-frequency category uses components with $\lambda < 0.0025$, the "medium"-frequency category uses components with $0.0025 < \lambda < 0.050$, and the "high"-frequency category uses components with $\lambda > 0.050$. The low- and high-frequency processes represent long-term

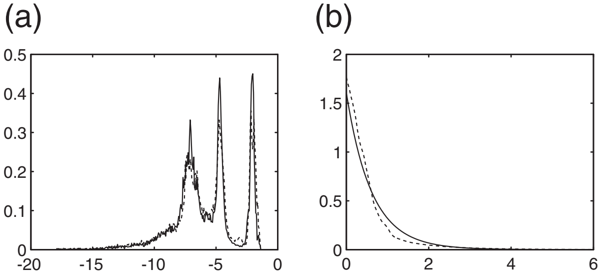

Figure 2 The S&P 500 index between April 28, 1989, and February 26, 2009: (a) the posterior mean of F_λ and (b) posterior mean of F_J using the Inf-Sup Gamma OU (solid line) and the Inf-Sup NP OU (dashed line).

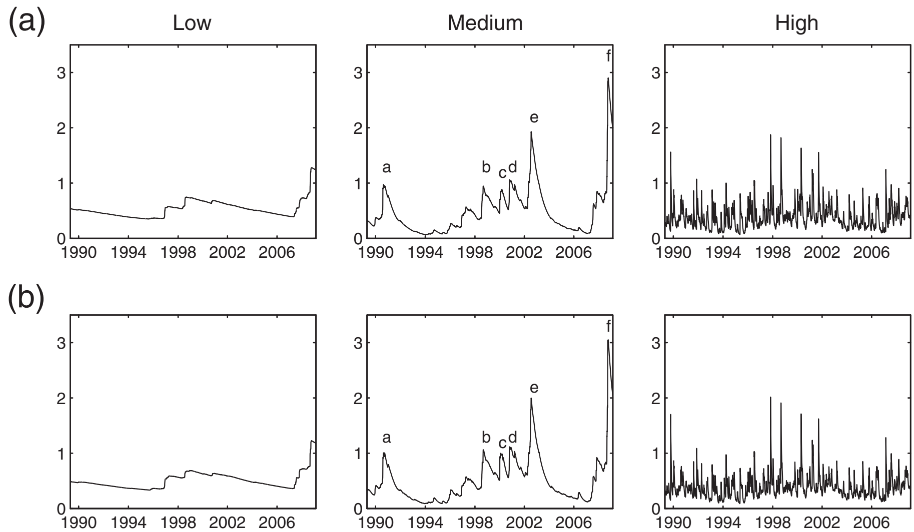

Figure 3 The S&P 500 index between April 28, 1989, and February 26, 2009: posterior mean volatility in the three frequency categories using (a) the Inf-Sup Gamma OU model and (b) the Inf-Sup NP OU model.

and short-term changes in the economy. The medium category contains the main large movements due to economic factors. The large jumps in volatility are marked on graphs for the medium category in Figure 3 and are linked to large movements in the index. Jump a is a 3% drop on August 28, 1990, due to fear of instability in the Middle East following Iraq's invasion of Kuwait, jump b is a 7% drop on August 31, 1998, due to the economic fear following the Russian default, jump c is a 4% drop on January 4, 2000, due to fear of a long-term rises in interest rises, jump d is on October 13, 2000, due to fears of oil price rises in the run-up to the Iraq war, jump e is 4% drop on July 18, 2002, due to market nervousness at the time of the collapse of WorldCom, and jump f is due to 9% drop following the rejection of bailout plan by the U.S. senate.

4.2 Other Indices

Three more indices were analyzed over the same period: the NASDAQ 100 (April 28, 1989, until February 26, 2009), FTSE 100 (from May 17, 1989, until February 26, 2009), and the Nikkei 225 (November 4, 1988, until February 26, 2009). All time series contained 5000 observations (the same length as the S&P 500 index analyzed in the previous subsection). Some additional results are presented in the Appendix A. Plots of the log returns for each index is shown in Figure 4.

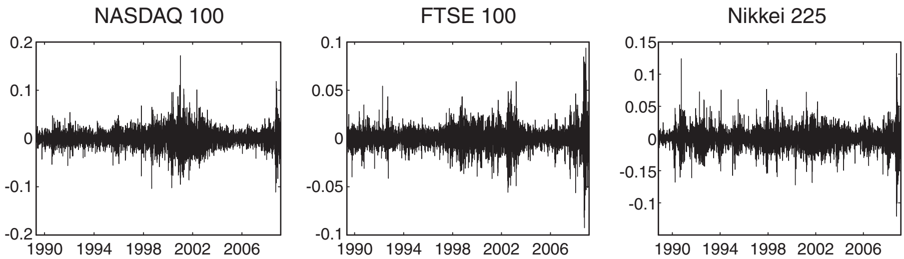

Figure 4 The returns for the NASDAQ 100, FTSE 100, and Nikkei 225 indices.

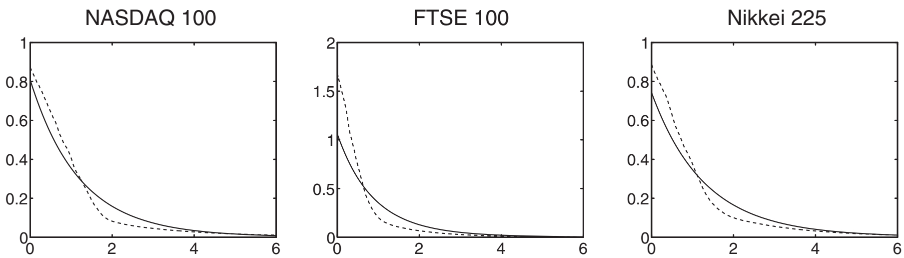

Figure 5 The posterior mean of F_J for the NASDAQ 100, FTSE 100, and Nikkei 225 indices using the Inf-Sup Gamma OU (solid line) and the Inf-Sup NP OU (dashed line).

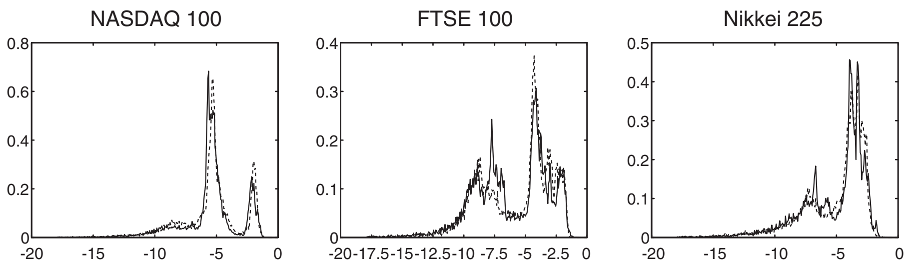

Figure 6 The posterior mean of F_λ for the NASDAQ 100, FTSE 100, and Nikkei 225 indices using the Inf-Sup Gamma OU (solid line) and the Inf-Sup NP OU (dashed line).

Figure 5 shows the posterior mean of F_J using the parametric and nonparametric mean for the three indices. The inference from the NASDAQ 100 for the distribution was similar using the parametric and nonparametric model. However, inference with the FTSE 100 index showed a difference between the estimate using the nonparametric model and the parametric model. The results showed a lot more mass close to zero and much less mass above 2. The inference about F_J with the Nikkei 225 also showed more mass close to zero and less mass for larger values of J.

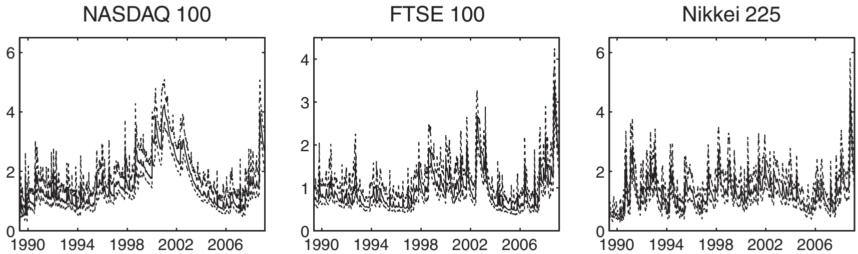

Figure 7 Posterior median of the volatility (solid line) with 95% credible interval (dashed line) for the NASDAQ 100, FTSE 100, and Nikkei 225 indices using the Inf-Sup NP OU.

Figure 6 shows the posterior mean of F_λ for the three indices under the parametric and nonparametric models. The results were very close under the two models with each data set and had a multimodal shape (like the results for the S&P 500 index). The distribution for the NASDAQ 100 was again trimodal but the mass around each mode was rather different. The NASDAQ 100 has a much smaller mode around the values of -7.5 and -2.5 and a much larger mass around -5. The estimates for the FTSE 100 index again showed a trimodal shape with modes around -7.5, -5, and -2.5 which fell in similar proportion to the estimates for the S&P 500 index under the parametric model. However, in contrast to the results for the S&P 500 index, there are clear differences between the parametric and nonparametric estimates. The nonparametric estimate placed more mass around -5. The results for the Nikkei 225 index showed a substantial mode around 2.5 and a smaller mode around -7.5 which was closer to the results for the FTSE 100 than the S&P 500 and NASDAQ 100. This suggested that a lot of the dynamics in these indices were caused by relatively short-lived effects which last on the order of 2 weeks.

Figure 7 shows point-wise posterior estimates of σ_n for the three indices. The results for the FTSE 100 were similar to results for the S&P 500 index. However, the results for NASDAQ 100 were rather different to the S&P 500, with the period 1998 particularly different with volatility growing to about 5% in 2001. Interestingly, jumps in volatility tended to occur at similar times but the effects seem to decay at a slower rate. This was consistent with the larger mass around -5 and the smaller mass around -2.5 in the posterior mean of F_λ compared with the S&P 500 index.

Table 2 shows the posterior mean of the mass placed on the different frequency categories in the posterior mean of F_λ with the four stock indices. The NASDAQ 100 and Nikkei 225 placed much more mass in the medium category and much less mass in the low-frequency category than the FTSE 100 and S&P 500 indices.

The posterior mean volatilities assigned to the medium-frequency category are shown in Figure 8 with the same posterior mean for the S&P 500 index plotted as a dotted line for comparison (the posterior means for three indices for

Table 2 The posterior mean mass that F_λ places on the different categories for the four indices.

Index	Low	Medium	High
S&P 500	0.47	0.32	0.21
NASDAQ 100	0.27	0.55	0.18
FTSE 100	0.41	0.42	0.17
Nikkei 225	0.29	0.54	0.16

all categories are shown in the Appendix A). Comparing the estimates from the NASDAQ 100 and the S&P 500 indices shows that jumps in volatility occurred at the same times (in response to the same economic factors). However, the NASDAQ 100 had a much higher overall volatility and different estimated jump sizes. Comparing the six large jumps identified for the S&P 500 index (which are shown as events (a)–(f)) leads to the following observations. The total effect on volatility around event (a) was to double volatility (from 0.5 to 1.0 for the S&P 500 index and from 0.7 to 1.3 for the NASDAQ 100) over a period of a month (which includes several smaller jumps). Around event (b), volatility jumps from 0.4 to 1.1 for the S&P 500 from 1.1 to 2.3 for the NASDAQ 100. The absolute effect of the Russian default and the Asian crisis was larger for the NASDAQ 100 but the relative effect was smaller (with the volatility of the S&P 500 almost trebling whereas the volatility of the NASDAQ 100 only doubled). The effects of events (c) and (d) were much larger for the NASDAQ 100 than the S&P 500. The volatility of the NASDAQ 100 jumped from 1.5 to 3.4 whereas the volatility of the S&P 500 jumped from 0.5 to 0.9 around event (c). The volatility jumped from 2.6 to 4.0 whereas the volatility of the S&P 500 jumped from 0.6 to 1.1 around event (d). Conversely, the change

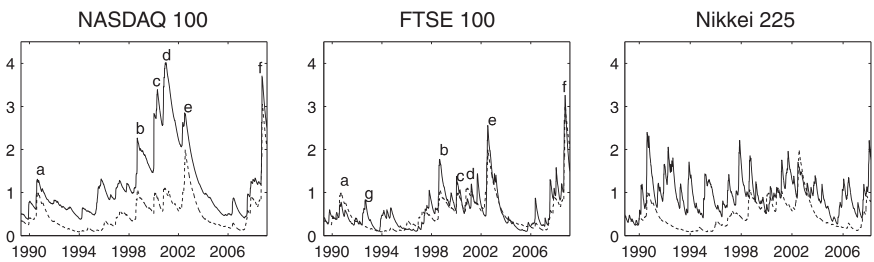

Figure 8 The posterior mean volatility in the medium-frequency category for the NASDAQ 100, FTSE 100, and Nikkei 225 (solid line) with the posterior mean volatility in the medium-frequency category for the S&P 500 index (dashed line).

in volatility around event (e) showed a change in the volatility of S&P 500 from 0.6 to 2.0 whereas the volatility of the NASDAQ 100 jumped from 2.1 to 2.9. Finally, around event (f), the volatility of the NASDAQ 100 jumped from 1.2 to 3.7 and the volatility of the S&P 500 jumped from 0.7 to 3.0.

In contrast to the NASDAQ 100, the volatility of the FTSE 100 followed the volatility of the S&P 500 much more closely with a similar overall level of volatility. However, there were some noticeable differences in the period 1992 to 1996 where the volatility of the S&P 500 consistently fell, unlike the FTSE 100 which showed two periods of increased volatility (which are discounted to similar levels as the S&P 500). In particular, there is an extra large jump which is marked g and is associated with "Black Wednesday" when Britain withdrew from the European Exchange Rate Mechanism. Clearly, this event would affect the British economy but not the American economy.

5 DISCUSSION

This paper proposes a Bayesian nonparametric approach to estimating a general OU-type stochastic volatility model and develops computational methods necessary to fit it. The volatility process is assumed to be a superposition of an infinite number of OU processes. This provides a flexible specification for the dynamics of the volatility process and avoids the need to prespecify the number of components included in the finite superposition model. The introduction of a nonparametric jump distribution allows flexible inference about the marginal distribution of the volatility. A novel MCMC approach is used to draw inference about the volatility process with a minimum of truncation. The analysis of four stock indices show that the mixing distribution of the superposition, F_λ, is multimodal (and often trimodal) if drawn on the log scale. This suggests decomposing volatility into subprocesses which are associated with each mode and so characterized by their dynamic behavior. These subprocesses work on very different timescales. One subprocess shows long-range dependence which represents a slow-moving volatility process, and a second subprocess shows a quickly decaying dependence which has a half-life of about 2 weeks. A third process, which has a medium-range dependence, is the most interesting and has jumps which can be associated with economic events.

The model could be developed by allowing other aspects of the model such as the risk-premium or leverage effect to differ between components. The extra flexibility would come at the cost of specifying an informative prior which makes the extension nontrivial. The model could also be applied to other financial time series. It would be particularly interesting to apply the model to high-frequency data.

A. ADDITIONAL RESULTS

Table 3 Posterior estimate of parameters for the NASDAQ 100 index.

	Inf-Sup Gamma OU	Inf-Sup NP OU
$E[\sigma_n^2]$	2.9 (2.0, 5.7)	4.9 (2.6, 9.4)
$SD[\sigma_n^2]$	3.7 (1.9, 9.5)	9.5 (3.2, 26.2)
$E[J]$	1.2 (0.9, 1.9)	2.0 (1.2, 2.9)
$E[\lambda]$	0.024 (0.011, 0.046)	0.028 (0.012, 0.049)
μ	0.086 (0.032, 0.139)	0.094 (0.042, 0.147)
β	−0.015 (−0.037, 0.001)	−0.017 (−0.039, 0.004)
ρ	−0.82 (−1.28, −0.48)	−0.64 (−1.08, −0.29)
M_λ	2.0 (0.7, 4.3)	2.4 (0.7, 5.1)
λ_J	0.010 (0.002, 0.037)	0.012 (0.002, 0.043)
σ_J^2	3.0 (1.6, 6.9)	3.0 (1.5, 6.9)

Table 4 Posterior estimate of parameters for the FTSE 100 index.

	Inf-Sup Gamma OU	Inf-Sup NP OU
$E[\sigma_n^2]$	1.6 (0.8, 4.0)	2.1 (1.2, 4.6)
$SD[\sigma_n^2]$	1.5 (0.6, 6.1)	2.1 (0.9, 6.7)
$E[J]$	1.0 (0.6, 1.7)	1.0 (0.6, 1.8)
$E[\lambda]$	0.021 (0.006, 0.050)	0.024 (0.009, 0.050)
μ	0.047 (0.001, 0.085)	0.048 (0.012, 0.085)
β	−0.031 (−0.072, 0.009)	−0.031 (−0.073, 0.009)
ρ	−0.78 (−1.26, −0.39)	−0.84 (−1.40, −0.46)
M_λ	2.0 (0.7, 4.6)	2.4 (0.8, 5.4)
λ_J	0.0047 (0.0009, 0.0164)	0.0045 (0.0009, 0.0146)
σ_J^2	3.7 (1.9, 8.9)	3.8 (1.9, 9.2)

Table 5 Posterior estimate of parameters for the Nikkei 225 index.

	Inf-Sup Gamma OU	Inf-Sup NP OU
$E[\sigma_n^2]$	2.6 (1.8, 5.4)	3.9 (2.3, 7.8)
$SD[\sigma_n^2]$	3.6 (1.8, 9.6)	7.0 (2.4, 19.4)
$E[J]$	1.3 (0.9, 2.1)	1.8 (1.0, 2.8)
$E[\lambda]$	0.023 (0.011, 0.041)	0.024 (0.013, 0.038)
μ	0.052 (0.001, 0.101)	0.053 (0.002, 0.103)
β	−0.039 (−0.071, −0.008)	−0.040 (−0.071, −0.008)
ρ	−0.48 (−0.92, −0.30)	−0.55 (−0.88, −0.26)
M_λ	2.15 (0.42, 5.32)	2.72 (1.10, 5.33)
λ_J	0.018 (0.006, 0.043)	0.018 (0.007, 0.041)
σ_J^2	4.6 (2.3, 10.4)	4.6 (2.3, 10.1)

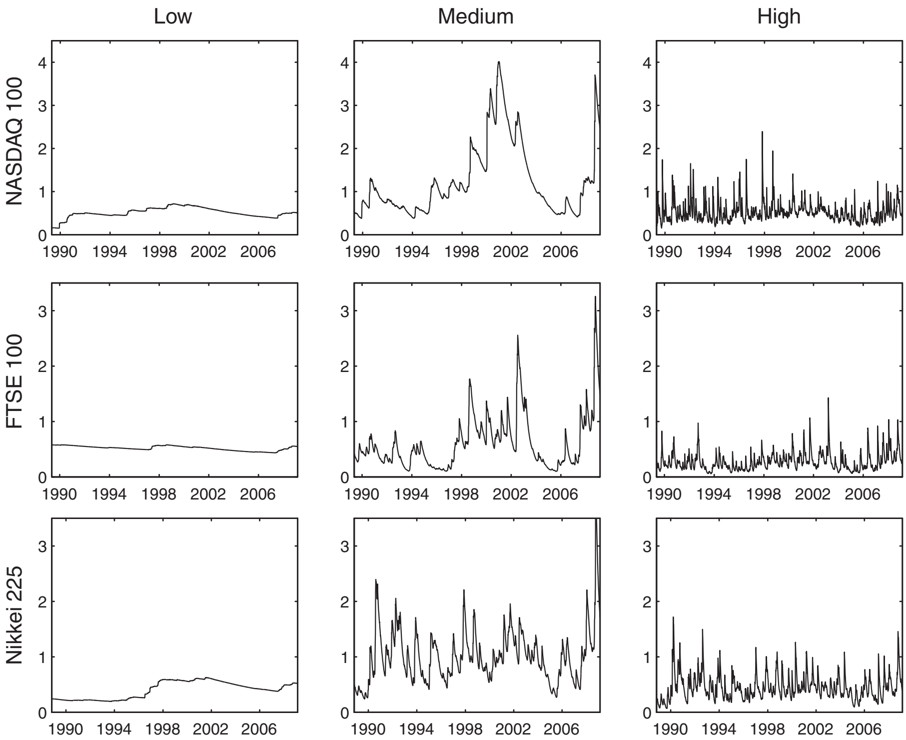

Figure 9 The posterior mean volatility in the three frequency categories for the NASDAQ 100, FTSE 100, and Nikkei 225 using the Inf-Sup NP OU model.

Received March 24, 2010; revised June 9, 2010; accepted June 14, 2010.

REFERENCES

Alizadeh, S., M. Brandt, and F. Diebold. 2002. Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance* 57: 1047–1091.

Barndorff-Nielsen, O. E. 2001. Superposition of Ornstein–Uhlenbeck Type Processes. *Theory of Probability and its Applications* 45: 175–194.

Barndorff-Nielsen, O. E., and N. Leonenko. 2005. Spectral Properties of Superpositions of Ornstein–Uhlenbeck Type Processes. *Methodology and Computing in Applied Probability* 7: 335–352.

Barndorff-Nielsen, O. E., and N. Shephard. 2001. Non-Gaussian OU Based Models and Some of Their Uses in Financial Economics (with Discussion). *Journal of the Royal Statistical Society B* 63: 167–241.

Barndorff-Nielsen, O. E., and N. Shephard. 2002. Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society B* 64: 253–280.

Brix, A. 1999. Generalized Gamma Measures and Shot-Noise Cox Processes. *Advances in Applied Probability* 31: 929–953.

Chernov, M., R. Gallant, E. Ghysels, and G. Tauchen. 2003. Alternative Models for Stock Price Dynamics. *Journal of Econometrics* 116: 225–257.

Cont, R., and P. Tankov. 2003. *Financial Modelling with Jump Processes*. Boca Raton: Chapman & Hall.

Eraker, B., M. Johannes, and N. Polson. 2003. The Impact of Jumps in Volatilities and Returns. *Journal of Finance* 58: 1269–1300.

Escobar, M. D., and M. West. 1995. Bayesian Density Estimation and Inference Using Mixtures. *Journal of the American Statistical Association* 90: 577–588.

Ferguson, T. S. 1973. A Bayesian Analysis of Some Nonparametric Problems. *Annals of Statistics* 1: 209–230.

Ferguson, T. S. 1974. Prior Distributions on Spaces of Probability Measures. *Annals of Statistics* 2: 615–629.

Ferguson, T. S., and M. J. Klass. 1972. A Representation of Independent Increment Processes without Gaussian Components. *Annals of Mathematical Statistics* 43: 1634–43.

Frühwirth-Schnatter, S., and L. Sögner. 2009. Bayesian Estimation of Stochastic Volatility Models based on OU processes with Marginal Gamma Law. *The Annals of the Institute of Statistical Mathematics* 61: 159–179.

Gander, M. P. S., and D. A. Stephens. 2007a. Simulation and Inference for Stochastic Volatility Models Driven by Lévy Processes. *Biometrika* 94: 627–646.

Gander, M. P. S., and D. A. Stephens. 2007b. Stochastic Volatility Modelling with General Marginal Distributions: Inference, Prediction and Model Selection. *Journal of Statistical Planning and Inference* 137: 3068–3081.

Griffin, J. E., and M. F. J. Steel. 2006. Inference with Non-Gaussian Ornstein–Uhlenbeck Processes for Stochastic Volatility. *Journal of Econometrics* 134: 605–644.

Griffin, J. E., and M. F. J. Steel. 2010. Bayesian Inference with Stochastic Volatility Models Using Continuous Superpositions of Non-Gaussian Ornstein–Uhlenbeck processes. *Computational Statistics and Data Analysis* (in press).

Hanson, T. 2006. Inference for Mixtures of Finite Polya Tree Models. *Journal of the American Statistical Association* 101: 1548–1565.

James, L. F., A. Lijoi, and I. Prünster. 2009. Posterior Analysis for Normalized Random Measures with Independent Increments. *Scandinavian Journal of Statistics* 36: 76–97.

Jensen, M. J., and J. M. Maheu. 2010. Bayesian Semiparametric Stochastic Volatility Modeling. *Journal of Econometrics* (in press).

Jongbloed, G., F. H. Van Der Meulen, and A. W. Van Der Vaart. 2005. Nonparametric Inference for Lévy-Driven Ornstein–Uhlenbeck Processes. *Bernoulli* 11: 759–791.

Kalli, M., S. G. Walker, and P. Damien. 2009. Bayesian Semiparametric Density Estimation Models with Applications to Financial Asset Returns. Technical report, University of Kent.

Lavine, M. 1992. Some Aspects of Polya Tree Distributions for Statistical Modeling. *The Annals of Statistics* 20: 1222–1235.

Lavine, M. 1994. More Aspects of Polya Tree Distributions for Statistical Modeling. *The Annals of Statistics* 22: 1161–1176.

LeBaron, B. 2001. Stochastic Volatility as a Simple Generator of Apparent Financial Power Laws and Long Memory. *Quantitative Finance* 1: 621–631.

Lijoi, A., R. H. Mena, and I. Prünster. 2007. Controlling the Reinforcement in Bayesian Nonparametric Mixture Models. *Journal of the Royal Statistical Society Series B* 69: 715–740.

Maudlin, R. D., W. D. Sudderth, and S. C. Williams. 1992. Polya Trees and Random Distributions. *The Annals of Statistics* 20: 1203–1221.

Molina, G., C.-H. Han, and J.-P. Fouque. 2008. "MCMC Estimation of Multiscale Stochastic Volatility Models." In C. F. Lee and A. C. Lee (Eds.), *Handbook of Quantitative Finance and Risk Management*. Heidelber: Springer.

Müller, P., and F. A. Quintana. 2004. Nonparametric Bayesian Data Analysis. *Statistical Science* 19: 95–110.

Newton, M. A., and A. E. Raftery. 1994. Approximate Bayesian Inference by Weighted Likelihood Bootstrap (with Discussion). *Journal of the Royal Statistical Society B* 56: 3–48.

Regazzini, E., A. Guglielmi, and G. Di Nunno. 2002. Theory and Numerical Analysis for Exact Distributions of Functionals of a Dirichlet Process. *The Annals of Statistics* 5: 1376–1411.

Regazzini, E., A. Lijoi, and I. Prünster. 2003. Distributional Results for Means of Normalized Random Measures with Independent Increments. *The Annals of Statistis* 31: 560–585.

Roberts, G. O., O. Papaspiliopoulos, and P. Dellaportas. 2004. Bayesian Inference for Non-Gaussian Ornstein–Uhlenbeck Stochastic Volatility Processes. *Journal of the Royal Statistical Society B* 66: 369–393.

Taufer, E., N. Leonenko, and M. Bee. 2009. Characteristic Function Estimation of Ornstein–Uhlenbeck-Based Stochastic Volatility Models. Technical report, Università Degli Studi di Trento.

Valdivieso, L., W. Schoutens, and F. Tuerlinckx. 2009. Maximum Likelihood Estimation in Processes of Ornstein–Uhlenbeck Type. *Statistical Inference for Stochastic Processes* 12: 1–19.

Woerner, J. H. C. 2007. Inference in Lévy-Type Stochastic Volatility Models. *Advances in Applied Probability* 39: 531–549.