School of Social Sciences
CITY UNIVERSITY LONDON

Department of Economics

Jump Diffusion Models: Estimation of Fit and Predictive Power

Laura Delaney[1]
City University London

Department of Economics
Discussion Paper Series
No. 11/11

[1] Department of Economics, City University London, Social Sciences Bldg, Northampton Square, London EC1V 0HB, UK.
Email: laura.delaney.1@city.ac.uk. T: +44 (0)20 7040 4129 F: +44 (0)20 7040 8580.

Jump Diffusion Models: Estimation of Fit and Predictive Power

Laura Delaney *

December 1, 2011

Abstract

In this paper I extend the familiar geometric Brownian motion (GBM) and Merton's log-normal jump diffusion (LJD) models of option pricing to a specification consisting of a Gamma distribution and a Beta distribution (GBJD). This specification separates the "good" news from the "bad" news components, which adds to the realism behind the intuition of the two other specifications. I compare the fit and, in particular, the predictive power of the three models and find that in terms of both, the GBJD is the preferred model over the GBM and the LJD.

Keywords: Jump Diffusions, Gamma Distribution, Gaussian Distribution, Log-Normal Distribution.

Mathematics Subject Classification (2000): 60G15, 60G51, 62P20.

JEL Classification: C46, C52, G13.

*Department of Economics, City University London, Northampton Square, London EC1V 0HB, UK. Email: Laura.Delaney.1@city.ac.uk, Tel. +44 (0)207 040 4129

1 Introduction

Most of the models and option pricing techniques employed in applied areas of finance are rooted in the well known Black-Scholes model. The classical model used for stock prices or indices, and which became the basis for the Black-Scholes option pricing theory is the geometric Brownian motion (GBM hereafter). However, almost as universally accepted as the Black-Scholes model itself, are its weaknesses. For instance, GBM is based on the predominant assumption that observations follow a Gaussian (Normal) distribution. This is largely due to the fact that the normal distribution, as well as the continuous-time process it generates, has nice analytic properties. However, a look at data from various areas of finance such as equity, fixed income, foreign exchange or credit, clearly reveals that by assuming normality, one gets a model which is only a poor approximation of reality.

One of the central assumptions underlying a sequence of events that lead to a Gaussian distribution is that there are no "wild" jumps or uncertainty as to step size. However, recent events have shown that stock prices have been subject to abrupt and unanticipated large changes or "jumps", and have become highly unstable and volatile in nature. Black-Scholes based models, such as GBM, fall apart in environments with rapid movements in the underlying assets. This is because the key distributional assumption is that the price of an asset follows a diffusion, that is, a stochastic process that generates a continuous sample path. In fact, this assumption implies normality, so that over a short interval of time, the stock price cannot change by much. The recent stock market volatility, during the 2007/08 Credit-Liquidity Crisis, and the example cited above of October 1987, provide evidence that diffusions inadequately characterize asset price movements and that processes allowing for jumps would be more appropriate.

From a practical viewpoint, financial decision making using models which

are based on a continuous-time setting will be satisfactory only if reasonable specifications of the stock price process are built upon. As well as this, the extent of skewness and the presence of outliers in an actual return distribution are important inputs to hedging and risk management decisions, as well as for option pricing. However, in a Black-Scholes based framework, such inputs are unaccounted for.

Although the GBM had served as a convenient and tractable framework, as the empirical evidence against GBM accumulated (see Sundaresan [2000]), Merton's jump diffusion (LJD) representation gained widespread acceptance, primarily because it was shown to be more consistent with the empirical returns distributions; that is, it produced a higher peak and accounted for excess kurtosis and skewness.

Recent work has stressed the importance of continuing to search for models based on processes that admit jumps and thereby providing a more accurate fit to the observed data. Concerted efforts are also being made to ensure such models are meaningful and mathematically plausible. In extant literature, a wide range of continuous-time models have been constructed by choosing different theoretical structures for the drift, the diffusion and the jump component of the process (for example, stochastic volatility and mean reversion). Also many variations have been proposed to enhance the jump specification by, for example, including different distributional assumptions for the jump magnitudes. Andersen and Andreasen [2000] proposed a model in which a non-random volatility structure is combined with log-normally distributed Poisson jumps. Kou [2002] proposes the Double Exponential Jump Diffusion (DEJD) model whereby a single Poisson process with fixed intensity generates the jumps in prices, but the jump amplitudes are drawn from two independent exponential distributions. The point of such models is to improve derivative pricing and portfolio optimization. These extensions are intended to reduce the deviation between model and reality. Adding a jump component should improve the fit

to the observed time series of returns, since the jumps may accommodate outliers as well as asymmetry in the return distribution. The presence of outliers depends upon the magnitude and variability of the jump component, while the asymmetry is controlled by the average magnitude of the jump.

While much of the literature has focussed primarily on improving the fit of these specifications with the data, to my knowledge, there has been very little effort, if indeed any, to assess their predictive power. The aim of this research is to go beyond the process introduced by Merton [1976] and, through a selection analysis, by assuming a different distributional choice to model the jump component, in particular a Gamma Distribution for the "good" news component and a Beta Distribution for the "bad" news component, whether the fit of the model will be enhanced, and/or more importantly, whether its forecasting accuracy will be improved.

There are several economic justifications for distinguishing between so called "good" and "bad" news. One such justification is provided by Milgrom [1981], who has formalized the notion of good and bad news and shown that such a distinction plays an important role for models used in information economics. Because information economics is the study of situations in which different economic agents have access to different information, signalling models, such as that of Spence [1973], are typically applied to deal with such information asymmetries. In such signalling models, the analysis is driven by a monotonicity property; for example, more talented workers buy more education. In order to incorporate the important role of monotonicity in models of rational expectations, which form the basis of information economics, a model which makes the distinction between good and bad news is necessary to obtain reliable results; specifically, in a rational expectations model, the rise in a firm's stock price should be attributable to the arrival of good news about the firm's prospects, and vice versa.

I use daily returns data for the ISEQ, FTSE100 and S&P500 indices and

also from a number of stocks included in one or other of these indices, in conjunction with cumulant moment matching to fit the models. The data source is Bloomberg. I will utilize the Schwartz Bayesian Information Criterion (BIC) to assess the relative performance of the models as outlined in Ramezani and Zeng [2007].[1] In order to assess the predictive performance, I calculate the Root Mean Square Error (RMSE) and the Janus Coefficient (J). In terms of both criteria, I find that the GBJD specification is the preferred model over the GBM and the LJD for the series of returns that I examine.

In the next section the standard GBM model and Merton's LJD model are discussed. In Section 3 the motivation of the so called Gamma-Beta specification (GBJD) is explained. Sections 4 through to 6 describe the model estimation, the selection process and the assessment of its predictive power in some detail. Sections 7 and 8 examine the data and results obtained. Section 9 finally concludes.

2 The Log-Normal Jump Diffusion Model

In order to establish a benchmark, I initially estimate a representation that is compatible with the Black-Scholes model. The classical model of GBM is given by the Stochastic Differential Equation (SDE):

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t \tag{1}$$

which represents the price process of the stock with price S_t, at time t, and $S_{t-} := \lim_{dt \downarrow 0} S_t dt$. All processes are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with filtration $(\mathscr{F}_t)_{t \geq 0}$.

The two terms are familiar from the Black-Scholes model where the drift rate (or instantaneous expected return on the stock) is denoted by μ, volatility

[1] The more commonly used Akaike criterion (AIC) for model selection tends to over-parameterize the models (see Schwartz [1978]), hence I opt to utilize BIC.

σ, conditional on no arrivals of important new information (both μ and σ are assumed to be constant), and the random walk (Wiener process) by W_t.

As a first extension of the GBM representation, I estimate Merton's [1976] model. His papers suggest that asset price dynamics may be modeled as jump-diffusion processes and that they provide the foundations under which to value contingent claims. In particular, he asserted that the true process of the underlying asset is a combination of a log-normal process and a jump process. These JD class of representations of an asset's returns process may be decomposed into three building blocks; a linear drift, a Brownian motion (diffusion component) representing "normal" price variations due to, for example, changes in the economic outlook or other new information that causes marginal changes in the asset's value, and a compound Poisson process (jump component) that generates "news" arrivals leading to "abnormal" (or more than marginal) changes in prices. It is assumed that this important new information arrives only at discrete points in time, and it is reasonable to expect that the volatile and more "abnormal" periods are random.

Upon arrival of news, jump magnitudes are determined by sampling from an independent and identically distributed (i.i.d.) random variable. Merton's special case has become the most important representation of the JD process. In the jump diffusion model posited by Merton, the GBM model (or Black-Scholes specification) is extended to accommodate a jump component, thus representing the process like so:

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t + \sum_{j=1}^{J} Y_{j,t} dN_{j,t}. \tag{2}$$

The last term represents the jumps, with J being the number of Poisson processes denoted by $N_{j,t}$, and the number of stochastic or deterministic jump amplitudes are denoted by $Y_{j,t}$. Hence, there is an instantaneous jump in the relative stock price of size $Y_{j,t}$ conditional on an increment in $N_{j,t}$. $N_{j,t}$ has constant jump intensity λ_j for $j = 1 \dots J$; that is, λ_j is the mean number of

arrivals of important new information per unit time and $\mathbb{P}(dN_{j,t} = 1) = \lambda_j dt$. The jump amplitude $Y_{j,t}$ may follow any distribution, but Merton assumes an i.i.d. log-normal (α, δ^2) distribution (that is, $\log(1 + Y_t) \backsim N(\alpha, \delta^2)$) and Poisson (λ) arrival process. Hence, the model is more commonly referred to as the log-normal jump diffusion (LJD) model.[2] It is further assumed that all processes are independent and that $Y_{j,t} > -1$ for all j, which ensures that stock prices are non-negative.

The presence of a jump component provides additional flexibility in capturing the salient features of equity returns, including skewness and leptokurtosis.

By applying Ito's formula (for semi-martingales) to $\partial \ln S_t$ and, through integration and the Fundamental Theorem of Calculus, a solution for (2) is obtained and given by[3]:

$$S_t = S_0 \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\} \prod_{0<s<t} \prod_{j=1}^{J} (1 + Y_{j,s} dN_{j,s}). \tag{3}$$

The density function for the log-return, $r_{i+1} = \log(S_{i+1}) - \log(S_i)$, is:

$$p(x; \Theta) = \sum_{j=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^j}{j!} \varphi(x; (\mu - \frac{1}{2}\sigma^2)t + j\alpha, \sigma^2 t + j\delta^2), \tag{4}$$

where $\varphi(x; \mu, \sigma^2)$ is the density for a normally distributed random variable with mean μ and variance σ^2 and t is the sampling frequency. The log-likelihood function can be written as:

$$l(x_1 \dots x_T; \Theta) = \sum_{i=1}^{T} \log p(x_i; \Theta). \tag{5}$$

As a second extension to the GBM, I assume a different modeling structure for the jump components and refer to this new formation as the Gamma-Beta jump diffusion (GBJD).

[2] Another strong assumption of the model is that the jump component of the underlying stock's return represents non-systematic or diversifiable risk, and therefore, all of the stock's systematic or non-diversifiable risk is contained within the continuous component.

[3] Alternatively, an explicit solution is provided by the Doleans-Dade formula outlined in Protter [1991].

3 Gamma-Beta Model

While Merton's LJD model is an improvement on the Black-Scholes based GBM specification, it too is based on underlying assumptions which are clear deviations from reality. For example, the model makes no distinction between the intensity or distributional characteristics of news that cause an upward jump in prices, that is, "good news", and news that cause downward jumps, or "bad news". It simply has a single jump component that captures the impact of news on security prices.

This limitation of Merton's LJD model has led me to propose another specification of the jump diffusion model in which the upward and downward jumps are distinguished in terms of their distributional characteristics. I model the upward jump amplitudes via the Gamma distribution and the downward jumps via a Beta distribution and investigate if this choice of distributions will improve the fit or, more importantly, the forecasting accuracy over the other models.

The separation of good from bad news implies that a constraint must be imposed upon the range of values for the random return series (or equivalently, the percentage change in price). Firstly, the returns due to bad news must be bounded on the downside by -100% because stocks represent limited liability. Secondly, the returns due to good news must be positive. These constraints imply that we cannot model either the "up" or "down" jump amplitudes by assuming a log-normal distribution. They also indicate which distributions may be plausibly utilized when modeling upward and downward movements. The Gamma distribution requires that the underlying random variables are necessarily positive, and thus, I model the jump amplitudes for good news by a Gamma$(1, \alpha_u^{-1})$ distribution, and assume a Beta$(1, \alpha_d)$ distribution to model those amplitudes that are due to bad news. As with many other statistical distributions, the Beta distribution does not require the underlying random

variables to be necessarily positive or negative. However, I opt to use this distribution to model the downward jump amplitudes because of its compatibility with the Gamma distribution for modeling purposes.

Equation (2) then becomes:

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t + \sum_{j=1}^{J_u} Y_{j_u,t} dN_{j_u,t} + \sum_{j=1}^{J_d} Y_{j_d,t} dN_{j_d,t}. \tag{6}$$

where the third term is a summation over the upward jumps (see subscripts u) and the fourth over the downward movements (see subscripts d). The solution of which is given by:

$$S_t = S_0 \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\} \prod_{0<s<t} \prod_{j=u,d}^{G,B} (1 + Y_{j,s} dN_{j,s}). \tag{7}$$

It should be noted that G and B relate to the number of observations with up movements and down movements respectively.

From this model, the parameters $\Theta = (\mu, \sigma, \lambda_u, \lambda_d, \alpha_u, \alpha_d)$ must be estimated.

In essence, the Gamma-Beta model reduces to a single JD model, such as Merton's, when $\lambda = \lambda_u + \lambda_d$ and has a jump amplitude which takes the form of a mixed distribution of Gamma(1, α_u^{-1}), with probability $\frac{\lambda_u}{\lambda}$, and Beta(1, α_d) with probability $\frac{\lambda_d}{\lambda}$.

4 Estimation

One method of obtaining parameter estimates for the model is to match cumulants; that is, to equate the first six population cumulants with the first six sample cumulants, since the model has six unknown parameters that need to be estimated $(\mu, \sigma, \lambda_u, \lambda_d, \alpha_u, \alpha_d)$. Then the resulting equation for the parameters must be solved. I follow the procedures outlined in Kendall and Stuart [1977].

I used STATA10 to calculate the first six sample cumulants, $\bar{M}_s, s = 1 \dots 6$ from the sample moments, $\bar{m}_s = \mathbb{E}[X^s]$, in each of the data sets via the equation $\bar{M}_s = \mathbb{E}[(X - \mathbb{E}[X])^s] = \mathbb{E}[(X - m_1)^s]$, $s = 1, \dots 6$.

I define $U := \sum_{i=1}^{G_s} \ln(Y_i^u)$ and $D := \sum_{i=1}^{B_s} \ln(Y_i^d)$ for the up and down jumps respectively. These summations are compound Poisson processes, where G_s and B_s have intensities λ_u and λ_d respectively. The Moment Generating Function (MGF) of these processes is given by $\phi_{\ln(Y_i^j)}(s) = \exp \lambda(\phi_{\ln(Y_i^j)}(s) - 1)$, $j = u, d$. Thus, the Cumulant Generating Function (CGF), given by the logarithm of the MGF, is $\kappa_J(s) = \lambda_j(\phi_{\ln(Y_i^j)}(s) - 1) = \lambda_j \sum_{i=1}^{\infty} \frac{\mathbb{E}[(\ln(Y_i^j))^i]}{i!}$, $J = U, D; j = u, d$.

From this information, equations for the first six population moments, $M_p, p = 1 \dots 6$ can be derived:

$$M_1 = s(\mu - \frac{1}{2}\sigma^2 + \frac{\lambda_u}{\alpha_u} - \frac{\lambda_d}{\alpha_d}). \tag{8}$$

$$M_2 = s(\sigma^2 + 2\frac{\lambda_u}{\alpha_u^2} + 2\frac{\lambda_d}{\alpha_d^2}). \tag{9}$$

$$M_3 = s(6\frac{\lambda_u}{\alpha_u^3} - 6\frac{\lambda_d}{\alpha_d^3}). \tag{10}$$

$$M_4 = s(24\frac{\lambda_u}{\alpha_u^4} + 24\frac{\lambda_d}{\alpha_d^4}). \tag{11}$$

$$M_5 = s(120\frac{\lambda_u}{\alpha_u^5} - 120\frac{\lambda_d}{\alpha_d^5}). \tag{12}$$

$$M_6 = s(720\frac{\lambda_u}{\alpha_u^6} + 720\frac{\lambda_d}{\alpha_d^6}). \tag{13}$$

Then by setting $M_p = \bar{M}_s$, I obtain the six estimates, which can be easily solved using simultaneous equations and a program such as Mathematica. Some simple algebraic manipulation produces equations which are easier to handle numerically. One such equation being:

$$\left(\frac{\bar{M}_5^2}{100} - \frac{\bar{M}_4\bar{M}_6}{120}\right)\hat{\alpha_u^2} + \left(\frac{\bar{M}_4\bar{M}_5}{20} + \frac{\bar{M}_3\bar{M}_6}{30}\right)\hat{\alpha_u} + \left(\frac{\bar{M}_4^2}{4} - \frac{\bar{M}_3\bar{M}_5}{5}\right) = 0 \tag{14}$$

Solving for the positive root of this equation yields an estimate of α_u, denoted $\hat{\alpha}_u$, and then by substitution into the remaining equations, the following expressions are obtained:

$$\frac{5\bar{M}_4\hat{\alpha_u} - 20\bar{M}_3}{-5\bar{M}_5\hat{\alpha_d} + 5\bar{M}_4} = \hat{\alpha_d} \tag{15}$$

$$\frac{\hat{\alpha_d}^4(\frac{\bar{M}_4}{24s}\hat{\alpha_u} - \frac{\bar{M}_3}{6s})}{\hat{\alpha_u} + \hat{\alpha_d}} = \hat{\lambda_d} \tag{16}$$

$$\hat{\alpha_u^3}\left(\frac{\bar{M}_3}{6} + \frac{\hat{\alpha_d}(\frac{\bar{M}_4}{24s}\hat{\alpha_u} - \frac{\bar{M}_3}{6s})}{\hat{\alpha_u} + \hat{\alpha_d}}\right) = \hat{\lambda_u} \tag{17}$$

$$\frac{\bar{M}_2}{s} - 2\frac{\hat{\lambda_u}}{\hat{\alpha_u}^2} - 2\frac{\hat{\lambda_d}}{\hat{\alpha_d}^2} = \hat{\sigma}^2 \tag{18}$$

$$\frac{\bar{M}_1}{s} + \frac{1}{2}\hat{\sigma}^2 - \frac{\hat{\lambda_u}}{\hat{\alpha_u}} + \frac{\hat{\lambda_d}}{\hat{\alpha_d}} = \hat{\mu} \tag{19}$$

This method was used by Beckers [1981] to obtain parameter estimates for the LJD model. The results for the Gamma-Beta model are presented in Table 5 in the Appendix.

However, while the estimates that cumulant matching yields are consistent, it should be noted that they are inefficient. As well as this, as Press [1968] discusses, the cumulants are functions of the sample moments and therefore, the distributions of the cumulant estimators in large samples will be Normal. For this reason, it may be advisable to only use cumulant matching to obtain initial values for Maximum Likelihood Estimation (MLE). The MLE method can, theoretically, be used since the proposed Gamma-Beta model is a first order stochastic differential equation of generalized Ito type. With equally-spaced sample data, the log-likelihood function given N returns observations is:

$$L(r; \mu, \sigma, \lambda_u, \lambda_d, \alpha_u, \alpha_d) = \sum_{i=1}^{N} \ln(f(r_i)) \tag{20}$$

where the derivation of $f(r_i)$ is given in Appendix A.

However, the unconditional distribution of returns is a mixture density, that is, the sum of four conditional densities, each of which are assigned Poisson

weights, but for such densities, a global maximum of the likelihood function does not exist (see Kiefer [1978]). This is due to the fact that the likelihood function tends towards infinity at a point where for the ith observation, $r_i = \mu$ and $\sigma_i \to 0$. However, Hamilton [1994] (page 689) shows that such problems are not a major obstacle once the choice of numerical optimization procedure converges to a local maxima.

The Newton-Raphson procedure is the most widely used optimization method for jump-diffusion models. However, this method requires the calculation of first and second order derivatives of the log-likelihood function, which are difficult to obtain for the GBJD specification. While there are numerical optimization procedures which do not necessitate the use of derivatives, such as Powell's method, these too are not without their caveats.

The estimation of the GBJD, as well as the calculation of the standard errors via Powell's method, would be computationally extremely time consuming because the likelihood function involves double infinite summations and double improper integrals. A high-performance computer would also be needed to carry out the calculations.

Another challenge that is likely to arise during the estimation process is that the likelihood function may not converge, and hence a type of switching algorithm may need to be used, that is, to combine Powell's method with another algorithm, such as method of Steepest Descent, so that explosion of the likelihood function may be avoided.

Hence, I have chosen to stick with the estimates obtained via the Cumulant Matching method, and despite its shortcomings, the estimates still yield informative results.

5 Model Selection

For model selection I have utilized the Bayesian Information Criterion (BIC) proposed by Schwartz [1978]. The advantage of this method over likelihood tests is that the BIC allows the comparison of more than two model specifications simultaneously and it does not require that the alternatives be nested.

Suppose that the ith model M_i has a vector of parameters θ_i, where θ_i has n_i independent parameters to be estimated. $\hat{\theta}_i$ denotes the estimator of θ_i. Then the BIC for model M_i is defined by:

$$BIC_i = -2 \log f(D|\hat{\theta}_i; M_i) + n_i \log(T) \tag{21}$$

where T is the number of observations in the data set D and $f(D|\hat{\theta}_i; M_i)$ estimated function. The best "fit" model is the one with the smallest BIC. However, while fit is important in terms of model selection, the critical test of any economic or financial model is its ability to forecast future returns.

6 Predictive Power

The ultimate test of the quality of a fitted model is the accuracy of the forecasts of the conditional distribution of future observations given past observations on a variable. While stock returns are affected by the usual fluctuations in economic variables, a significant driver of returns is the innate human propensity to swing between euphoria and fear. While this behavior of market participants is heavily influenced by economic events, often extreme actions are not underpinned by any fundamental factor. Owing to this extent of unpredictable human behavior on determining stock returns, no such model is ever going to be able to accurately predict future observations. That is, a fitted model is at best only an estimate or approximation of the process that generates the underlying data set, and thus, the model and the forecasts that it produces are

subject to identification and estimation errors. It is also crucial to be aware that forecasting assumes that the data generating process remains stable into the future.

The forecast of returns that will have the minimum mean square forecast error, that is, the optimal predictor of returns, is the expected value of r_{T+l}, $l = 1, 2, \ldots$ conditional on the information available at time T i.e. $\hat{r}_{T+l} = \mathbb{E}(r_{T+l}|\mathcal{F}_T)$.

GBM Forecast Estimates

From equation (3) above, the log-return $r_{T+l} = \log S_{T+l} - \log S_{T+l-1}$ is given by:

$$r_{T+l} = \left(\mu - \frac{1}{2}\sigma^2\right) + \sigma\left[W_{T+l} - W_{T+l-1}\right], \tag{22}$$

and thus

$$\hat{r}_{T+l} = \mathbb{E}\left(r_{T+l}|\mathcal{F}_T\right) = \hat{\mu} - \frac{1}{2}\hat{\sigma}^2 \tag{23}$$

since $W_i, i = 1, 2, \ldots$ is a Brownian motion and therefore has expected value zero.

6.1 LJD Forecast Estimates

For the LJD specification, the log-return for the l-period ahead forecast is given by:

$$r_{T+l} = \left(\mu - \frac{1}{2}\sigma^2\right) + \sigma\left[W_{T+l} - W_{T+l-1}\right] + \sum_{j=1}^{J} \log\left(1 + Y_{j,T+l} dN_{j,T+l}\right) \tag{24}$$

and

$$\hat{r}_{T+l} = \mathbb{E}(r_{T+l}|\mathcal{F}_T) = \hat{\mu} - \frac{1}{2}\hat{\sigma}^2 + \hat{\lambda}\hat{\alpha}. \tag{25}$$

6.2 GBJD Forecast Estimates

Through similar reasoning,

$$r_{T+l} = \left(\mu - \frac{1}{2}\sigma^2\right) + \sigma\left[W_{T+l} - W_{T+l-1}\right] + \sum_{j=u}^{J_u} \log\left(1 + Y_{j,T+l} dN_{j,T+l}\right) + \sum_{j=d}^{J_d} \log\left(1 + Y_{j,T+l} dN_{j,T+l}\right) \tag{26}$$

and

$$\hat{r}_{T+l} = \mathbb{E}(r_{T+l}|\mathcal{F}_T) = \hat{\mu} - \frac{1}{2}\hat{\sigma}^2 + \frac{\hat{\lambda_u}}{\hat{\alpha_u}} - \frac{\hat{\lambda_d}}{\hat{\alpha_d}}. \tag{27}$$

6.3 Assessing Forecasting Accuracy

In order to assess the accuracy of the forecasts, two popular measures are employed; the Root Mean Square Error (RMSE) and the Janus Coefficient (J).[4] The Mean Absolute Error (MAE) is another popular measure of forecasting adequacy, but it does not penalize large forecasting errors as much as the RMSE measure does. Hence, I calculate only the RMSE measure and J. Small values of the RMSE and J indicate good forecasting performance. The values of r_i, $i = 1, 2, \ldots T$ are the returns in the original sample and for r_i, $i = T+1, \ldots T+l$, I used the return values starting from the day immediately following the last observation in the sample for each stock until July 6th, 2009.

7 Sample Statistics

Tables 1 and 2 below give some summary statistics for the sample taken of daily returns from the ISEQ, FTSE100 and S&P500 composite indices and a list of nine individual stocks. The stocks that I select are very liquid, an important characteristic given the event driven nature of the jump diffusion models. The indices are broad indicators of the equity market, and the daily

[4] $RMSE = \sqrt{\frac{1}{l}\sum_{i=T+1}^{T+l}(\hat{r}_i - r_i)^2}$ and $J = \frac{\frac{1}{l}\sum_{i=T+1}^{T+l}(\hat{r}_i - r_i)^2}{\frac{1}{T}\sum_{i=1}^{T}(\hat{r}_i - r_i)^2}$.

sampling frequency captures high-frequency fluctuations in the returns process that are critical for identification of jump components.[5] However, since none of the indices reported are adjusted for dividends, it is more correct to note that I model the observed series of log-price differences and refer to it hereafter as the "return process".

The ISEQ index spans the period 01/2000 through 12/2008 ($N = 2319$), the FTSE100 spans the period 10/1985 through 10/2008 ($N = 5849$) and the S&P500 series spans the period 12/1988 through 12/2008 ($N = 5042$). The number of days in the sample with positive returns (Up Freq.), no change in returns and negative returns (Down Freq.) appear in the last three columns of the table. For the indices there are a comparable number of days with positive and negative returns. However, there are significantly fewer days with no change in the returns in each instance. The range of returns ((−.3850, .3755) for the ISEQ, (−.1303, .8469) for the FTSE100 and (−.0947, .1096) for S&P500) is large for all indices. This is indicative of the occurrence of significant booms and busts during the sample periods. The excess kurtosis is induced by stochastic volatility in the returns process and exceeds what can be rationalized by the GBM model.

Indeed, the summary statistics for the stocks produce a very similar picture to those of the indices, and hence, do not require a separate discussion in this section.

8 Results

Tables 3, 4 and 5, reported in Appendix B, are estimates of the parameters for the GBM, Merton and Gamma-Beta models respectively. All of the estimated parameters are of daily sizes.

[5]The daily sampling frequency avoids modeling intra-day return dynamics which are confounded by market microstructure effects and trading frictions.

Table 1: Summary Statistics

	Mean	SD	Min	Max
FTSE100	.0002	.0111	−.1303	.0847
S&P500	.0002	.0112	−.0947	.1096
ISEQ	−.0002	.0174	−.3850	.3755
Intel	.0006	.0270	−.2489	.1833
Vodafone	.0004	.0232	−.1590	.1371
HSBC	.0004	.0185	−.1455	.1442
British American Tobacco	.0006	.0218	−.2202	.1872
General Electric	.0003	.0175	−.1368	.1276
Johnsons&Johnsons	.0005	.0151	−.1725	.1154
Bank of Ireland	−.0016	.0326	−.2725	.3443
Fyffes	−.0012	.0295	−.5766	.1335
Elan Corp.	.0005	.0600	−1.140	.2377

Table 2: Summary Statistics cont.

	Skewness	Kurtosis	Up Freq.	No Change	Down Freq.
FTSE100	.6918	13.82	3054	13	2782
S&P500	.2664	13.38	2675	3	2364
ISEQ	−.0877	19.74	1179	65	1073
Intel	−.3944	8.17	2512	120	2409
Vodafone	.0205	6.13	2443	231	2379
HSBC	−.1643	8.81	2080	118	1982
British American Tobacco	.3535	15.15	1295	69	1265
General Electric	−.1443	9.69	2426	195	2420
Johnsons&Johnsons	−.1856	9.66	2471	164	2406
Bank of Ireland	.6082	28.52	703	100	756
Fyffes	−.6489	11.27	584	359	616
Elan Corp.	−.6971	11.63	741	99	719

The second last column in each table reports the BIC values for the three specifications. As mentioned previously, the model with the smallest BIC value provides the best fit to the data. The values in boldface are the smallest values. As expected, the GBM specification does not provide a better fit over the jump diffusion specifications for any of the return series examined. This is consistent with intuition and expectations. The GBJD is superior to the LJD for eight of the returns series, with the LJD beating the GBJD for only four cases. This provides strong support for the GBJD specification in terms of model fit.

The last column in these tables reports the Janus Coefficient value, J. Similarly, the smallest values are in boldface. For brevity I do not to report the RMSE value, since these results yielded the same conclusions to those produced by J. From a forecasting adequacy perspective, the GBJD specification also appears to perform better than the other two specifications. For the three indices, the GBJD provides both a better fit and appears to have greater predictive power than the GBM and LJD models. For the individual stocks, the link between forecasting adequacy and fit is more opaque. For HSBC and Bank of Ireland, the LJD is strongest in terms of fit and predictive power and for General Electric, the GBJD seems to be the preferred model for both. For the rest of the stocks, the model which gives the best fit for a particular stock is not the same model that has the strongest predictive power for that stock. Indeed, the GBM appears to have the greatest forecasting ability for Intel and Vodafone and Elan Corp. Overall, the results for forecasting adequacy of the models are not very informative. However, the GBJD has a greater predictive power for five of the twelve returns series studied, whereas the GBM and LJD have greater predictive power for only three and four of the twelve series respectively.

In discussing the estimate results, I focus on the daily returns for the FTSE100 (first row in the tables). This is largely because the series spans the period which includes the market crisis of October 1987, the years 1997 and

1998 (Russian default and Long-Term Capital Management collapse), September 11th 2001 and some of the 2007/08 credit and liquidity crisis. That is, it incorporates the rare events that had a significant impact on stock prices and causes dramatic consequences on the market volatility. As well as this, the sample is big enough to encompass the periods where the market stabilized after these rare events (excepting the most recent crisis, however).

Considering firstly, the LJD parameters, it appears that a jump in the return process occurs approximately once every $3,300$ days (λ^{-1}). As expected, the estimated volatility associated with the continuous component of the jump diffusion specifications is smaller than the corresponding estimate under GBM. However, turning to the GBJD specification of the FTSE100, it appears that "good" and "bad" news arrive once every 625 (λ_u^{-1}) and 87 (λ_d^{-1}) days respectively, which is considerably more realistic than a jump occurring approximately once every 10 years, as the LJD specification suggests.

Overall, the parameter estimates obtained for the GBJD, the model of most interest, have reasonable values which are informative. One of the most notable things is that the volatility parameter for all of the returns series is significantly reduced by separating the jump components into up jumps and down jumps, as expected.

9 Concluding Remarks

This paper extends the standard geometric Brownian motion (GBM) and log-normal jump diffusion (LJD) models of option pricing to a new specification which distinguishes between upward and downward jumps in returns. This new specification, namely the Gamma-Beta jump diffusion (GBJD) model, is derived and it is compared with the other two specifications in terms of model fit and forecasting power. It is found that the separation of the jump component to distinguish between upward and downward movements clearly

improves the characterization of the empirical distribution of returns. In terms of both fit and predictive performance, the GBJD specification is the preferred model over the GBM and the LJD for the series of returns that are examined.

Theoretically, the GBJD specification can be applied to assess the dynamics of a wide number of other economic variables, not simply stock returns. Such variables include inflation, short-term interest rates and foreign exchange.

Jumps arise for many reasons, such as sudden financial turmoil, as witnessed globally in August 2007, litigation issues or incomplete accounting information. Hence, searching for models that account for such jumps, or "abnormal" movements in the underlying assets, are becoming increasingly important in terms of financial modeling.

References

[1] Andersen, L., Andreasen, J. "Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Pricing." *Review of Derivatives Research*, 4 (2000), pp. 231–262.

[2] Beckers, S. "A Note on Estimating the Parameters of the Jump Diffusion Model of Stock Returns." *Journal of Financial and Quantitative Analysis,* Vol. 16, No. 1 (1981), pp. 127–140.

[3] Hamilton, J.D. *Time Series Analysis,* Princeton University Press, Princeton, NJ., 1994.

[4] Kendall, M. G., Stuart, A. *The Advance Theory of Statistics,* London: Griffen and Co., 1977.

[5] Kiefer, N. M. "Discrete Parameter Variation: Efficient Estimation of a Switching Regression Model." *Econometrica,* Vol. 46, No. 2 (1978), pp. 427–434.

[6] Kou, S.G. "A Jump-Diffusion Model for Option Pricing." *Management Science,* Vol. 48, No. 8 (2002), pp. 1086–1101.

[7] Lewis, M. *The Story of Modern Financial Insanity,* London: Penguin, 2008.

[8] Merton, R. "The Impact on Option Pricing of Specification Error in the Underlying Stock Price Returns." *Journal of Finance,* Vol. 31, No. 2 (1976), pp. 333–350.

[9] Milgrom, P. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics,* 12 (1981), pp. 380–391.

[10] Press, S. J. "A Modified Compound Poisson Process with Normal Compounding." *Journal of American Statistical Association*, (1968), pp. 607–613.

[11] Protter, P. *Stochastic Integration and Differential Equations,* 2nd. ed., Berlin: New York: Springer, 2005.

[12] Ramezani, C.A., Zeng, Y. "Maximum Likelihood Estimation of the Double Exponential Jump Diffusion Process." *Annals of Finance,* 3 (2007), pp. 487–507.

[13] Schwartz, G. "Estimating the Dimension of a Model." *Annals of Statistics,* Vol. 6, No. 2 (1978), pp. 461–464.

[14] Spence, M. "Job Market Signalling". *The Quarterly Journal of Economics*, Vol. 87, No. 3 (1973), pp. 355–374.

[15] Sundaresan, S.M. "Continuous-Time Methods in Finance: A Review and an Assessment." *Journal of Finance,* Vol. 55, No. 4 (2000), pp. 1569–1622.

[16] Walck, C. Gamma Distribution. *Handbook on Statistical Distributions for Experimentalists*. Particle Physics Group, University of Stockholm, 1996.

Appendix

A Derivation of the Unconditional Density Function

Let $G_s = m$ be the number of upward jumps and $B_s = n$ the number of downward jumps over the time span $t = 0, \ldots, s$.

By letting $U := \sum_{i=1}^{G_s} \ln(Y_i^u) > 0$, $D := \sum_{i=1}^{B_s} \ln(Y_i^d) < 0$ and $T := U + D$, the s period return can be written as:

$$r(s) = (\mu - \frac{1}{2}\sigma^2)s + Z(s) + U + D. \tag{A.1}$$

As shown in Walck [1996], $\exp(\alpha) = \Gamma(1, \alpha^{-1})$. Thus, $\ln(Y^u) \sim \Gamma(1, \alpha_u^{-1}) = \exp(\alpha_u)$. Thus, for $G_s = m \geq 1$ the conditional distribution of U is given by $f_{U|m} \sim \Gamma(m, \alpha_u^{-1})$. Also, if $Y^d \sim Beta(1, \alpha_d)$, then $\ln(Y^d) \sim \exp(-\alpha_d) = \Gamma(1, -\alpha_d)$. Then for $B_s = n \geq 1$, the conditional distribution of $D < 0$ is denoted by $f_{-D|n} \sim \Gamma(n, -\alpha_d)$.

Therefore, if $m, n \geq 1$, the conditional density for $r(s)$ is an independent sum of $\Gamma(m, \alpha_u)$, $\Gamma(n, -\alpha_d)$ and $N((\mu - \frac{1}{2}\sigma^2)s, \sigma^2 s)$:

$$f_{r(s)|m,n}(r) = \int_{-\infty}^{\infty} f(r-t)\Big(\int_{-\infty}^{0} f_{-D}(x) f_U(t-x)dx\Big)dt, \tag{A.2}$$

where $f(r-t) \sim N((\mu - \frac{1}{2}\sigma^2)s, \sigma^2 s)$.

It is also necessary to derive the unconditional density of $s = 1$ period returns as it plays a crucial role in estimation and hypothesis testing. The function can be written as a Poisson weighted sum of the four conditional densities; i.e. $f(r)$ for $m = n = 0$, $f(r)$ for $m \geq 1, n = 0$, $f(r)$ for $m = 0, n \geq 1$ and $f(r)$ for $m, n \geq 1$. where $P(n, \lambda) = \frac{e^{-\lambda}\lambda^n}{n!}$.

$$f(r) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} P(m, \lambda_u)P(n, \lambda_d) f_{m,n}(r) \tag{A.3}$$

where $P(n, \lambda) := \frac{e^{-\lambda}\lambda^n}{n!}$.

B Tables

Table 3: GBM

	$\hat{\mu}$	$\hat{\sigma}$	BIC	J
FTSE100	.0003	.0111	−30219	0.62259
S&P500	.0003	.0112	−30946	0.69740
ISEQ	−.00001	.0174	−12177	0.15453
Intel	.0010	.0265	−21723	**0.32605**
Vodafone	.0007	.0228	−23216	**0.33777**
HSBC	.0005	.0185	−21492	1.30006
British American Tobacco	.0008	.0218	−12650	0.18880
General Electric	.0005	.0161	−26599	1.53640
Johnsons&Johnsons	.0006	.0147	−27493	0.36612
Bank of Ireland	−.0010	.0326	−6235	1.64332
Fyffes	−.0006	.0295	−6553	0.40311
Elan Corp.	.0023	.0600	−4335	**0.85912**

Table 4: LJD

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\delta}$	BIC	J
FTSE100	.0003	.0109	.0003	−.1362	.0456	−36395	0.62247
S&P500	.0003	.0119	.0000	−.2818	.1435	−31206	0.69738
ISEQ	.0001	.0046	.0172	−.0047	.1375	−12991	0.15444
Intel	.0011	.0260	.0004	−.2460	.1016	−21911	0.32606
Vodafone	.0007	.0228	.0000	1.386	.6112	−23351	0.33778
HSBC	.0005	.0184	.0001	−.3918	.2246	**-21564**	**1.30005**
British American Tobacco	.0000	.0217	.0000	1.3849	.4679	**-12759**	0.18881
General Electric	.0005	.0161	.0000	−.6793	.4911	−26636	1.53639
Johnsons&Johnsons	.0006	.0146	.0051	−.2309	.0913	−27629	**0.36602**
Bank of Ireland	.0002	.0281	.0062	.0261	.2071	**-8574**	**1.63810**
Fyffes	.0012	.0210	.0055	−.1814	.2111	−7583	**0.40203**
Elan Corp.	.0032	.0400	.0069	−.3456	.4123	**-6575**	0.85957

Table 5: GBJD

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda_u}$	$\hat{\alpha_u}$	$\hat{\lambda_d}$	$\hat{\alpha_d}$	BIC	J
FTSE100	.0005	.0103	.0016	32.24	.0115	38.08	**-36452**	**0.62244**
S&P500	.0006	.0078	.0646	68.01	.0903	69.57	**-35784**	**0.69737**
ISEQ	−0.0001	.0083	.0109	13.57	.0077	11.45	**-13551**	**0.15443**
Intel	.0016	.0247	.0032	18.64	.0164	20.84	**-21945**	0.36797
Vodafone	−.0012	.0162	.3316	67.08	.1785	56.91	**-23513**	0.33789
HSBC	.0008	.0144	.0732	49.76	.0834	47.64	−21511	1.30027
British American Tobacco	−.0008	.0142	.1150	37.11	.0517	31.43	−12741	**0.18879**
General Electric	.00003	.0146	.0443	55.91	.0093	31.55	**-26676**	**1.52794**
Johnsons&Johnsons	.0007	.0143	.0008	22.64	.0028	23.22	**-27800**	0.36629
Bank of Ireland	−.0047	.0170	.2412	32.49	.0424	16.37	−8295	1.63811
Fyffes	.0106	.0177	.00001	1.16	.2040	15.62	**-7835**	0.40328
Elan Corp.	.0172	.0305	.00003	.6232	.1432	8.10	−6142	0.86977