The hypothesis that asset returns are normally distributed has been widely rejected. The literature has shown that empirical asset returns are highly skewed and leptokurtic. The affine jump-diffusion (AJD) model improves upon the normal specification by adding a jump component to the price process. Two important extensions proposed by Ramezani and Zeng (1998) and Kou (2002) further improve the AJD specification by having two jump components in the price process, resulting in the asymmetric affine jump-diffusion (AAJD) specification. The AAJD specification allows the probability distribution of the returns to be asymmetrical. That is, the tails of the distribution are allowed to have different shapes and densities. The empirical literature on the "leverage effect" shows that the impact of innovations in prices on volatility is asymmetric: declines in stock prices are accompanied by larger increases in volatility than the reverse. The asymmetry in AAJD specification indirectly accounts for the leverage effect and is therefore more consistent with the empirical distributions of asset returns. As a result, the AAJD specification has been widely adopted in the portfolio choice, option pricing, and other branches of the literature. However, because of their complexity, empirical estimation of the AAJD models has received little attention to date. The primary objective of this paper is to contribute to the econometric methods for estimating the parameters of the AAJD models. Specifically, we develop a Bayesian estimation technique. We provide a comparison of the estimated parameters under the Bayesian and maximum likelihood estimation (MLE) methodologies using the S&P 500, the NASDAQ, and selected individual stocks. Focusing on the most...
recent spectacular market bust (2007–2009) and boom (2009–2010) periods, we examine how the parameter estimates differ under distinctly different economic conditions.

**Keywords:** Asset price processes; affine jump-diffusion; double exponential jump-diffusion; Markov chain Monte Carlo; Bayesian econometrics.

1. Introduction

Almost every aspect of modern finance theory, from valuations and portfolio choice to option pricing and corporate finance, as well as the ever expanding field of mathematical finance, critically depends upon the form of the probability distribution governing the asset returns. Although geometric brownian motion (GBM) had served as a convenient paradigm for some time, as the empirical evidence against GBM accumulated, the affine jump-diffusion (AJD) representation, pioneered by Merton (1976), gained wide acceptance primarily because it was shown to be consistent with the empirical features of asset returns (higher mode and excess kurtosis and skewness).

While the simple AJD specification admits a leptokurtic and asymmetric return distribution, it has proved to be inadequate in fully matching the sample moments of asset returns. In its most popular form, the AJD price process has a single jump component that captures the impact of news on security prices (Merton, 1976). News that leads to positive jumps in prices — good news — and news that leads to negative jumps in prices — bad news — are not distinguished by their intensity or distributional characteristics. This potential limitation of the simple jump-diffusion framework has led to two alternative specifications.

Under the Pareto–Beta jump-diffusion (PBJD) proposed by Ramezani and Zeng (1998), good and bad news are generated by two independent Poisson processes where the jump magnitudes are drawn from the Pareto and Beta distributions. Alternatively, Kou (2002) proposed the double exponential jump-diffusion (DEJD) where a single Poisson process with fixed intensity generates news, but the jump magnitudes representing abnormal up- and down-price movements are drawn from two independent exponential distributions. As Ramezani and Zeng (2007) have shown, the two models are closely related in that the parameters of one model may be retrieved from the other. Given the close kinship between these models, henceforth we will refer to them as the asymmetric affine jump-diffusion (AAJD) specifications.

The AAJD representation has gained popularity primarily because of its distributional flexibility. Furthermore, as Kou (2007) and others have shown, the AAJD specifications lead to nearly analytical option pricing formulas for certain exotic and path dependent options. This is a significant advantage as most of the existing methods for pricing options under the jump-diffusion processes are confined to plain vanilla European options. Because of these and other advantageous features (Kou, 2007, p. 86), the applications of the AAJD framework have been expanding in the literature. Kou (2007) and Ramezani and Zeng (2007) provide a survey of important applications of AAJD framework. More recent extensions of the AAJD representations to other areas of economics and finance include Bertrand and Prigent (2011), Bo et al. (2012), Dao and Jeanblanc (2006), Deng et al. (2012), Dotis et al. (2006), Moazeni et al. (2011) and Zhang et al. (2012).

Despite the growing interest in AAJD specification, estimation and empirical assessment of this model has received sparse attention to date. In practice, most studies have arbitrarily assumed “reasonable” parameter values under each specification and proceeded to carry out their intended analysis. A notable exception is Ramezani and Zeng (2007), who utilize maximum likelihood estimation (MLE) to provide parameter estimates for the AAJD specification. The empirical tests performed by these authors suggest that the AAJD specification provides a superior fit to asset returns, relative to Merton’s (1976) single jump and the classical GBM specifications.

1 In recent years, Bayesian estimation methods have been developed for the AJD models (Johannes and Polson, 2009), but these procedures have not been extended to the AAJD specifications. The primary objective of this paper is to fill this gap in the literature. To assess the accuracy of our proposed Bayesian estimation method, we provide a comparison of the parameter estimates with those obtained via MLE. This comparison is done using daily returns for identical sample periods; 1962–2003 for the S&P 500, 1973–2003 for the NASDAQ and 1999–2003 for various individual stocks. To our knowledge, this is the first such comparison for the AAJD models. We find that our Bayesian estimation procedure produces comparable parameter estimates, with significantly smaller standard errors. In Sec. 2, we briefly present the details of the AAJD specifications. In Sec. 3, we discuss our Bayesian estimation method. Using a simulation exercise, we demonstrate the reliability of our proposed method to estimate the parameters of AAJD. In Sec. 4, we present our empirical results and provide a comparison of the Bayesian and the MLE estimates. We also use our Bayesian estimation to estimate the parameters in the boom and bust periods of 2008–2010. In particular, we examine the differences in the parameter estimates using the bear, bull and combined market periods. The paper concludes with suggestions for future enhancements to this line of research.

---

For a survey of estimation procedures for AJD class of representations see Zhao (2008), Six (2006) and the papers in Ait-Sahalia and Hansen (2004).
recent spectacular market bust (2007–2009) and boom (2009–2010) periods, we examine how the parameter estimates differ under distinctly different economic conditions.

Keywords: Asset price processes; affine jump-diffusion; double exponential jump-diffusion; Markov chain Monte Carlo; Bayesian econometrics.

1. Introduction

Almost every aspect of modern finance theory, from valuations and portfolio choice to option pricing and corporate finance, as well as the ever expanding field of mathematical finance, critically depends upon the form of the probability distribution governing the asset returns. Although geometric brownian motion (GBM) had served as a convenient paradigm for some time, as the empirical evidence against GBM accumulated, the affine jump-diffusion (AJD) representation, pioneered by Merton (1976), gained wide acceptance primarily because it was shown to be consistent with the empirical features of asset returns (higher mode and excess kurtosis and skewness).

While the simple AJD specification admits a leptokurtic and asymmetric return distribution, it has proved to be inadequate in fully matching the sample moments of asset returns. In its most popular form, the AJD price process has a single jump component that captures the impact of news on security prices (Merton, 1976). News that leads to positive jumps in prices — good news — and news that leads to negative jumps in prices — bad news — are not distinguished by their intensity or distributional characteristics. This potential limitation of the simple jump-diffusion framework has led to two alternative specifications.

Under the Pareto–Beta jump-diffusion (PBJD) proposed by Ramezani and Zeng (1998), good and bad news are generated by two independent Poisson processes where the jump magnitudes are drawn from the Pareto and Beta distributions. Alternatively, Kou (2002) proposed the double exponential jump-diffusion (DEJD) where a single Poisson process with fixed intensity generates news, but the jump magnitudes representing abnormal up- and down-price movements are drawn from two independent exponential distributions. As Ramezani and Zeng (2007) have shown, the two models are closely related in that the parameters of one model may be retrieved from the other. Given the close kinship between these models, henceforth we will refer to them as the asymmetric affine jump-diffusion (AAJD) specifications.

The AAJD representation has gained popularity primarily because of its distributional flexibility. Furthermore, as Kou (2007) and others have shown, the AAJD specifications lead to nearly analytical option pricing formulas for certain exotic and path dependent options. This is a significant advantage as most of the existing methods for pricing options under the jump-diffusion processes are confined to plain vanilla European options. Because of these and other advantageous features (Kou, 2007, p. 86), the applications of the AAJD framework has been expanding in the literature. Kou (2007) and Ramezani and Zeng (2007) provide a survey of important applications of AAJD framework. More recent extensions of the AAJD representations to other areas of economics and finance include Bertrand and Prigent (2011), Bo et al. (2012), Dao and Jeanblanc (2006), Deng et al. (2012), Dotis et al. (2006), Moazeni et al. (2011) and Zhang et al. (2012).

Despite the growing interest in AAJD specification, estimation and empirical assessment of this model has received sparse attention to date. In practice, most studies have arbitrarily assumed “reasonable” parameter values under each specification and proceeded to carry out their intended analysis. A notable exception is Ramezani and Zeng (2007), who utilize maximum likelihood estimation (MLE) to provide parameter estimates for the AAJD specification. The empirical tests performed by these authors suggests that the AAJD specification provides a superior fit to asset returns, relative to Merton’s (1976) single jump and the classical GBM specifications.1

In recent years, Bayesian estimation methods have been developed for the AJD models (Johannes and Polson, 2009), but these procedures have not been extended to the AAJD specifications. The primary objective of this paper is to fill this gap in the literature. To assess the accuracy of our proposed Bayesian estimation method, we provide a comparison of the parameter estimates with those obtained via MLE. This comparison is done using daily returns for identical sample periods; 1962–2003 for the S&P 500, 1973–2003 for the NASDAQ and 1999–2003 for various individual stocks. To our knowledge, this is the first such comparison for the AAJD models. We find that our Bayesian estimation procedure produces comparable parameter estimates, with significantly smaller standard errors. In Sec. 2, we briefly present the details of the AAJD specifications. In Sec. 3, we discuss our Bayesian estimation method. Using a simulation exercise, we demonstrate the reliability of our proposed method to estimate the parameters of AAJD. In Sec. 4, we present our empirical results and provide a comparison of the Bayesian and the MLE estimates. We also use our Bayesian estimation to estimate the parameters in the boom and bust periods of 2008–2010. In particular, we examine the differences in the parameter estimates using the bear, bull and combined market periods. The paper concludes with suggestions for future enhancements to this line of research.

1For a survey of estimation procedures for AJD class of representations see Zhao (2008), Sin (2006) and the papers in Ait-Sahalia and Hansen (2004).
2. Asymmetric Affine Jump-Diffusion Models

In this section, we present two versions of the AAJD specifications for asset return processes. As noted, these extensions to the AJD specification have gained popularity because of their desirable theoretical and empirical features. Although these models both appear in the literature, they are equivalent in that both imply the same probability law for daily returns. Indeed, given parameters corresponding to one model, Ramezani and Zeng (2007) show how to derive the parameters for the other specification that yield identical dynamics. However, these two model specifications each reflect a particular intuition as to the underlying role played by jumps in the daily return process, and both have been found useful in applications. Therefore, in this paper we discuss the implementation of our proposed econometric methodology in both contexts.

2.1. Pareto–Beta jump diffusion

The PBJD posits that the return process contains two jump components, representing good and bad news arrival. Each type of news is generated by an independent Poisson process that lead to up- or down-jumps in prices, whose magnitudes are drawn from the Pareto and Beta distributions, respectively. This formulation is consistent with Milgrom (1981), who formalized the notion of "good" and "bad" news and showed that the arrival of good (bad) news about a firm’s prospects will always lead to a rise (fall) in its share price. Ramezani and Zeng (1998) present other plausible economic justifications for invoking a distinction between good and bad news. They note that for individual stocks, discontinuous up- and down-price movements may be a consequence of shocks to earnings expectations, significant changes in the operations and the financial structure of the firm, unexpected changes in its competitive environment or its organizational form, and sudden shifts in its corporate plans. Similarly, at the stock index level, major political events, macroeconomic policy decisions, such as a cut in interest rates by the Federal Reserve serves as the unexpected good news expectation, significant changes in the operations and the financial structure of the firm’s prospects will always lead to a rise (fall) in its share price. Ramezani and Zeng (2007) show that the distribution of Y is

\[ f_Y(y) = \left( \frac{\lambda_u}{\lambda_u + \lambda_d} \right) \eta_u e^{-\eta_u y} I(y \geq 0) + \left( \frac{\lambda_d}{\lambda_u + \lambda_d} \right) \eta_d e^{\eta_d y} I(y < 0). \]

Let \( S(t) \) denote the price of stock at time \( t \), the PBJD process can be represented by

\[ \frac{dS(t)}{S(t)} = \mu dt + \sigma dZ(t) + \sum_{j=1}^{\infty} (V^+_{i,j} - 1) dN_i(t), \]

where \( \mu \) and \( \sigma \) are the drift and volatility terms, \( Z(t) \) is a standard Wiener process, \( V^+ \) is the jump magnitude, and \( N_i(t) \) are independent Poisson processes with intensity parameters \( \lambda_j \) (\( j = u, d \) represent up- and down-jumps respectively). It is assumed that the up-jump magnitudes \( V^+ \) are distributed Pareto \( (\eta_u, 1) \) and the down-jump magnitudes \( V^- \) are distributed Beta \( (\eta_d, 1) \). Letting \( Y = \ln(V) \), Ramezani and Zeng (2007) show that the distribution of \( Y \) is

\[ f_Y(y) = \left( \frac{\lambda_u}{\lambda_u + \lambda_d} \right) \eta_u e^{-\eta_u y} I(y \geq 0) + \left( \frac{\lambda_d}{\lambda_u + \lambda_d} \right) \eta_d e^{\eta_d y} I(y < 0). \]

Let \( S(t) \) be the closing price at time (day) \( t \). The daily return for period \( i \) is defined as \( \tau_i = \ln \left( \frac{S(t_i)}{S(t_{i-1})} \right) \), and \( \epsilon = (\tau_1, \ldots, \tau_M) \) are the returns over the period \( i = 1, \ldots, M \). Under the PBJD specification, daily returns are independent and identically distributed (IID) with probability density function \( (pdf) \), \( f(\tau_i|\theta_{PBJD}) \), given by

\[ f(\tau_i|\theta_{PBJD}) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} P(m|\lambda_u)P(n|\lambda_d)f_{\text{Poisson}}(\tau_i|\theta_{PBJD}). \]

where \( \theta_{PBJD} = (\mu, \sigma, \lambda_u, \lambda_d, \eta_u, \eta_d) \) and \( P(j|\lambda) = \frac{e^{-\lambda} \lambda^j}{j!} \). The density \( f(\tau_i|\theta_{PBJD}) \) is the Poisson weighted mixture of the densities, \( f_{\text{Poisson}}(\tau_i|\theta_{PBJD}) \), for a return comprised of \( m \) up-jumps and \( n \) down-jumps. For our empirical analysis, we use time discretization of the continuous returns process to estimate the AAJD models. For this purpose, we follow the standard practice and calculate daily returns using close-to-close prices (\( \tau \)). As such, Eraker et al. (2003) find that this form of "time-discretizations" of the AJD price processes does not introduce any biases in the Bayesian parameter estimates and the discretization bias will be negligible when data is sampled at high frequencies such as daily.

2.2. Double exponential jump diffusion

The economic interpretation of the DEJD, proposed by Kou (2002), differs from the PBJD. Under the DEJD specification, a single Poisson process generates news and the jump magnitudes representing abnormal up- and down-price movements...
2. Asymmetric Affine Jump-Diffusion Models

In this section, we present two versions of the AAJD specifications for asset return processes. As noted, these extensions to the AID specification have gained popularity because of their desirable theoretical and empirical features. Although these models both appear in the literature, they are equivalent in that both imply the same probability law for daily returns. Indeed, given parameters corresponding to one model, Ramezani and Zeng (2007) show how to derive the parameters for the other specification that yield identical dynamics. However, these two model specifications each reflect a particular intuition as to the underlying role played by jumps in the daily return process, and both have been found useful in applications. Therefore, in this paper we discuss the implementation of our proposed econometric methodology in both contexts.

2.1. Pareto–Beta jump diffusion

The PBJD posits that the return process contains two jump components, representing good and bad news arrival. Each type of news is generated by an independent Poisson process that lead to up- or down-jumps in prices, whose magnitudes are drawn from the Pareto and Beta distributions, respectively. This formulation is consistent with Milgrom (1981), who formalized the notion of “good” and “bad” news and showed that the arrival of good (bad) news about a firm’s prospects will always lead to a rise (fall) in its share price. Ramezani and Zeng (1998) present other plausible economic justifications for invoking a distinction between good and bad news. They note that for individual stocks, discontinuous up- and down-price movements may be a consequence of shocks to earnings expectations, significant changes in the operations and the financial structure of the firm, unexpected changes in its competitive environment or its organizational form, and sudden shifts in its corporate plans. Similarly, at the stock index level, major political events, macroeconomic policy decisions, such as a cut (increase) in interest rates by the Federal Reserve serves as the unexpected good news arrival. In this paper, we discuss the implementation of our proposed econometric methodology in both contexts.

Let \( S(t) \) denote the price of stock at time \( t \), the PBJD process can be represented by

\[
\frac{dS(t)}{S(t^-)} = \mu dt + \sigma dZ(t) + \sum_{j=0}^{m} \left( V^u_j - 1 \right) dN_j(t),
\]

where \( \mu \) and \( \sigma \) are the drift and volatility terms, \( Z(t) \) is a standard Wiener process, \( V^u_j \) is the jump magnitude, and \( N_j(t) \) are independent Poisson processes with intensity parameters \( \lambda_j \) (\( j = u,d \) represent up- and down-jumps respectively). It is assumed that the up-jump magnitudes \( V^u \) are distributed Pareto \((\eta_u, \alpha_u)\) and the down-jump magnitudes \( V^d \) are distributed Beta \((\eta_d, 1)\). Letting \( Y = \ln(V) \), Ramezani and Zeng (2007) show that the distribution of \( Y \) is

\[
f_Y(y) = \begin{cases} \frac{\lambda_u}{\lambda_u + \lambda_d} \eta_u e^{-\eta_u y} I(y \geq 0) + \frac{\lambda_d}{\lambda_u + \lambda_d} \eta_d e^{\eta_d y} I(y < 0). \end{cases}
\]

Let \( S(t) \) be the closing price at time (day) \( t \). The daily return for period \( i \) is defined as \( r_i = \ln \left( \frac{S(t_i)}{S(t_{i-1})} \right) \), and \( \varepsilon = (r_1, \ldots, r_M) \) are the returns over the period \( i = 1, \ldots, M \). Under the PBJD specification, daily returns are independent and identically distributed (IID) with probability density function (pdf), \( f(r_i|\theta_{PBJD}) \), given by

\[
f(r_i|\theta_{PBJD}) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} P(m|\lambda_u)P(n|\lambda_d)P(m,n|r_i|\theta_{PBJD}).
\]

where \( \theta_{PBJD} = (\mu, \sigma, \lambda_u, \lambda_d, \eta_u, \eta_d) \) and \( P(j|\lambda) = \frac{\lambda^j e^{-\lambda}}{j!} \). The density \( f(r_i|\theta_{PBJD}) \) is the Poisson weighted mixture of the densities, \( f_{m,n}(r_i|\theta_{PBJD}) \), for a return comprised of \( n \) up-jumps and \( m \) down-jumps. For our empirical analysis, we use time discretization of the continuous returns process to estimate the AID models. For this purpose, we follow the standard practice and calculate daily returns using close-to-close prices \((r)\). As such, Eraker et al. (2003) find that this form of “time-discretizations” of the AID price processes does not introduce any biases in the Bayesian parameter estimates and the discretization bias will be negligible when data is sampled at high frequencies such as daily.

2.2. Double exponential jump diffusion

The economic interpretation of the DEJD, proposed by Kou (2002), differs from the PBJD. Under the DEJD specification, a single Poisson process generates news and the jump magnitudes representing abnormal up- and down-price movements.

\(^3\)Maher and McCurdy (2004) show that expansionary and contractionary economic periods are accompanied with unequal frequency of good and bad news arrivals. Section 4 provides direct evidence in support of this conjecture using recent data that span the dramatic bull and bear market periods (2007–2010).
are drawn from two independent exponential distributions, with the return process given by

\[
\frac{dS(t)}{S(t-)} = \mu dt + \sigma dZ(t) + \sum_{i=1}^{N(\lambda)}(V_i - 1)dN(\lambda),
\]

where \(N(\lambda)\) is an independent Poisson processes with intensity parameter \(\lambda\) and \(V_i\) is a sequence of IID random variables. Kou (2002) assumes that \(Y = \ln(V)\) has an IID mixture distribution of Exponentials is given by

\[
f_j(y) = p \cdot \eta e^{-\eta y}I(y \geq 0) + (1 - p) \cdot \eta e^{\eta y}I(y < 0),
\]

where \(p\) is the mixing probability. Under this specification, daily returns are IID with pdf, \(f(r|\theta_{DEJD})\), given by

\[
f(r|\theta_{DEJD}) = \sum_{k=0}^{\infty} P(k|\lambda)f_k(r|\theta_{DEJD}).
\]

where \(\theta_{DEJD} = (\mu, \sigma, \lambda, p, \eta, A)\) and the density \(f(r|\theta_{DEJD})\) is the Poisson weighted mixture of the densities, \(f_k(r|\theta_{DEJD})\), for a return comprised of \(k\) jumps.

2.3. The equivalence between PBJD and DEJD

Ramezani and Zeng (2007) established the connection between the PBJD and DEJD specifications. Relying on the fact the logarithm of Pareto and Beta distributed random variables is exponentially distributed and assuming that \(\lambda_u\) and \(\lambda_d\) are independent Poisson processes, they show that the parameters of one model may be retrieved from the other by imposing the following mathematical restrictions:

\[
\lambda = \lambda_u + \lambda_d,
\]

\[
p = \frac{\lambda_u}{\lambda}.
\]

For estimation purposes, both models have the same number of parameters, though their interpretation is different; the single jump rate governing news arrival under DEJD, \(\lambda\), is simply the sum of the two independent up- and down-jump news arrival rates under PBJD. Moreover, the probability of a draw from the upper tail of the double exponential distribution, \(\rho\), is determined by the relative arrival rate of the up-jumps to the total arrival rate. Note that relative to the Normal distribution, both specification are capable of generating a higher peak, positive or negative skewness, and positive kurtosis. Therefore, both models are likely to better match the empirical moments of returns.

It should be noted that when the price processes, \(S(t)\), follows the same probability law under both PBJD and DEJD specifications (i.e. the jumps magnitudes are exponentially distributed and the arrival rates are independent), the above mathematical restrictions will hold exactly. In that case, the two models are simply different parameterization of the same underlying process. Moreover, as our empirical results show, parameter estimates implied by observed data do not in general satisfy the above mathematical relationships. So, it does make a difference in terms of both implementation and interpretation which model formulation is used in practical applications.

2.4. The leverage effect

The “leverage effect” refers to the negative correlation between innovation in prices and both the implied and the realized volatility. The most cited explanation, due to Black (1976), ties this relationship to the effect a change in market valuation of a firm’s equity has on the degree of leverage in its capital structure, with an increase in leverage producing an increase in stock volatility.

Figlewski and Want (2000) use both returns and directly measured leverage to examine the hypothetical explanation for the leverage effect as it applies to the individual stocks in the S&P100 (OEX) index, and to the index itself. They find a strong leverage effect associated with declining stock prices, but also numerous anomalies that call into question its existence. They conclude that the leverage effect is “much weaker or non-existent” when positive price changes reduce leverage. They also note that the effect of leverage is too small for individual firms, too large for the index, and the impact of leverage on volatility appears to die out over a few months. Most importantly, they find that there is no effect on volatility when leverage changes because of a change in outstanding debt or shares, only when stock prices change. They conclude that the leverage effect is a “down market effect” that may have little direct connection to firm leverage.

At the modeling stage, there has been some work toward integrating the leverage effect into the AJD specification prior to undertaking any empirical work. This is often achieved by formally specifying a relationship between price innovations and instantaneous volatility of returns. For example, by allowing for stochastic volatility, stochastic jump intensity, or stochastic skewness. Alternatively, a stochastic volatility model with time-varying correlation between return and volatility innovations may be specified. Other methods include the constant elasticity of variance model or GARCH models that incorporate time-varying skewness. Figlewski and Want (2000) and Ait-Sahalia et al. (2013) provide a discussion of alternative specifications and estimation issues that arise.
are drawn from two independent exponential distributions, with the return process given by
\[ \frac{dS(t)}{S(t-)} = \mu dt + \sigma dZ(t) + \sum_{i=1}^{\text{~}} (V_i - 1) dN(\lambda t), \]
where \( N(\lambda t) \) is an independent Poisson processes with intensity parameter \( \lambda \) and \( V_i \) is a sequence of IID random variables. Kou (2002) assumes that \( Y = \ln(V) \) has an IID mixture distribution of Exponentials is given by
\[ f_y(y) = p \eta_y e^{-\eta_y y} I(y \geq 0) + (1-p) \eta_y e^{\eta_y y} I(y < 0), \]
where \( p \) is the mixing probability. Under this specification, daily returns are IID with pdf, \( f(r;\theta_{\text{DEJD}}) \), given by
\[ f(r;\theta_{\text{DEJD}}) = \sum_{k=0}^{\infty} P(k|\lambda) f_k(r;\theta_{\text{DEJD}}). \]
where \( \theta_{\text{DEJD}} = (\mu, \sigma, \lambda, p, \eta_y, \eta_d) \) and the density \( f(r;\theta_{\text{DEJD}}) \) is the Poisson weighted mixture of the densities, \( f_k(r;\theta_{\text{DEJD}}) \), for a return comprised of \( k \) jumps.

2.3. The equivalence between PBJD and DEJD

Ramezani and Zeng (2007) established the connection between the PBJD and DEJD specifications. Relying on the fact the logarithm of Pareto and Beta distributed random variables is exponentially distributed and assuming that \( \lambda_u \) and \( \lambda_d \) are independent Poisson processes, they show that the parameters of one model may be retrieved from the other by imposing the following mathematical restrictions:
\[ \lambda = \lambda_u + \lambda_d, \]
\[ p = \frac{\lambda_u}{\lambda}. \]

For estimation purposes, both models have the same number of parameters, though their interpretation is different; the single jump rate governing news arrival under DEJD, \( \lambda \), is simply the sum of the two independent up- and down-jump news arrival rates under PBJD. Moreover, the probability of a draw from the upper tail of the double exponential distribution, \( p \), is determined by the relative arrival rate of the upper jumps to the total arrival rate. Note that relative to the Normal distribution, both specification are capable of generating a higher peak, positive or negative skewness, and positive kurtosis. Therefore, both models are likely to better match the empirical moments of returns.

It should be noted that when the price processes, \( S(t) \), follows the same probability law under both PBJD and DEJD specifications (i.e. the jumps magnitudes are exponentially distributed and the arrival rates are independent), the above mathematical restrictions will hold exactly. In that case, the two models are simply different parameterization of the same underlying process. Moreover, as our empirical results show, parameter estimates implied by observed data do not in general satisfy the above mathematical relationships. So, it does make a difference in terms of both implementation and interpretation which model formulation is used in practical applications.

2.4. The leverage effect

The “leverage effect” refers to the negative correlation between innovation in prices and both the implied and the realized volatility. The most cited explanation, due to Black (1976), ties this relationship to the effect a change in market valuation of a firm’s equity has on the degree of leverage in its capital structure, with an increase in leverage producing an increase in stock volatility.

Figlewski and Want (2000) use both returns and directly measured leverage to examine the hypothetical explanation for the leverage effect as it applies to the individual stocks in the S&P100 (OEX) index, and to the index itself. They find a strong leverage effect associated with declining stock prices, but also numerous anomalies that call into question its existence. They conclude that the leverage effect is “much weaker or non-existent” when positive price changes reduce leverage. They also note that the effect of leverage is too small for individual firms, too large for the index, and the impact of leverage on volatility appears to die out over a few months. Most importantly, they find that there is no effect on volatility when leverage changes because of a change in outstanding debt or shares, only when stock prices change. They conclude that the leverage effect is a “down market effect” that may have little direct connection to firm leverage.

At the modeling stage, there has been some work toward integrating the leverage effect into the AJD specification prior to undertaking any empirical work. This is often achieved by formally specifying a relationship between price innovations and instantaneous volatility of returns. For example, by allowing for stochastic volatility, stochastic jump intensity, or stochastic skewness. Alternatively, a stochastic volatility model with time-varying correlation between return and volatility innovations may be specified. Other methods include the constant elasticity of variance model or GARCH models that incorporate time-varying skewness. Figlewski and Want (2000) and Ait-Sahalia et al. (2013) provide a discussion of alternative specifications and estimation issues that arise.
The AAJD specifications do not formally account for the leverage effect. However, these specifications can be easily extended to capture this effect by incorporating time-varying parameters (i.e., volatility, jump size and intensities). For example, since the leverage effect implies that “bad news” about returns are more likely to be associated with increases in volatility, we can simply link instantaneous volatility to the down-jump components of AAJD model. It is important to note that such additional “model flexibility” aimed at capturing empirical feature of the data will considerably increase the complexity of the estimation procedure, leading to significant trade-offs in terms of the statistical validity of the estimated parameters (see Ait-Sahalia et al., 2013). To conclude, the extent empirical literature shows that the impact of innovations in prices on volatility is generally asymmetric: other things equal, declines in stock prices are accompanied by larger increases in volatility than the decline in volatility that accompanies rising stock markets. We believe that the AAJD indirectly captures this asymmetry since it uses different distributions for the tails of return distribution.

3. Bayesian Parameter Estimation

Bayesian estimation methods have become more attractive and broadly used in finance and economic research in recent years (Johannes and Polson, 2009). Bayesian estimation methods that are well suited to the AJD processes have been studied by Eraker et al. (2003), Li et al. (2008), Goncalves and Roberts (2012), Jacquier et al. (2007), and Johannes and Polson (2009). In this section, we present our proposed Bayesian estimation method for the above discretely observed jump-diffusion models. Our proposed methodology parallels that of Eraker et al. (2003) and Johannes and Polson (2009) in every respect, except we allow for two jump components in the return processes. This estimation approach consists of an MCMC algorithm to obtain a sample from each model’s joint posterior distribution of the parameters and the latent variables. Our main contribution to the literature concerns the development of an algorithm that is particularly well suited to the AAJD class of specifications. As we demonstrate below, Bayesian estimation of these discretely observed jump-diffusion models requires careful development of a model truncation scheme, as well as the development of an appropriate Gibbs Sampler (GS).

The MCMC method breaks the joint distribution of model parameters into its conditional, which are of lower dimension and easier to sample from. To see the details of our approach, let us focus on the PBJD specification and assume that the elements of \( \theta_{PBJD} \) are mutually independent. The posterior distribution, \( f(\theta_{PBJD}|y,t) \), summarizes the information about the parameters, given the returns data \( (t) \) and the unobserved states variables \( (Y) \). After integrating out the latent variables, we utilize the Bayes’ rule to write the posterior distribution for the parameters of the PBJD specification as

\[
\begin{align*}
    f(\theta_{PBJD}|y,t) &\propto L(t|\theta_{PBJD}) f(\theta_{PBJD}) \\
    &\propto L(t|\theta_{PBJD}) f(\mu, \sigma^2, \lambda_d, \lambda_u, \eta_d) \\
    &\propto L(t|\theta_{PBJD}) f(\mu|\theta_{PBJD}) f(\sigma^2|\theta_{PBJD}) f(\lambda_d|\theta_{PBJD}) f(\lambda_u|\theta_{PBJD}) f(\eta_d|\theta_{PBJD})
\end{align*}
\]

where the likelihood function is \( L(t|\theta_{PBJD}) = \prod_{i=1}^{M} L(t_i|\theta_{PBJD}) \) and \( L(t_i|\theta_{PBJD}) \) is the marginal likelihood function given by Eq. (1). Relying on the independence assumption, the joint distribution of the parameters is written as the product of the marginal prior distributions.

Bayesian estimation requires that the researcher propose a prior distribution for the model parameters. For our models the choice of the priors is very important, as each distribution serves critical economic and statistical roles. Moreover, the choice allows us to incorporate important economic information such as positivity of key parameters (i.e. \( \sigma > 0, \lambda > 0 \), and \( p \in [0,1] \)). Statistically, the specification of the priors can also help impose “stationarity” or to separate up- and down-jump components.

Following Eraker et al. (2003), Johannes and Polson (2009), Jacquier and Polson (2011) and the extant Bayesian AJD estimation literature, we assume uninformed (improper, vague) priors for the distribution of the model parameters. This decision is dictated by the fact that, aside from “toy models”, it is very difficult to specify informative priors for complex and high dimensional models such as the AAJD. An advantage of assuming uninformed priors is that it relieves the researcher from having to select from many alternative priors (as well as estimating hyper parameters). Assuming uninformed, on the other hand, places the burden of parameter estimation entirely on the observed data.

Even with uninformed priors, it is extremely difficult, if not impossible, to analytically establish the posterior distribution, \( f(\theta_{PBJD}|y,t) \), because of the complexities associate with Eq. (1). Hence, a suitable “proposal distribution” for the joint posterior distribution of the model parameters rarely exists and consequently even an “appropriate” Metropolis-Hasting (MH) algorithm would have low acceptance rates and would be computationally unattractive (Eraker et al., 2003).

Estimating AAJD type models is a challenging task under both the MLE and the Bayesian estimation. The return density given by Eq. (1) contains infinite summation. In practice, the summations are truncated after a few iterations when the value of the likelihood function does not change “significantly”. To avoid the infinite summation problem, Ball and Torous (1983) proposed a Bernoulli approximation to the density function for a process with a single jump component.

\[f(\theta_{PBJD}|y,t) \propto \prod_{i=1}^{M} f(t_i|\theta_{PBJD}) \]

\[f(t_i|\theta_{PBJD}) = \prod_{j=1}^{J} p_i(t_i) \]

where \( p_i(t_i) \) is the distribution of the return at time \( t_i \), and \( J \) is the number of jumps.

For a survey of MCMC methods and recent developments in their application see Greyer (2011).
The AAJD specifications do not formally account for the leverage effect. However, these specifications can be easily extended to capture this effect by incorporating time-varying parameters (i.e., volatility, jump size and intensities). For example, since the leverage effect implies that “bad news” about returns are more likely to be associated with increases in volatility, we can simply link instantaneous volatility to the down-jump components of AAJD model. It is important to note that such additional “model flexibility” aimed at capturing empirical feature of the data will considerably increase the complexity of the estimation procedure, leading to significant trade-offs in terms of the statistical validity of the estimated parameters (see Ait-Sahalia et al., 2013). To conclude, the extant empirical literature shows that the impact of innovations in prices on volatility is generally asymmetric: other things equal, declines in stock prices are accompanied by larger increases in volatility than the decline in volatility that accompanies rising stock markets. We believe that the AAJD indirectly captures this asymmetry since it uses different distributions for the tails of return distribution.

3. Bayesian Parameter Estimation
Bayesian estimation methods have become more attractive and broadly used in finance and economic research in recent years (Johannes and Polson, 2009). Bayesian estimation methods that are well suited to the AJD processes have been studied by Eraker et al. (2003), Li et al. (2008), Goncalves and Roberts (2012), Jacquier et al. (2007), and Johannes and Polson (2009). In this section, we present our proposed Bayesian estimation method for the above discrete-observed jump-diffusion models. Our proposed methodology parallels that of Eraker et al. (2003) and Johannes and Polson (2009) in every respect, except we allow for two jump components in the return processes. This estimation approach consists of a MCMC algorithm to obtain a sample from each model’s joint posterior distribution of the parameters and the latent variables.1 Our main contribution to the literature concerns the development of an algorithm that is particularly well suited to the AAJD class of specifications. As we demonstrate below, Bayesian estimation of these discrete-observed jump-diffusion models requires careful development of a model truncation scheme, as well as the development of an appropriate Gibbs Sampler (GS).

The MCMC method breaks the joint distribution of model parameters into its conditionals, which are of lower dimension and easier to sample from. To see that the elements of \( \theta_{\text{PJJD}} \) are mutually independent. The posterior distribution, \( f(\theta_{\text{PJJD}} | Y, t) \), summarizes the information about the parameters, given the returns data \( (r) \) and the unobserved states variables \( (Y) \). After integrating out the latent variables, we utilize the Bayes’ rule to write the posterior distribution for the parameters of the PBJD specification as

\[
f(\theta_{\text{PJJD}} | Y, t) \propto L(\theta_{\text{PJJD}}) f(\theta_{\text{PJJD}})
\propto L(\theta_{\text{PJJD}}) f(\mu, \sigma^2, \lambda_0, \lambda_d, \eta_0, \eta_d)
\propto L(\theta_{\text{PJJD}}) f(\mu) f(\sigma^2) f(\lambda_0) f(\lambda_d) f(\eta_0) f(\eta_d)
\]

where the likelihood function is \( L(\theta_{\text{PJJD}}) = \prod_{t=1}^{T} f(r_t | \theta_{\text{PJJD}}) \) and \( f(r_t | \theta_{\text{PJJD}}) \) is the marginal likelihood function given by Eq. (1). Relying on the independence assumption, the joint distribution of the parameters is written as the product of the marginal prior distributions.

Bayesian estimation requires that the researcher propose a prior distribution for the model parameters. For our models the choice of the priors is very important, as each distribution serves critical economic and statistical roles. Moreover, the choice allows us to incorporate important economic information such as positivity of key parameters (i.e. \( \sigma > 0, \lambda > 0, \) and \( p \in [0, 1] \)). Statistically, the specification of the priors can also help impose “stationarity” or to separate up- and down-jump components.

Following Eraker et al. (2003), Johannes and Polson (2009), Jacquier and Polson (2011) and the extant Bayesian AJD estimation literature, we assume un-informed (improper, vague) priors for the distribution of the model parameters. This decision is dictated by the fact that, aside from “toy models”, it is very difficult to specify informative priors for complex and high dimensional models such as the AAJD. An advantage of assuming uninformative priors is that it relieves the researcher from having to select from many alternative priors (as well as estimating hyper parameters). Assuming uninformative, on the other hand, places the burden of parameter estimation entirely on the observed data.

Even with uninformative priors, it is extremely difficult, if not impossible, to analytically establish the posterior distribution, \( f(\theta_{\text{PJJD}} | Y, t) \), because of the complexities associate with Eq. (1). Hence, a suitable “proposal distribution” for the joint posterior distribution of the model parameters rarely exists and consequently even an “appropriate” Metropolis-Hasting (MH) algorithm would have low acceptance rates and would be computationally unattractive (Eraker et al., 2003).

Estimating AJD type models is a challenging task under both the MLE and the Bayesian estimation. The return density given by Eq. (1) contains infinite summations. In practice, the summations are truncated, after a few iterations when the value of the likelihood function does not change significantly. To avoid the infinite summation problem, Ball and Torous (1983) proposed a Bernoulli approximation to the density function for a process with a single jump component.

---

1 For a survey of MCMC methods and recent developments in their application see Greayer (2011).
Such approximation requires that during the time period under consideration, either no-jump or at most a single jump occurs.

In the extant literature that relies on the approximate Bernoulli mixture model, \( \lambda \) is interpreted as the mixing parameter (probability), with its domain restricted to the closed interval \([0,1]\). This is clearly inconsistent with the interpretation of \( \lambda > 0 \) as the arrival rate of a jump process. In the case of DEJD specification which contains one arrival rate, Kou (2002, footnote 7, p. 1091) proposes the Bernoulli approximation, again requiring \( \lambda \in [0,1] \) to act as the mixing parameter. The PBJD model has two arrival rates and it is not clear if the Bernoulli approximation may be applied and whether it is an appropriate approximation to the full model. To summarize, the Bernoulli approximation for both AAJD models may be inadequate and it may have severe consequences for the estimated parameters.

To see this, note that in Eq. \((2)\), \( \lambda \) is the parameter governing the arrival rate of the Poisson jumps and not the probability of a jump, \( P(A=1|\lambda) = e^{-\lambda} \). Treating \( \lambda \) as a mixing probability constrains the estimation such that \( \lambda \in [0,1] \) rather than allowing \( \lambda \in \mathbb{R}^+ \). Furthermore, after values of \( \lambda \) are estimated, the resulting density function may not have mixing probabilities that sum to one because the summations are truncated and rarely normalized. Our experience with simulation has shown that the approximate Bernoulli mixture models do not reliably estimate values of \( \lambda \) greater than 0.5.

An important contribution of the present paper is to offer an elegant methodology to overcome this shortcoming, while retaining the simplicity of the Bernoulli approximation. Specifically, we extend the Bernoulli approach for the AID models to the AAJD models with a truncated trinomial mixture approximation to Eq. \((1)\). Our hypothesis for the random generation of the data includes three mutually exclusive states: either no jumps occur, or at least one up-jump and no down-jumps occur, or at least one down-jump and no up-jumps occur. Let \( N_u \) be the number of up-jumps with \( N_u \sim \text{Poisson}(\lambda_u) \) and \( N_d \) be the number of down-jumps with \( N_d \sim \text{Poisson}(\lambda_d) \). Assume \( N_u \) and \( N_d \) are independent. The probabilities of the three mutually exclusive states are

\[
P(N_u \geq 1, N_d = 0) = (1 - e^{-\lambda_d}) e^{-\lambda_u} = \lambda_u, \quad (5)
\]

\[
P(N_u = 0, N_d \geq 1) = e^{-\lambda_u} (1 - e^{-\lambda_d}) = \lambda_d, \quad (6)
\]

\[
P(N_u = 0, N_d = 0) = (1 - \lambda_u - \lambda_d). \quad (7)
\]

The purpose of this parameterization is to allow for more than one jump event to drive the discontinuous component of the price process. Since we are working with interday close-to-close returns, we assume that the “net jump size” in a given day can only be in one direction, up or down. Hence, in our scheme \( f_{0,m} \approx f_{0,1} \) and \( f_{0,0} \approx f_{1,0} \) for all \( m \in \mathbb{N} \) and \( f_{0,0} \) would represent a “wash” of the many up- and down-jumps from the intraday price movements. Note that the influence of the proposed approximation for \( f_{0,m} \) \( (f_{0,0}) \) is quite negligible. Our proposed approximation will result in a truncated daily return density function for the PBJD specification

\[
f(r|\theta^*_{\text{PBJD}}) = (1 - \lambda_u - \lambda_d) f_{0,0}(r) + \lambda_u f_{1,0}(r) + \lambda_d f_{0,1}(r),
\]

where \( \theta^*_{\text{PBJD}} = (\mu, \sigma^2, \lambda_u, \lambda_d, \eta_u, \eta_d) \) and

\[
f_{0,0}(r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(r - \mu)^2}
\]

\[
f_{1,0}(r) = \frac{\eta_u}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-\eta_u \frac{1}{2}(y - \mu)^2} dy,
\]

\[
f_{0,1}(r) = \frac{\eta_d}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^0 e^{\eta_d \frac{1}{2}(y - \mu)^2} dy,
\]

where \( f(r|\theta^*_{\text{PBJD}}) \) is an approximation for the full-model density in Eq. \((1)\). We utilize this approximate trinomial mixture model in conjunction with a Gibbs Sampler (GS) to estimate the parameters of the PBJD specifications. Path averaging of the inverse of the above state probabilities, Eqs. \((5)-(7)\), provides estimates of \( \lambda_u \) and \( \lambda_d \). Additionally, the estimates of \( \mu \) and \( \sigma \) are obtained by path averaging of Eqs. \((3)\) and \((4)\). Because of the mass preservation and reverse transformations, the interpretation of \( \lambda_u \) and \( \lambda_d \) is preserved and unlike the standard Bernoulli approximation, these parameters are not confined to the \([0,1]\) interval. Furthermore, \( \eta_u \) and \( \eta_d \) represent the average “net jump” amplitude per day. A summary of the full model, its approximation, and the GS algorithm are in Table 1. The details of the sampling schemes for the latent variables and the parameters are presented in the Appendix.

For the DEJD specification, we use the Bernoulli approach but sample from two different exponentially distributed jump magnitudes (see p. 1091 Kou, 2002). Specifically, it is assumed that in each time period, no jumps or at least one jump event may occur. The probability of these mutually exclusive events are

\[
P(N = 0) = e^{-\lambda} \equiv \lambda', \quad (8)
\]

\[
P(N \geq 1) = 1 - \lambda' \quad (9)
\]

4Jorion (1988) demonstrated the limitations of the Bernoulli approximation for the MLE estimation of parameters of the AID process.

5This is particularly true as the mixing probability weights associated with more than one jump per period decline exponentially.
Such approximation requires that during the time period under consideration, either no-jump or at most a single jump occurs.

In the extant literature that relies on the approximate Bernoulli mixture model, \( \lambda \) is interpreted as the mixing parameter (probability), with its domain restricted to the closed interval \([0,1]\). This is clearly inconsistent with the interpretation of \( \lambda > 0 \) as the arrival rate of a jump process. In the case of DEJD specification which contains one arrival rate, Kou (2002, footnote 7, p. 1091) proposes the Bernoulli approximation, again requiring \( \lambda \in [0,1] \) to act as the mixing parameter. The PBJD model has two arrival rates and it is not clear if the Bernoulli approximation may be applied and whether it is an appropriate approximation to the full model. To summarize, the Bernoulli approximation for both AAJD models may be inadequate and it may have severe consequences for the estimated parameters.

To see this, note that in Eq. (2), \( \lambda \) is the parameter governing the arrival rate of the Poisson jumps and not the probability of a jump. The Poisson distribution is given by

\[
P(N = k | \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}
\]

for \( k = 0, 1, 2, \ldots \), where \( \lambda > 0 \). Treating \( \lambda \) as a mixing probability constrains the estimation such that \( \lambda \in [0,1] \) rather than allowing \( \lambda \in \mathbb{R}^+ \). Furthermore, after values of \( \lambda \) are estimated, the resulting density function may not have mixing probabilities that sum to one because the summation of these mutually exclusive events are three mutually exclusive states: either no jumps occur, or at least one up-jump and no down-jumps occur, or at least one down-jump and no up-jumps occur. Let \( N_u \) be the number of up-jumps with \( N_u \sim \text{Poisson}(\lambda_u) \) and \( N_d \) be the number of down-jumps with \( N_d \sim \text{Poisson}(\lambda_d) \). Assume \( N_u \) and \( N_d \) are independent. The probabilities of the three mutually exclusive states are

1. No jumps occur, \( N_u = 0, N_d = 0 \):
   \[
P(N_u = 0, N_d = 0) = e^{-\lambda_u}e^{-\lambda_d} = \lambda_u^0 \lambda_d^0.
   \]
2. At least one up-jump and no down-jumps occur, \( N_u \geq 1, N_d = 0 \):
   \[
P(N_u \geq 1, N_d = 0) = 1 - e^{-\lambda_u} = \lambda_u^1 - \lambda_u^0.
   \]
3. At least one down-jump and no up-jumps occur, \( N_u = 0, N_d \geq 1 \):
   \[
P(N_u = 0, N_d \geq 1) = 1 - e^{-\lambda_d} = \lambda_d^1 - \lambda_d^0.
   \]

The purpose of this parameterization is to allow for more than one jump event to drive the discontinuous component of the price process. Since we are working with interday close-to-close returns, we assume that the “net jump size” in a given day can only be in one direction, up or down. Hence, in our scheme \( f_{0,m} \approx f_{0,1} \) and \( f_{0,0} \approx f_{1,0} \) for all \( n,m \in \mathbb{N} \), and \( f_{0,0} \) would represent “wash” of the many up- and down-jumps from the intraday price movements. Note that the influence of the proposed approximation for \( f_{0,m} (0,0) \) is quite negligible. Our proposed approximation will result in a truncated daily return density function for the PBJD specification

\[
f(r) \bigg| \theta_{\text{PBJD}} = (1 - \lambda_u - \lambda_d) f_{0,0}(r) + \lambda_u^1 f_{0,1}(r) + \lambda_d^1 f_{1,0}(r),
\]

where \( \theta_{\text{PBJD}} = (\mu, \sigma^2, \lambda_u, \lambda_d, \eta_u, \eta_d) \) and

\[
f_{0,0}(r) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(r - \mu)^2},
\]

\[
f_{0,1}(r) = \frac{\eta_u}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(y - \mu)^2} dy,
\]

\[
f_{1,0}(r) = \frac{\eta_d}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(y - \mu)^2} dy,
\]

where \( f(r | \theta_{\text{PBJD}}) \) is an approximation for the full-model density in Eq. (1). We utilize this approximate trinomial mixture model in conjunction with a Gibbs Sampler (GS) to estimate the parameters of the PBJD specifications. Path averaging of the inverse of the above state probabilities, Eqs. (5)-(7), provides estimates of \( \lambda_u \) and \( \lambda_d \). Additionally, the estimates of \( \lambda \) and \( p \) are obtained by path averaging of Eqs. (3) and (4). Because of the mass preservation and reverse transformation, the interpretation of \( \lambda_u \) and \( \lambda_d \) is preserved and unlike the standard Bernoulli approximation, these parameters are not confined to the \([0,1]\) interval. Furthermore, \( \eta_u \) and \( \eta_d \) represent the average “net jump” amplitude per day. A summary of the full model, its approximation, and the GS algorithm are in Table 1. The details of the sampling schemes for the latent variables and the parameters are presented in the Appendix.

For the DEJD specification, we use the Bernoulli approach but sample from two different exponentially distributed jump magnitudes (see p. 1091 Kou, 2002). Specifically, it is assumed that in each time period, no jumps or at least one jump event may occur. The probability of these mutually exclusive events are

\[
P(N = 0) = e^{-\lambda} \equiv \lambda',
\]

\[
P(N \geq 1) = 1 - \lambda'
\]

Footnote 7: Jorion (1988) demonstrated the limitations of the Bernoulli approximation for the MLE estimation of parameters of the AJD process.
and it is assumed that $f_i \approx f_0$ for all $n \in N$. Under these assumptions, the approximate density for the DEJD model is given by

$$f(r_i|\theta_{DEJD}) = \lambda f_0(r_i) + (1 - \lambda) f_i(r_i),$$

where $	heta_{DEJD} = (\mu, \sigma^2, \lambda, p, \eta_0, \eta_d)$ and

$$f_i(r_i|\theta_{DEJD}) = \pi \eta_0 \int_{0}^{\infty} e^{-\eta_0 y} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y - \mu - \lambda^2)^2}{2\sigma^2}} dy_i + (1 - \pi) \eta_d \int_{-\infty}^{0} e^{-\eta_d y} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(y + \mu - \lambda^2)^2}{2\sigma^2}} dy_i$$

for $f_0 = f_{0.0}$. We utilize this approximation in conjunction with the GS to estimate $	heta_{DEJD}$. Path averaging of the inverse of the above state probabilities, Eqs. (8) and (9), provide estimates of $\lambda$. Additionally, estimates of $\lambda_0$ and $\lambda_d$ are obtained by path averaging the inverse of Eqs. (3) and (4). Table 2 presents the full DEJD model, its approximation, and our proposed GS algorithm to estimate the parameters.

### Table 2. Bayesian estimation procedure for DEJD.

| Full model approximation | $f(r_i|\theta_{DEJD}) = \sum_{\eta_0} \sum_{\eta_d} P(\eta_0|\eta_d) P(\lambda|\eta_0) P(p|\lambda) P(\mu|\lambda) P(\sigma^2)$ |
|--------------------------|-------------------------------------------------------------------------------------------------|
| $f(r_i|\theta_{DEJD}) = f_0(r_i) + f_i(r_i)$ | $f(r_i|\theta_{DEJD}) = \lambda f_0(r_i) + (1 - \lambda) f_i(r_i)$ |

Pick starting values:
1. Sample $I_i, Y_i, \lambda_i$ for $i = 1, \ldots, M$
2. Sample $\lambda_i \sim \text{Beta}$ and $\mu \sim \text{Beta}$
3. Sample $\eta_0 \sim \text{Gamma}$, $\eta_d \sim \text{Gamma}$
4. Sample $\mu \sim \text{Normal}$, $\sigma^2$ with MH

Note: Repeat steps 1-4 until convergence.

### 3.1. Simulation study

In this section, we present the results from a Monte Carlo simulation designed to assess the efficacy of our Bayesian estimation algorithm. Undertaking the simulation study is an important step for the AJD class of models. As Johannes and Polson (2009) note, conducting a simulation study is important because time discretization of AJD process can potentially bias the parameter estimates. Given the complexity of AJD, simulation enables the researcher to verify that parameters and their standard errors can be reliably estimated for the given sample size. Moreover, simulation provides useful diagnostics and the chance to check the efficiency and convergence of the proposed algorithm and determine how long to run the proposed algorithm. Finally, simulation enables us to verify that the proposed method provides parameter estimates from an identifiable posterior distribution that we can reasonably reconstruct with our GS method.

Our study simulation based on different parametrizations of the AJD models and for each set of parameters, we simulate a data set with 2000 observations. Each sample path is generated using a range of "true" parameter values under the PBJD and DEJD specifications, as shown in Table 3. The assumed parameter values generate sample paths that are similar to the observed daily returns for the time period used for estimation. As Table 3 shows, we experiment with a variety of parameter values, including low and high jump intensity that correspond to infrequent and frequent discontinuous movement in prices. We use a suitable burn-in period, where we discard 25% of the generated samples, and no "mixing" of the paths is required. We conduct standard diagnostics and verify that the chains have converged.

Table 3 summarizes the results of the simulation study. It is clear that the proposed procedure produces accurate parameter estimates. It appears that some parameters, particularly the mean jump sizes, are estimated less precisely as we anticipated would happen given the assumptions used to construct our approximate trinomial mixture model. Overall, most estimated parameters are close to their true values and have small standard errors. It appears the simulation study validates our MCMC algorithm and the integrity of our code. The simulation also confirms that our distributional assumptions, both our priors and conditionals, are appropriate and our MCMC algorithm offers desirable convergence properties. It is also clear that the impact of discretization error on the parameter estimates is small.

### 3.2. Model diagnostics

A number of practical issues influence the development and application of MCMC algorithms. The theoretical convergence properties of the Markov Chain and the actual verification that the chain produced by the MCMC algorithm has converged, are of prime concern. Johannes and Polson (2009) discuss the basic theory of the
Bayesian Estimation of AJD Processes

3.1. Simulation study

In this section, we present the results from a Monte Carlo simulation designed to assess the efficacy of our Bayesian estimation algorithm. Undertaking the simulation study is an important step for the AJD class of models. As Johannes and Polson (2009) note, conducting a simulation study is important because time discretization of AJD process can potentially bias the parameter estimates. Given the complexity of AJD, simulation enables the researcher to verify that parameters and their standard errors can be reliably estimated for the given sample size. Moreover, simulation provides useful diagnostics and the chance to check the efficiency and convergence of the proposed algorithm and determine how long to run the proposed algorithm. Finally, simulation enables us to verify that the proposed method provides parameter estimates from an identifiable posterior distribution that we can reasonably reconstruct with our GS method.

Our study simulation based on different parametrizations of the AJD models and for each set of parameters, we simulate a data set with 2000 observations. Each sample path is generated using a range of “true” parameter values under the PBJD and DEJD specifications, as shown in Table 3. The assumed parameter values generate sample paths that are similar to the observed daily returns for the time period used for estimation. As Table 3 shows, we experiment with a variety of parameter values, including low and high jump intensity that correspond to infrequent and frequent discontinuous movement in prices. We use a suitable burn-in period, where we discard 25% of the generated samples, and no “mixing” of the paths is required. We conduct standard diagnostics and verify that the chains have converged.

Table 3 summarizes the results of the simulation study. It is clear that the proposed procedure produces accurate parameter estimates. It appears that some parameters, particularly the mean jump sizes, are estimated less precisely as we anticipated would happen given the assumptions used to construct our approximate trinomial mixture model. Overall, most estimated parameters are close to their true values, and have small standard errors. It appears the simulation study validates our parameter values under the assumptions, both our priors and conditionals, are appropriate and our MCMC algorithm offers desirable convergence properties. It is also clear that the impact of discretization error on the parameter estimates is small.

3.2. Model diagnostics

A number of practical issues influence the development and application of MCMC algorithms. The theoretical convergence properties of the Markov Chain and the actual verification that the chain produced by the MCMC algorithm has converged, are of prime concern. Johannes and Polson (2009) discuss the basic theory of the

---

Table 1. Bayesian estimation procedure for PBJD.

| Full model approximation | $f(r; \theta_{PBJD}) = \prod_{i=1}^{n} \prod_{j=1}^{m} P(r_i | \theta_{PBJD})$ |
|-------------------------|--------------------------------------------------|
| Pick starting values    | $\lambda_{D}^{o}, \lambda_{J}^{o}, \eta_{D}, \eta_{J}, \sigma^{2}$ |
| 1. Sample               | $L_{i}, Y_{i}, Y_{i}^{d}$ for $i = 1, \ldots, M$ |
| 2. Sample               | $L_{i} \sim$ Gamma, $\eta_{D} \sim$ Gamma |
| 3. Sample               | $\mu \sim$ Normal, $\sigma^{2}$ with MH |

Table 2. Bayesian estimation procedure for DEJD.

| Full model approximation | $f(r; \theta_{DEJD}) = \prod_{i=1}^{n} \prod_{j=1}^{m} P(r_i | \theta_{DEJD})$ |
|-------------------------|--------------------------------------------------|
| Pick starting values    | $\lambda_{D}^{o}, \lambda_{J}^{o}, \eta_{D}, \eta_{J}, \sigma^{2}$ |
| 1. Sample               | $L_{i}, Y_{i}, Y_{i}^{d}$ for $i = 1, \ldots, M$ |
| 2. Sample               | $L_{i} \sim$ Beta and $\mu \sim$ Normal |
| 3. Sample               | $\eta_{D} \sim$ Gamma, $\eta_{J} \sim$ Gamma |
| 4. Sample               | $\mu \sim$ Normal, $\sigma^{2}$ with MH |

Note: Repeat steps 1-4 until convergence.
Table 3. MCMC parameter estimates from simulated data.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \hat{\alpha} )</th>
<th>( \beta )</th>
<th>( \eta_0 )</th>
<th>( \eta_1 )</th>
<th>( \lambda_0 )</th>
<th>( \lambda_1 )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBID-true</td>
<td>0.3500</td>
<td>0.1429</td>
<td>10.000</td>
<td>10.000</td>
<td>0.0500</td>
<td>0.3000</td>
<td>-0.0060</td>
<td>0.0200</td>
</tr>
<tr>
<td>PBID</td>
<td>0.3543</td>
<td>0.1311</td>
<td>10.838</td>
<td>11.174</td>
<td>0.0464</td>
<td>0.3080</td>
<td>-0.0061</td>
<td>0.0200</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.3649</td>
<td>0.1464</td>
<td>10.713</td>
<td>11.084</td>
<td>0.0534</td>
<td>0.3115</td>
<td>-0.0062</td>
<td>0.0210</td>
</tr>
<tr>
<td>PBID-true</td>
<td>0.8750</td>
<td>0.4000</td>
<td>47.000</td>
<td>30.000</td>
<td>0.3500</td>
<td>0.0500</td>
<td>-0.0060</td>
<td>0.0200</td>
</tr>
<tr>
<td>PBID</td>
<td>0.8724</td>
<td>0.4171</td>
<td>44.368</td>
<td>28.047</td>
<td>0.3667</td>
<td>0.0504</td>
<td>-0.0065</td>
<td>0.0194</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.8571</td>
<td>0.4413</td>
<td>44.136</td>
<td>27.740</td>
<td>0.3810</td>
<td>0.0603</td>
<td>-0.0065</td>
<td>0.0195</td>
</tr>
<tr>
<td>PBID-true</td>
<td>0.3500</td>
<td>0.1429</td>
<td>10.000</td>
<td>10.000</td>
<td>0.0500</td>
<td>0.3000</td>
<td>-0.0060</td>
<td>0.0200</td>
</tr>
<tr>
<td>PBID</td>
<td>0.3156</td>
<td>0.1320</td>
<td>12.443</td>
<td>10.345</td>
<td>0.0417</td>
<td>0.2739</td>
<td>-0.0069</td>
<td>0.0194</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.3293</td>
<td>0.1472</td>
<td>12.619</td>
<td>10.317</td>
<td>0.0486</td>
<td>0.2807</td>
<td>-0.0069</td>
<td>0.0194</td>
</tr>
<tr>
<td>PBID-true</td>
<td>0.8750</td>
<td>0.4000</td>
<td>47.000</td>
<td>30.000</td>
<td>0.3500</td>
<td>0.0500</td>
<td>-0.0060</td>
<td>0.0200</td>
</tr>
<tr>
<td>PBID</td>
<td>0.9061</td>
<td>0.3978</td>
<td>44.135</td>
<td>31.071</td>
<td>0.3472</td>
<td>0.0461</td>
<td>-0.0062</td>
<td>0.0201</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.8817</td>
<td>0.3933</td>
<td>42.482</td>
<td>31.516</td>
<td>0.3472</td>
<td>0.0461</td>
<td>-0.0062</td>
<td>0.0201</td>
</tr>
</tbody>
</table>

Note: The table presents parameter estimates under the PBID and DEJD specifications using simulated data. Standard errors appear below the estimates.

convergence of Markov Chains and suggest simple tests to check for the convergence of the algorithm. These authors indicate that chains generated by the MH and the GS algorithms have special theoretical properties which assure convergence, without reference to the specifics of a particular algorithm. That is both algorithms generate Markov Chains that are “time-reversible” and have the target distribution as an “invariant distribution”. Given that our proposed MCMC also uses the MH and GS samplers, it also meets the theoretical requirements for the convergence of the chains.

While theoretical underpinnings of our proposed MCMC algorithm guarantees its convergence, as Johannes and Polson (2009) note, it is important to formally diagnose convergence from the realized output of the chain. There are a number of diagnostics that achieve this objective. Johannes and Polson (2009) suggest that looking at parameter trace plots, which show the history of the chain, enables the researcher to detect if the chain is stuck in a region of the state space and not converging. They also note that it is useful to look at the correlation structure of draws by computing the autocorrelation function (ACF), as algorithms that have low ACF may not converge. We conducted these diagnostic tests for our proposed MCMC algorithm (both the actual and simulated data). We find that our algorithm is efficient and convergent. The diagnostics are shown in Figs. 1 and 2, which show that parameter values converge after the initial “burn” period.

As Ait-Sahalia (2004) has shown, the ability to disentangle jumps from the diffusion component of returns is critical to estimating the parameters of the AJD processes, particularly the volatility parameter (\( \sigma \)). We believe it is also critical to disentangle jumps from the diffusion component to accurately estimate \( \lambda_\alpha \) and \( \lambda_\beta \). To our knowledge, there does not exist a Bayesian estimation method that adequately resolves the entanglement problem. In the absence of an accepted Bayesian jump detection procedure, researchers have relied upon the outcome of their simulation studies to determine if their proposed estimation algorithm is delivering an accurate estimate of the model parameters. This is, the approach we adopt in this paper as well. In fact our simulation study suggests that our proposed algorithm

---

6 Greyer (2011) provided additional diagnostic checks. We refrain from producing our trace plots and ACF graphs to save space. Both are available from the authors upon request.

7 In recent years Bayesian methods for detecting jumps have been proposed in Lee and Huang (2010) and Lee (2012).
Table 3. MCMC parameter estimates from simulated data.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\lambda$</th>
<th>$P$</th>
<th>$q_1$</th>
<th>$q_2$</th>
<th>$\lambda_d$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBJD-true</td>
<td>0.3500</td>
<td>0.1429</td>
<td>10.000</td>
<td>10.000</td>
<td>0.0500</td>
<td>0.3000</td>
<td>-0.0060</td>
</tr>
<tr>
<td>PBJD</td>
<td>0.3543</td>
<td>0.1311</td>
<td>10.838</td>
<td>11.174</td>
<td>0.0464</td>
<td>0.3080</td>
<td>-0.0001</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.3649</td>
<td>0.1464</td>
<td>10.713</td>
<td>11.084</td>
<td>0.0534</td>
<td>0.3115</td>
<td>-0.0062</td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.0000</td>
<td>0.026</td>
<td>0.011</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>PBJD-true</td>
<td>0.8750</td>
<td>0.4000</td>
<td>47.000</td>
<td>30.000</td>
<td>0.3500</td>
<td>0.0500</td>
<td>-0.0060</td>
</tr>
<tr>
<td>PBJD</td>
<td>0.8724</td>
<td>0.4171</td>
<td>44.368</td>
<td>28.047</td>
<td>0.3607</td>
<td>0.0504</td>
<td>-0.0005</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.8571</td>
<td>0.4413</td>
<td>44.136</td>
<td>27.740</td>
<td>0.3810</td>
<td>0.0603</td>
<td>-0.0065</td>
</tr>
<tr>
<td></td>
<td>0.0011</td>
<td>0.0026</td>
<td>0.0927</td>
<td>0.1219</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>DEJD-true</td>
<td>0.3500</td>
<td>0.1429</td>
<td>10.000</td>
<td>10.000</td>
<td>0.0500</td>
<td>0.3000</td>
<td>-0.0060</td>
</tr>
<tr>
<td>PBJD</td>
<td>0.3156</td>
<td>0.1320</td>
<td>12.443</td>
<td>10.345</td>
<td>0.0417</td>
<td>0.2739</td>
<td>-0.0069</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.3293</td>
<td>0.1472</td>
<td>12.619</td>
<td>10.317</td>
<td>0.0486</td>
<td>0.2807</td>
<td>-0.0069</td>
</tr>
<tr>
<td></td>
<td>0.0004</td>
<td>0.0005</td>
<td>0.0359</td>
<td>0.0106</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>DEJD-true</td>
<td>0.8750</td>
<td>0.4000</td>
<td>47.000</td>
<td>30.000</td>
<td>0.3500</td>
<td>0.0500</td>
<td>-0.0060</td>
</tr>
<tr>
<td>PBJD</td>
<td>0.9061</td>
<td>0.3978</td>
<td>44.135</td>
<td>31.071</td>
<td>0.3623</td>
<td>0.0355</td>
<td>-0.0066</td>
</tr>
<tr>
<td>DEJD</td>
<td>0.8817</td>
<td>0.3933</td>
<td>42.482</td>
<td>31.516</td>
<td>0.3472</td>
<td>0.0461</td>
<td>-0.0062</td>
</tr>
<tr>
<td></td>
<td>0.0011</td>
<td>0.0031</td>
<td>0.1075</td>
<td>0.1965</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Note: The table presents parameter estimates under the PBJD and DEJD specifications using simulated data. Standard errors appear below the estimates.

convergence of Markov Chains and suggest simple tests to check for the convergence of the algorithm. These authors indicate that chains generated by the MH and the GS algorithms have special theoretical properties which assure convergence, without reference to the specifics of a particular algorithm. That is both algorithms generate Markov Chains that are “time-reversible” and have the target distribution as an “invariant distribution”. Given that our proposed MCMC also uses the MH and GS samplers, it also meets the theoretical requirements for the convergence of the chains.

While theoretical underpinnings of our proposed MCMC algorithm guarantees its convergence, as Johannes and Polson (2009) note, it is important to formally diagnose convergence from the realized output of the chain. There are a number of diagnostics that achieve this objective. Johannes and Polson (2009) suggest that looking at parameter trace plots, which show the history of the chain, enables the researcher to detect if the chain is stuck in a region of the state space and not converging. They also note that it is useful to look at the correlation structure of draws by computing the autocorrelation function (ACF), as algorithms that have low ACF may not converge. We conducted these diagnostic tests for our proposed MCMC algorithm (both the actual and simulated data). We find that our algorithm is efficient and convergent. The diagnostics are shown in Figs. 1 and 2, which show that parameter values converge after the initial “burn” period.

As Ait-Sahalia (2004) has shown, the ability to disentangle jumps from the diffusion component of returns is critical to estimating the parameters of the AJD processes, particularly the volatility parameter ($\sigma$). We believe it is also critical to disentangle jumps from the diffusion component to accurately estimate $\lambda_d$ and $\lambda_g$. To our knowledge, there does not exist a Bayesian estimation method that adequately resolves the entanglement problem. In the absence of an accepted Bayesian jump detection procedure, researchers have relied upon the outcome of their simulation studies to determine if their proposed estimation algorithm is delivering accurate estimate of the model parameters. This is, the approach we adopt in this paper as well. In fact our simulation study suggests that our proposed algorithm

6 Greyer (2011) provided additional diagnostic checks. We refrain from producing our trace plots and ACF graphs to save space. Both are available from the authors upon request.

7 In recent years Bayesian methods for detecting jumps have been proposed in Lee and Hang (2010) and Lee (2012).
will be reasonably accurate in estimating the volatility and the jump intensity components of the return process.

4. Data and Results

The empirical analysis presented in this section has a number of objectives. First, utilizing the proposed Bayesian estimation method, we will present the estimated parameters for the AAJD models, using time series data for individual stocks and important broad market indexes, spanning the period 1962–2010 and selected sub-periods. Second, we will provide comparisons with previous studies by using the same data (same span and frequency). Finally, we provide a direct comparison of the parameter estimates under our proposed Bayesian estimation procedure versus the MLE approach.

To permit comparison with other studies, particularly the MLE analysis presented by Ramezani and Zeng (2007), we focus on daily (value weighted) returns for the S&P 500 and the NASDAQ composite indexes. The S&P 500 daily return series (dividend adjusted) spans the period 7/1962 through 12/2010 (N = 10446). The NASDAQ series spans the period 1/1973 through 12/2010 (N = 7828).\(^8\) We also use daily returns for five individual stocks, with large kurtosis (range of 3 to 10).

10. The data on individual stocks spans the period 1/1999 through 12/2003 (N = 1256). The selected firms, which trade on NASDAQ and NYSE, are followed by a large number of analysts, and are highly liquid. These characteristics are important given the event driven nature of the AAJD models. We use the proposed Bayesian estimation methodology to directly estimate the parameter of both specifications for each series.

Table 4 presents the sample statistics for the indexes and individual stock returns for the period under consideration. The large range of return values, particularly for the indexes, reflects significant booms and crashes that occurred during the sample period. All returns are highly skewed and have large kurtosis.

Table 5 reports both MLE and Bayesian parameter estimates for the S&P 500 and the NASDAQ composite indexes.\(^9\) Focusing on the parameter estimates, we find that estimates of \(\mu\) and \(\sigma\) are very similar across models and estimation techniques. This similarity is also true for the frequency of news arrivals, \(\lambda_e\) and \(\lambda_d\).\(^10\) Turning to the mean up- and down-jump amplitudes, \(\eta_e\) and \(\eta_d\), we find that these parameter estimates are invariant to model specification under Bayesian estimation. However, under MLE, \(\eta_e\) and \(\eta_d\) estimates are significantly larger.

---

\(^8\)No dividend adjusted series are available for NASDAQ index since few firms on this exchange pay dividends.

\(^9\)The MLE estimates are taken from the Ramezani and Zeng (2007) study.

\(^10\)The inverse, \(\lambda_e^{-1}\) and \(\lambda_d^{-1}\), provide an estimate of the inter-arrival times.
will be reasonably accurate in estimating the volatility and the jump intensity components of the return process.

4. Data and Results

The empirical analysis presented in this section has a number of objectives. First, utilizing the proposed Bayesian estimation method, we will present the estimated parameters for the AAJD models, using time series data for individual stocks and important broad market indexes, spanning the period 1962–2010 and selected sub-periods. Second, we will provide comparisons with previous studies by using the same data (same span and frequency). Finally, we provide a direct comparison of the parameter estimates under our proposed Bayesian estimation procedure versus the MLE approach.

To permit comparison with other studies, particularly the MLE analysis presented by Ramezani and Zeng (2007), we focus on daily (value weighted) returns for the S&P 500 and the NASDAQ composite indexes. The S&P 500 daily return series (dividend adjusted) spans the period 7/1962 through 12/2010 (N = 10446). The NASDAQ series spans the period 1/1973 through 12/2010 (N = 7828).\(^8\) We also use daily returns for five individual stocks, with large kurtosis (range of 3 to 10). The data on individual stocks spans the period 1/1999 through 12/31/2003 (N = 1256). The selected firms, which trade on NASDAQ and NYSE, are followed by a large number of analysts, and are highly liquid. These characteristics are important given the event driven nature of the AAJD models. We use the proposed Bayesian estimation methodology to directly estimate the parameter of both specifications for each series.

Table 4 presents the sample statistics for the indexes and the individual stock returns for the period under consideration. The large range of return values, particularly for the indexes, reflects significant booms and crashes that occurred during the sample period. All returns are highly skewed and have large kurtosis.

Table 5 reports both MLE and Bayesian parameter estimates for the S&P 500 and the NASDAQ composite indexes.\(^9\) Focusing on the parameter estimates, we find that estimates of \(\mu\) and \(\sigma\) are very similar across models and estimation techniques. This similarity is also true for the frequency of news arrivals, \(\lambda_u\) and \(\lambda_d\).\(^10\) Turning to the mean up- and down-jump amplitudes, \(\eta_u\) and \(\eta_d\), we find that these parameter estimates are invariant to model specification under Bayesian estimation. However, under MLE, \(\eta_u\) and \(\eta_d\) estimates are significantly larger.

\(^8\)No dividend adjusted series are available for NASDAQ index since few firms on this exchange pay dividends.

\(^9\)The MLE estimates are taken from the Ramezani and Zeng (2007) study.

\(^10\)The inverse, \(\tilde{\epsilon}_u\) and \(\tilde{\epsilon}_d\), provide an estimate of the inter-arrival times.

---

![Figure 2. MCMC paths of the parameter samples for the NASDAQ return series from 1/1973 through 12/2010.](image)
Consequently, we expect to obtain different estimates of the jump parameters when (i.e. Bull or Bear phases of the markets). As Ramezani and Zeng (2007) observed, the relative magnitude of the jump parameters, $\lambda_1$, $\lambda_2$, $\eta_a$, and $\eta_d$, can lead to large positive (Bull) or negative (Bear) adjustments in the drift of the return processes. Consequently, we expect to obtain different estimates of the jump parameters when implying smaller mean jump amplitudes. Finally, the standard errors of the estimate are significantly lower under Bayesian estimation across the board.

Figure 3 presents the estimated PBJD distribution for the S&P 500 using the Bayesian parameter estimates in Table 5. The top panel show the daily return distribution. The bottom panel decomposes the return distribution into the GBM, up-jump, and down-jump components. As the bottom panel shows, the asymmetry and leptokurtosis of the returns is captured by the estimated PBJD model. Note that the left-tail reflects the 1987 market crash and other significant drops.

Table 6 presents similar comparison of MLE versus Bayesian estimation using individual stock data. Again, we find that estimates of $\mu$ and $\sigma$ are very similar across models and estimation techniques. However, we find that estimates of the jump components, $\lambda_1$, $\lambda_2$, $\eta_a$, and $\eta_d$, vary significantly across the models and the estimation techniques. This is to be expected as total volatility for stocks contains a significant idiosyncratic component and stock returns are highly skewed and leptokurtic.

The foregoing comparisons of Bayesian and the MLE estimates used historical data without regards to prevailing market conditions during the selected data period (i.e. Bull or Bear phases of the markets). As Ramezani and Zeng (2007) observed, the relative magnitude of the jump parameters, $\lambda_1$, $\lambda_2$, $\eta_a$, and $\eta_d$, can lead to large positive (Bull) or negative (Bear) adjustments in the drift of the return processes. Consequently, we expect to obtain different estimates of the jump parameters when we use data that contains one Bull or Bear market epoch. To explore this conjunction, we focus on highly volatile recent data, spanning the period May 2007 through December 2010. We use our Bayesian estimation procedure to estimate the parameters of PBJD and DEJD for three periods: The Bear Market period (05/2007 through 03/2009), the Bull Market period (04/2009 through 12/2010), and the combined cycle (05/2007 through 12/2010).

Table 7 reports the results for the recent period and the Bull and the Bear sub-periods for the S&P 500 and the NASDAQ indexes. First note that the drift ($\mu$) and the volatility of the Brownian motion ($\sigma$) are comparable across all three epochs. We find the down-jump arrival rate, $\lambda_d$, is significantly larger than the up-jump arrival rate, $\lambda_u$, during the Bear period. During this same period, the mean jump amplitudes, $\eta_a$ and $\eta_d$ are similar but notably larger than the combined and the Bull period. These results confirm the conjecture of Ramezani and Zeng (2007) noted above. However, we find that the proportion of total volatility due to the jump component is markedly larger during the Bear period than the combined and the Bull period. The estimated parameters for the Bear period lead to negative skewness in both the risk neutral and the physical returns distributions, suggesting that the probability of a large decrease in stock prices exceeds the probability of a large increase. Jackwerth and Rubinstein (1996) termed this phenomenon as "crashophobia". The economic rationale for crashophobia is that put options are used as hedging

![Table 5. Comparison of MCMC and MLE for indexes.](image)

![Figure 3. PBJD distribution for S&P 500 (7/1/1962–12/31/2003) using the MCMC parameter estimates from Table 4.](image)
Consequently, we expect to obtain different estimates of the jump parameters when positive (Bull) or negative (Bear) adjustments in the drift of the return processes (i.e., Bull or Bear phases of the markets). As Ramezani and Zeng (2007) noted above, we find that the proportion of total volatility due to the jump component is markedly larger during the Bear period than the combined and the Bull period.

The estimated parameters for the Bear period lead to negative skewness in both the risk neutral and the physical returns distributions, suggesting that the probability of a large decrease in stock prices exceeds the probability of a large increase. Jackwerth and Rubinstein (1996) termed this phenomenon as "crashophobia." The economic rationale for crashophobia is that put options are used as hedging手段...
Bayesian Estimation of AID Processes

Table 6. Comparison of MCMC and MLE for individual stocks.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Model — method</th>
<th>$P$</th>
<th>$\lambda$</th>
<th>$\eta_0$</th>
<th>$\eta_1$</th>
<th>$\lambda_0$</th>
<th>$\lambda_1$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>PRJD—MLE</td>
<td>0.8755</td>
<td>0.3969</td>
<td>47.25</td>
<td>28.97</td>
<td>0.3475</td>
<td></td>
<td>0.0059</td>
<td>0.0210</td>
</tr>
<tr>
<td></td>
<td>PRJD—MCMC</td>
<td>0.8508</td>
<td>0.5894</td>
<td>46.95</td>
<td>36.83</td>
<td>0.3312</td>
<td>0.1282</td>
<td>0.0065</td>
<td>0.0190</td>
</tr>
<tr>
<td></td>
<td>DEJD—MCMC</td>
<td>0.8028</td>
<td>0.5645</td>
<td>41.42</td>
<td>37.27</td>
<td>0.1113</td>
<td>-0.061</td>
<td>0.0206</td>
<td></td>
</tr>
<tr>
<td>HYBD</td>
<td>PRJD—MLE</td>
<td>0.3921</td>
<td>1.1884</td>
<td>19.41</td>
<td>33.25</td>
<td>0.6600</td>
<td>0.7224</td>
<td>-0.0133</td>
<td>0.0230</td>
</tr>
<tr>
<td></td>
<td>PRJD—MCMC</td>
<td>0.4585</td>
<td>1.1312</td>
<td>18.54</td>
<td>24.60</td>
<td>0.5186</td>
<td>0.6126</td>
<td>0.0030</td>
<td>0.0131</td>
</tr>
<tr>
<td></td>
<td>DEJD—MCMC</td>
<td>0.4536</td>
<td>1.4258</td>
<td>16.97</td>
<td>25.05</td>
<td>0.6467</td>
<td>0.7791</td>
<td>0.0005</td>
<td>0.0218</td>
</tr>
<tr>
<td>INTC</td>
<td>PRJD—MLE</td>
<td>0.8475</td>
<td>0.4000</td>
<td>47.22</td>
<td>24.49</td>
<td>0.3390</td>
<td>0.0610</td>
<td>-0.0366</td>
<td>0.0281</td>
</tr>
<tr>
<td></td>
<td>PRJD—MCMC</td>
<td>0.5176</td>
<td>0.8357</td>
<td>47.85</td>
<td>49.39</td>
<td>0.4362</td>
<td>0.4031</td>
<td>0.0002</td>
<td>0.0142</td>
</tr>
<tr>
<td></td>
<td>DEJD—MCMC</td>
<td>0.4858</td>
<td>1.1026</td>
<td>46.69</td>
<td>49.34</td>
<td>0.5356</td>
<td>0.5670</td>
<td>0.0005</td>
<td>0.0144</td>
</tr>
<tr>
<td>LCBM</td>
<td>PRJD—MLE</td>
<td>0.7051</td>
<td>0.4276</td>
<td>24.22</td>
<td>20.57</td>
<td>0.5015</td>
<td>0.1261</td>
<td>-0.0551</td>
<td>0.0234</td>
</tr>
<tr>
<td></td>
<td>PRJD—MCMC</td>
<td>0.4846</td>
<td>1.1437</td>
<td>31.73</td>
<td>38.98</td>
<td>0.5106</td>
<td>0.6332</td>
<td>0.0005</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>DEJD—MCMC</td>
<td>0.4705</td>
<td>1.5225</td>
<td>29.75</td>
<td>37.92</td>
<td>0.7257</td>
<td>0.8168</td>
<td>0.0008</td>
<td>0.0109</td>
</tr>
<tr>
<td>MFRI</td>
<td>PRJD—MLE</td>
<td>0.4885</td>
<td>1.4696</td>
<td>33.07</td>
<td>35.67</td>
<td>0.7179</td>
<td>0.7517</td>
<td>0.0002</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>PRJD—MCMC</td>
<td>0.4985</td>
<td>1.6921</td>
<td>29.26</td>
<td>32.03</td>
<td>0.5444</td>
<td>0.5477</td>
<td>0.0002</td>
<td>0.0065</td>
</tr>
<tr>
<td></td>
<td>DEJD—MCMC</td>
<td>0.4984</td>
<td>1.9761</td>
<td>29.28</td>
<td>32.00</td>
<td>0.9849</td>
<td>0.9912</td>
<td>0.0002</td>
<td>0.0064</td>
</tr>
</tbody>
</table>

Table 7. MCMC parameter estimates for indexes.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Model — method</th>
<th>$P$</th>
<th>$\lambda$</th>
<th>$\eta_0$</th>
<th>$\eta_1$</th>
<th>$\lambda_0$</th>
<th>$\lambda_1$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>PRJD</td>
<td>0.3784</td>
<td>1.1489</td>
<td>50.67</td>
<td>26.01</td>
<td>0.3348</td>
<td>0.7144</td>
<td>0.0025</td>
<td>0.0035</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.4045</td>
<td>1.6277</td>
<td>78.75</td>
<td>70.88</td>
<td>0.6938</td>
<td>1.0039</td>
<td>0.0024</td>
<td>0.0042</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PRJD</td>
<td>0.2191</td>
<td>1.3805</td>
<td>63.69</td>
<td>60.11</td>
<td>0.2980</td>
<td>1.0825</td>
<td>0.0050</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.2764</td>
<td>2.6841</td>
<td>61.94</td>
<td>60.54</td>
<td>0.7931</td>
<td>2.0710</td>
<td>0.0024</td>
<td>0.0042</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PRJD</td>
<td>0.4039</td>
<td>1.2635</td>
<td>104.84</td>
<td>103.35</td>
<td>0.6241</td>
<td>0.6394</td>
<td>0.0015</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.4803</td>
<td>3.0172</td>
<td>102.69</td>
<td>102.35</td>
<td>1.0432</td>
<td>1.5940</td>
<td>0.0017</td>
<td>0.0051</td>
</tr>
</tbody>
</table>

Note: The tables present estimates for five stocks under the PRJD and DEJD specifications. The daily returns span the period 1/1/1999-12/31/2003 (N = 1,256). Standard errors appear below the estimates.

Figure 4 presents the decomposed PRJD estimated distributions for the S&P 500 and NASDAQ using the Bayesian parameter estimates for the Bear, Bull and the combined periods from Table 7. For both indexes, the figures validate the existence of the "crashophobia", as shown by the significant contribution of the down-jump component in the model.

For S&P 500's Bull period, we find $\lambda_d$ is comparable to $\lambda_u$. Whereas, both $\eta$s are smaller and similar than the combined and the Bear period estimates. Consequently, the jump components together have a less significant impact on returns, effectively offsetting one another, and allowing $\mu$ to be the dominant force that pushes up the index level (see Fig. 4). For the NASDAQ's Bull period, $\lambda_d$ is larger than $\lambda_u$. That is the arrival of “good news” further amplifies the positive drift of this index (see Fig. 4). Again, both $\eta$s are similar but smaller than the combined and the
Table 6. Comparison of MCMC and MLE for individual stocks.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Model — method</th>
<th>( P )</th>
<th>( \lambda )</th>
<th>( \eta_s )</th>
<th>( \eta_d )</th>
<th>( \lambda_d )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>PBJD — MLE</td>
<td>0.8755</td>
<td>0.3996</td>
<td>47.35</td>
<td>28.97</td>
<td>0.3475</td>
<td>0.0494</td>
<td>-0.0059</td>
</tr>
<tr>
<td></td>
<td>PBJD — MCMC</td>
<td>0.8508</td>
<td>0.3894</td>
<td>46.98</td>
<td>43.88</td>
<td>0.7312</td>
<td>0.1282</td>
<td>-0.0065</td>
</tr>
<tr>
<td></td>
<td>DEJD — MCMC</td>
<td>0.8038</td>
<td>0.5645</td>
<td>41.42</td>
<td>37.27</td>
<td>0.4532</td>
<td>0.1113</td>
<td>-0.0061</td>
</tr>
<tr>
<td></td>
<td>HYBD — MLE</td>
<td>0.3921</td>
<td>1.1884</td>
<td>19.41</td>
<td>33.25</td>
<td>0.4660</td>
<td>0.7224</td>
<td>-0.0013</td>
</tr>
<tr>
<td></td>
<td>PBJD — MCMC</td>
<td>0.4585</td>
<td>1.1312</td>
<td>18.54</td>
<td>24.60</td>
<td>0.5186</td>
<td>0.6126</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>DEJD — MCMC</td>
<td>0.4536</td>
<td>1.4258</td>
<td>16.97</td>
<td>25.05</td>
<td>0.6467</td>
<td>0.7791</td>
<td>0.0002</td>
</tr>
<tr>
<td>INTC</td>
<td>PBJD — MLE</td>
<td>0.8475</td>
<td>0.4000</td>
<td>47.22</td>
<td>24.49</td>
<td>0.3390</td>
<td>0.0610</td>
<td>-0.0036</td>
</tr>
<tr>
<td></td>
<td>PBJD — MCMC</td>
<td>0.5176</td>
<td>0.8337</td>
<td>47.85</td>
<td>49.39</td>
<td>0.4326</td>
<td>0.4031</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>DEJD — MCMC</td>
<td>0.4958</td>
<td>1.1126</td>
<td>46.69</td>
<td>49.34</td>
<td>0.5356</td>
<td>0.5670</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>LCBM — PBJD</td>
<td>0.7051</td>
<td>0.4276</td>
<td>24.22</td>
<td>20.57</td>
<td>0.3015</td>
<td>0.1261</td>
<td>-0.0051</td>
</tr>
<tr>
<td></td>
<td>PBJD — MCMC</td>
<td>0.4464</td>
<td>1.1437</td>
<td>31.73</td>
<td>39.18</td>
<td>0.5105</td>
<td>0.6332</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>DEJD — MCMC</td>
<td>0.4705</td>
<td>1.5425</td>
<td>29.75</td>
<td>37.92</td>
<td>0.7257</td>
<td>0.8168</td>
<td>0.0001</td>
</tr>
<tr>
<td>MFRI</td>
<td>PBJD — MLE</td>
<td>0.4885</td>
<td>1.4696</td>
<td>33.07</td>
<td>35.67</td>
<td>0.7179</td>
<td>0.7517</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>PBJD — MCMC</td>
<td>0.4985</td>
<td>1.0921</td>
<td>29.26</td>
<td>32.03</td>
<td>0.5444</td>
<td>0.5477</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>DEJD — MCMC</td>
<td>0.4985</td>
<td>1.9761</td>
<td>29.28</td>
<td>32.00</td>
<td>0.9849</td>
<td>0.9912</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Note: The table presents estimates for five stocks under the PBJD and DEJD specifications. The daily returns span the period 1/1/1999–12/31/2003 (N = 1,256). Standard errors appear below the estimates.

Table 7. MCMC parameter estimates for indexes.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Model — method</th>
<th>( P )</th>
<th>( \lambda )</th>
<th>( \eta_s )</th>
<th>( \eta_d )</th>
<th>( \lambda_d )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>PBJD</td>
<td>0.3784</td>
<td>1.1489</td>
<td>80.67</td>
<td>72.16</td>
<td>0.4348</td>
<td>0.7141</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.4087</td>
<td>1.6977</td>
<td>78.75</td>
<td>70.88</td>
<td>0.6938</td>
<td>1.0309</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td>NASD</td>
<td>0.4008</td>
<td>0.9542</td>
<td>73.92</td>
<td>66.72</td>
<td>0.8000</td>
<td>1.1239</td>
<td>0.0021</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PBJD</td>
<td>0.2159</td>
<td>1.3805</td>
<td>63.69</td>
<td>60.11</td>
<td>0.2980</td>
<td>1.0825</td>
<td>0.0050</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.2769</td>
<td>2.8641</td>
<td>61.94</td>
<td>60.54</td>
<td>0.7931</td>
<td>2.0761</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td>NASD</td>
<td>0.4939</td>
<td>1.2635</td>
<td>104.84</td>
<td>103.03</td>
<td>0.6241</td>
<td>0.6394</td>
<td>0.0015</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PBJD</td>
<td>0.4803</td>
<td>3.0172</td>
<td>102.69</td>
<td>102.35</td>
<td>1.4732</td>
<td>1.5940</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.4021</td>
<td>0.0242</td>
<td>0.17</td>
<td>0.17</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>NASD</td>
<td>PBJD</td>
<td>0.3514</td>
<td>1.1352</td>
<td>76.75</td>
<td>70.63</td>
<td>0.3989</td>
<td>0.7363</td>
<td>0.0032</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.3499</td>
<td>1.9247</td>
<td>74.93</td>
<td>70.38</td>
<td>0.6735</td>
<td>1.2512</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>NASD</td>
<td>0.0013</td>
<td>0.0140</td>
<td>0.16</td>
<td>0.10</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PBJD</td>
<td>0.1999</td>
<td>1.3374</td>
<td>61.58</td>
<td>60.55</td>
<td>0.2673</td>
<td>1.0701</td>
<td>0.0055</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.2720</td>
<td>2.2792</td>
<td>60.00</td>
<td>62.29</td>
<td>0.6199</td>
<td>1.6593</td>
<td>0.0050</td>
</tr>
<tr>
<td></td>
<td>NASD</td>
<td>0.5428</td>
<td>1.2989</td>
<td>99.80</td>
<td>96.46</td>
<td>0.7051</td>
<td>0.5938</td>
<td>0.0012</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>PBJD</td>
<td>0.5674</td>
<td>3.7589</td>
<td>99.47</td>
<td>95.19</td>
<td>2.1328</td>
<td>1.6261</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>DEJD</td>
<td>0.0010</td>
<td>0.0217</td>
<td>0.12</td>
<td>0.13</td>
<td>-</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Note: The table presents estimates for S&P-500 and NASDAQ indexes under the PBJD and DEJD specifications using recent data. Standard errors appear below the estimates.

Figure 4 presents the decomposed PBJD estimated distributions for the S&P 500 and NASDAQ using the Bayesian parameter estimates for the Bear, Bull and the combined periods from Table 7. For both indexes, the figures validate the existence of the “crashophobia”, as shown by the significant contribution of the down-jump component in the model.

For S&P 500’s Bull period, we find \( \lambda_d \) is comparable to \( \lambda_s \). Whereas, both \( \eta_s \) are similar and smaller than the combined and the Bear period estimates. Consequently, the jump components together have a less significant impact on returns, effectively offsetting one another, and allowing \( \mu \) to be the dominant force that pushes up the index level (see Fig. 4). For the NASDAQ’s Bull period, \( \lambda_s \) is larger than \( \lambda_d \). That is the arrival of “good news” further amplifies the positive drift of this index (see Fig. 4). Again, both \( \eta_s \) are similar but smaller than the combined and the
Bear period estimates. The good news component accelerates the drift of the return process.

The parameter estimates for the alternative AAJD specifications should be nearly identical when all the underlying distributional assumptions are satisfied, as shown in Sec. 2.3. In our simulation exercise, the noted assumptions are satisfied and as a result the parameter estimates reported in Table 3 are nearly identical. When utilizing actual return data, our analysis shows that the majority of the estimated parameters found in Tables 5–7 are notably different under PBJD and DEJD. This divergence of the the parameter estimates suggests at least one violation of the underlying assumptions: the independence of the returns, parameters, and jump events; their distributional forms; and the equivalency of the mixture model components.

5. Conclusion

In this study, we develop a Bayesian estimation method for estimating the parameters of the Asymmetric Affine Jump-Diffusion models. We provide an empirical assessment of these models using daily returns for the S&P 500 and the NASDAQ indexes, as well as specific stocks. We complete our analysis by providing a comparison of the estimated parameters under Bayesian estimation and MLE. We find that, in general, the Bayesian estimates are consistent with the MLE estimates. However, our Bayesian estimation approach is computationally more efficient and yields smaller standard errors.

We find that the AAJD models are consistent with the empirical features of return processes. We study the behavior of the models’ parameters during different market epochs (Bull and Bear periods) and find that with the introduction of the jump components, the volatility component due to the GBM part of the return process ($\sigma$) is constant across epochs. This finding is in contrast to the stochastic volatility models that show persistence in $\sigma$. As Eraker et al. (2003) and others have shown, stochastic volatility is an important component of the return process and should be formally integrated into the AAJD specification. Then, one can formally test the conjecture that there is no persistence in $\sigma$, after adjustments for jumps. With the Bayesian estimation approach, it may be simpler to determine whether stochastic volatility remains important when the jump components of return process has a more complex structure like the AAJD specifications. This is an interesting future line of research we plan to pursue.

There are other interesting directions to extend this work. As a starting point, other estimation techniques, such as the generalized method of moments and its variants may be utilized to obtain estimates of the AAJD parameters. Time-varying jump intensities, as proposed by Andersen et al. (2002), offers another way to enhance the AAJD specification. Finally, integrating the jump detection techniques suggested by Lee and Hanng (2010) and Lee (2012) into the Bayesian estimation of the AAJD represent a challenging but potentially valuable direction for future research.

Acknowledgments

We thank Sanjiv Das, Garland Durham, Xin Guo, Maciej Kostrewski, Stathis Tompaidis, Yong Zeng, the editor and an anonymous referee for their valuable comments and suggestions. We retain full duopoly rights over any error or omission.

Appendix A

The appendix provides information about our choice of candidate distributions and the specifics of our sampling procedures. The majority of our distributional choice are guided by previous studies, particularly Eraker et al. (2003), Jacquier et al. (2007), Johannes and Polson (2009) and Jacquier and Polson (2011).

A.1. Sampling the latent variable

Sampling the latent variable is the most important part of our Bayesian estimation method. This sampling step constitutes the “ownership” step of the algorithm,
Figure 4. PB JD distributions for S&P-500 and NASDAQ (bull, bear, and combined periods) using the MCMC parameter estimates from Table 7.

Bear period estimates. The good news component accelerates the drift of the return process.

The parameter estimates for the alternative A AJD specifications should be nearly identical when all the underlying distributional assumptions are satisfied, as shown in Sec. 2.3. In our simulation exercise, the noted assumptions are satisfied and as a result the parameter estimates reported in Table 3 are nearly identical. When utilizing actual return data, our analysis shows that the majority of the estimated parameters found in Tables 5–7 are notably different under PB JD and DE JD. This divergence of the parameter estimates suggests at least one violation of the underlying assumptions: the independence of the returns, parameters, and jump events; their distributional forms; and the equivalency of the mixture model components.

5. Conclusion

In this study, we develop a Bayesian estimation method for estimating the parameters of the Asymmetric Affine Jump-Diffusion models. We provide an empirical assessment of these models using daily returns for the S&P 500 and the NASDAQ indexes, as well as specific stocks. We complete our analysis by providing a comparison of the estimated parameters under Bayesian estimation and MLE. We find that, in general, the Bayesian estimates are consistent with the MLE estimates. However, our Bayesian estimation approach is computationally more efficient and yields smaller standard errors.

We find that the A AJD models are consistent with the empirical features of return processes. We study the behavior of the models’ parameters during different market epochs (Bull and Bear periods) and find that with the introduction of the jump components, the volatility component due to the GBM part of the return process (σ) is constant across epochs. This finding is in contrast to the stochastic volatility models that show persistence in σ. As Eraker et al. (2003) and others have shown, stochastic volatility is an important component of the return process and should be formally integrated into the A AJD specification. Then, one can formally test the conjecture that there is no persistence in σ, after adjustments for jumps. With the Bayesian estimation approach, it may be simpler to determine whether stochastic volatility remains important when the jump components of return process has a more complex structure like the A AJD specifications. This is an interesting future line of research we plan to pursue.

There are other interesting directions to extend this work. As a starting point, other estimation techniques, such as the generalized method of moments and its variants may be utilized to obtain estimates of the A AJD parameters. Time-varying jump intensities, as proposed by Andersen et al. (2002), offers another way to enhance the A AJD specification. Finally, integrating the jump detection techniques suggested by Lee and Hanng (2010) and Lee (2012) into the Bayesian estimation of the A AJD represent a challenging but potentially valuable direction for future research.

Acknowledgments

We thank Sanjiv Das, Garland Durham, Xin Guo, Maciej Kostrewski, Stathis Tompaidis, Yong Zeng, the editor and an anonymous referee for their valuable comments and suggestions. We retain full duopoly rights over any error or omission.

Appendix A

The appendix provides information about our choice of candidate distributions and the specifics of our sampling procedures. The majority of our distributional choice are guided by previous studies, particularly Eraker et al. (2003), Jacquier et al. (2007), Johannes and Polson (2009) and Jacquier and Polson (2011).

A.1. Sampling the latent variable

Sampling the latent variable is the most important part of our Bayesian estimation method. This sampling step constitutes the “ownership” step of the algorithm.
temporarily defining which mixture model component generates each of the observed returns (i.e., the no, up-, or down-jump component of the approximate trinomial mixture model). Let \( I \in \{-1, 0, 1\} \) indicate a down-, no, or an up-jump events respectively at times \( i = 1, \ldots, M \). The conditional probability mass function of \( I_i \) is

\[
P(I_i = -1 | \theta_{PRID}, r_i) = \frac{\lambda_d^2 f_0 (r_i)}{\lambda_u^2 f_0 (r_i) + (1 - \lambda_u^2 - \lambda_d^2) f_0 (r_i) + \lambda_u^2 f_0 (r_i)},
\]

\[
P(I_i = 0 | \theta_{PRID}, r_i) = \frac{(1 - \lambda_u^2 - \lambda_d^2) f_0 (r_i) + \lambda_u^2 f_0 (r_i)}{\lambda_u^2 f_0 (r_i) + (1 - \lambda_u^2 - \lambda_d^2) f_0 (r_i) + \lambda_u^2 f_0 (r_i)},
\]

\[
P(I_i = 1 | \theta_{PRID}, r_i) = \frac{\lambda_u^2 f_0 (r_i)}{\lambda_u^2 f_0 (r_i) + (1 - \lambda_u^2 - \lambda_d^2) f_0 (r_i) + \lambda_u^2 f_0 (r_i)}.
\]

By our construction of \( \lambda_u^2 \) and \( \lambda_d^2 \), the probability mass function for \( I_i \) is proper (the weights sum to one) and can be easily sampled from. Conditional on sampling an up-jump event \( I_i = 1 \), or down-jump event \( I_i = -1 \), we must also sample the jump amplitudes from the Pareto or Beta distributions.

Let \( Y^u_i \) and \( Y^d_i \) be the conditional up- and down-jump amplitudes for return \( r_i \). Sampling \( Y^u_i \) and \( Y^d_i \) is a challenging step, because we are not able to draw samples from the posterior directly. The posterior distribution of the conditional jumps are

\[
f(Y^u_i | \mu, \sigma^2, I_i = 1, \eta_u, r_i) \propto e^{-\eta_u Y^u_i} \frac{(r_i - \mu + \frac{\sigma^2}{2})^2}{2 \sigma^2},
\]

\[
f(Y^d_i | \mu, \sigma^2, I_i = -1, \eta_d, r_i) \propto e^{-\eta_d Y^d_i} \frac{(r_i - \mu + \frac{\sigma^2}{2})^2}{2 \sigma^2}.
\]

There does not exist a prior distribution for \( f(\eta_u) \) and \( f(\eta_d) \) that will yield recognizable proper conditionals, \( f(Y^j_i | \cdot, j = u, d \) from which to draw samples of \( Y^j_i \). Given this circumstance, it is possible to rely on MH sampling to draw samples of \( Y^j_i \). However, it is not clear what the optimal proposal distribution should be (Johannes and Polson, 2009, p. 16). Two reasonable choices are the Gamma or Normal distributions. After experimenting with both candidates, we found that the Gamma distribution is a poor choice leading to very low acceptance rates, whereas the Normal distribution yields the highest acceptance rates ranging from 83% to 97%. However, the MH sampling of jumps is computationally expensive for returns with large numbers of up- and down-jump events. In practice, we found it efficient and satisfactory to directly sample from the Normal distribution without the MH step. We therefore sample the up- and down-jumps using

- Conditional on \( I_i = 1 \): Sample \( Y^u_i \) from \( N(r_i - \mu + \frac{\sigma^2}{2}, \sigma^2) \).
- Conditional on \( I_i = -1 \): Sample \( Y^d_i \) from \( N(r_i - \mu + \frac{\sigma^2}{2}, \sigma^2) \).

Sampling the latent variable results in a 4-tuple of “complete information”, \( R_i = (r_i, I_i, Y^u_i, Y^d_i) \), for \( i = 1, \ldots, M \). The remaining sampling steps for the elements of \( \theta_{PRID} \) will have posterior distributions which rely on the following summary statistics

\[
U = \sum_{i=1}^{M} Y^u_i, \quad n = \sum_{i=1}^{M} |I_i| = 1, \quad D = -\sum_{i=1}^{M} Y^d_i, \quad m = \sum_{i=1}^{M} |I_i| = -1.
\]

where \( U \) and \( D \) are the cumulative sum of the sampled up- and down-jumps, and \( n \) and \( m \) are the number of sampled up- and down-jumps. The next component of the GS method is to sample the elements of \( \theta_{PRID} \) conditional on the complete data. Let \( R \) be the complete set of all 4-tuples, \( R = (R_1, \ldots, R_M) \). The complete likelihood is

\[
L(\theta_{PRID} | R) \propto \prod_{i=1}^{M} \left\{ \lambda_u^2 e^{-\eta_u Y^u_i} \frac{(r_i - \mu + \frac{\sigma^2}{2})^2}{2 \sigma^2} \right\} \times \prod_{i=1}^{M} \left\{ \lambda_d^2 e^{-\eta_d Y^d_i} \frac{(r_i - \mu + \frac{\sigma^2}{2})^2}{2 \sigma^2} \right\} \times \prod_{i=1}^{M} \left\{ (1 - \lambda_u - \lambda_d) e^{-\frac{1}{2} \sigma^2} \right\}.
\]

The subscript in each product, \( I_i \), indicates the product for up-, down-, and no-jump events. We use the complete likelihood to iteratively sample the parameters in \( \theta_{PRID} \) for the remainder of the GS.

A.2. Sampling the jump event parameters

The arrival rate of the up- and the down-jumps is governed by the Poisson processes with generating rates \( \lambda_u \) and \( \lambda_d \). We assume an uninformative prior \( \pi(\lambda_u, \lambda_d) \propto 1 \), \( \lambda_u \in [0, 1] \), \( \lambda_d \in [0, 1] \), and only require information about the cumulative number of up and down-jumps (\( U \) and \( D \) above). This information is contained in the summary statistics of \( R \). The posterior distribution of \( (\lambda_u, \lambda_d) \) conditional on \( R \) is

\[
f(\lambda_u, \lambda_d | R) \propto \prod_{i=1}^{M} \lambda_u^{i} \prod_{i=1}^{M} \lambda_d^{i} \prod_{j=1}^{M} (1 - \lambda_u - \lambda_d) \propto (\lambda_u^{i})^{n-1} (\lambda_d^{i})^{m-1} (1 - \lambda_u - \lambda_d)^{M-n-m-i-1}
\]

which is the kernel of a Dirichlet distribution, hence we sample \( (\lambda_u, \lambda_d) \sim \text{Dirichlet}(n+1, m+1, M-n-m+1) \).
temporarily defining which mixture model component generates each of the observed returns (i.e., the no, up-, or down-jump component of the approximate trinomial mixture model). Let \( l \in \{ -1, 0, 1 \} \) indicate a down-, no, and an up-jump events respectively at times \( i = 1, \ldots, M \). The conditional probability mass function of \( l_i \) is

\[
P(l_i = -1 | \theta_{\text{PRJD}}, r_i) = \frac{\lambda_u^2 f_{01}(r_i) + (1 - \lambda_u - \lambda_d) f_{00}(r_i) + \lambda_d f_{10}(r_i)}{\lambda_u^2 f_{01}(r_i) + (1 - \lambda_u - \lambda_d) f_{00}(r_i) + \lambda_d f_{10}(r_i)},
\]

\[
P(l_i = 0 | \theta_{\text{PRJD}}, r_i) = \frac{(1 - \lambda_u - \lambda_d) f_{00}(r_i)}{\lambda_u^2 f_{01}(r_i) + (1 - \lambda_u - \lambda_d) f_{00}(r_i) + \lambda_d f_{10}(r_i)},
\]

\[
P(l_i = 1 | \theta_{\text{PRJD}}, r_i) = \frac{\lambda_d^2 f_{11}(r_i)}{\lambda_u^2 f_{01}(r_i) + (1 - \lambda_u - \lambda_d) f_{00}(r_i) + \lambda_d f_{10}(r_i)}.
\]

By our construction of \( \lambda_u^2 \) and \( \lambda_d^2 \), the probability mass function for \( l_i \) is proper (the weights sum to one) and can be easily sampled from. Conditional on sampling an up-jump event \( l_i = 1 \), or down-jump event \( l_i = -1 \), we must also sample the jump amplitudes from the Pareto or Beta distributions.

Let \( Y_i^u \) and \( Y_i^d \) be the conditional up- and down-jump amplitudes for return \( r_i \). Sampling \( Y_i^u \) and \( Y_i^d \) is a challenging step, because we are not able to draw samples from the posterior distribution of the conditional jumps.

The conditional probability mass function for \( l_i \) is

\[
f(Y_i^u | \mu, \sigma^2, l_i = 1, \eta_u, r_i) \propto e^{-\eta_u(Y_i^u - \mu)^2} - (r_i - \mu + \sigma^2)^{\eta_u},
\]

\[
f(Y_i^d | \mu, \sigma^2, l_i = -1, \eta_d, r_i) \propto e^{-\eta_d(Y_i^d - \mu)^2} - (r_i - \mu + \sigma^2)^{\eta_d}.
\]

There does not exist a prior distribution for \( f(\eta_u) \) and \( f(\eta_d) \) that will yield recognizable proper conditionals, \( f(Y_i^u | \cdot) \), \( j = u, d \), from which to draw samples of \( Y_i^j \).

Given this circumstance, it is possible to rely on MH sampling to draw samples of \( Y_i^j \). However, it is not clear what the optimal proposal distribution should be (Johannes and Polson, 2009, p. 16). Two reasonable choices are the Gamma or Normal distributions. After experimenting with both candidates, we found that the Gamma distribution is a poor choice leading to very low acceptance rates, whereas the Normal distribution yields the highest acceptance rates ranging from 83% to 97%.

However, the MH sampling of jumps is computationally expensive for returns with large numbers of up- and down-jump events. In practice, we found it efficient and satisfactory to directly sample from the Normal distribution without the MH step. We therefore sample the up- and down-jumps using

- **Conditional on** \( l_i = 1 \), \( Y_i^u = 0 \) **and sample** \( Y_i^u \) **from** \( N(r_i - \mu + \frac{\sigma^2}{2}, \sigma^2) \).
- **Conditional on** \( l_i = -1 \), \( Y_i^u = 0 \) **and sample** \( Y_i^d \) **from** \( N(r_i - \mu + \frac{\sigma^2}{2}, \sigma^2) \).

Sampling the latent variable results in a 4-tuple of “complete information”, \( R_i = (r_i, l_i, Y_i^u, Y_i^d) \), for \( i = 1, \ldots, M \). The remaining sampling steps for the elements of \( \theta_{\text{PRJD}} \) will have posterior distributions which rely on the following summary statistics

\[
U = \sum_{i=1}^{M} Y_i^u, \quad n = \sum_{i=1}^{M} I(l_i = 1), \quad D = -\sum_{i=1}^{M} Y_i^d, \quad m = \sum_{i=1}^{M} I(l_i = -1),
\]

where \( U \) and \( D \) are the cumulative sum of the sampled up- and down-jumps, and \( n \) and \( m \) are the number of sampled up- and down-jumps. The next component of the GS method is to sample the elements of \( \theta_{\text{PRJD}} \) conditional on the complete data. Let \( \mathbf{R} \) be the complete set of all 4-tuples, \( \mathbf{R} = (R_1, \ldots, R_M) \). The complete likelihood is

\[
L(\theta_{\text{PRJD}} | \mathbf{R}) \propto \prod_{i=1}^{M} \left\{ \lambda_u^2 e^{-\eta_u(Y_i^u - \mu)^2} + \lambda_d^2 e^{-\eta_d(Y_i^d - \mu)^2} \right\} \\
\quad \times \prod_{i=1}^{M} \left\{ \lambda_u^2 e^{-\eta_u(Y_i^u - \mu)^2} + \lambda_d^2 e^{-\eta_d(Y_i^d - \mu)^2} \right\} \\
\quad \times \prod_{i=1}^{M} \left( 1 - \lambda_u - \lambda_d \right) e^{-\frac{(r_i - \mu)^2}{\sigma^2}}.
\]

The subscript in each product, \( I_i \), indicates the product for up-, down-, and no-jump events. We use the complete likelihood to iteratively sample the parameters in \( \theta_{\text{PRJD}} \) for the remainder of the GS.

**A.2. Sampling the jump event parameters**

The arrival rate of the up- and the down-jumps is governed by the Poisson processes with generating rates \( \lambda_u \) and \( \lambda_d \). We assume an uninformative prior \( \pi(\lambda_u, \lambda_d) \propto 1_{\lambda_u, \lambda_d \in [0, 1]} \), and only require information about the cumulative number of up and down-jumps (\( U \) and \( D \) above). This information is contained in the summary statistics of \( \mathbf{R} \). The posterior distribution of \( (\lambda_u, \lambda_d) \) conditional on \( \mathbf{R} \) is

\[
f(\lambda_u, \lambda_d | \mathbf{R}) \propto \prod_{i=1}^{M} \lambda_u^2 \prod_{i=1}^{M} \lambda_d^2 \prod_{i=1}^{M} (1 - \lambda_u - \lambda_d) \\
\propto (\lambda_u^2)^{n+1} (\lambda_d^2)^{m+1} (1 - \lambda_u - \lambda_d)^{M-n-m+1}
\]

which is the kernel of a Dirichlet distribution, hence we sample \( (\lambda_u, \lambda_d) \sim \text{Dirichlet}(n + 1, m + 1, M - n - m + 1) \).
A.3. Sampling the jump size parameters

The parameters that govern the generating process for the size of the up- and down-jumps are $\eta_u$ and $\eta_d$ respectively. All of the information needed to develop the posterior distribution of $\eta_u$ is contained in the summary statistics of $\mathbf{R}$. The posterior distribution of $\eta_u$ is

$$f(\eta_u | \mathbf{R}) \propto \prod_{t=1}^M \eta_u e^{-\eta_u r_t - \frac{1}{2\sigma^2} (r_t - Y^u - \mu + \frac{\sigma^2}{2})^2}$$

which is the kernel of a Gamma distribution, hence we sample $\eta_u \sim \Gamma(n + 1, U)$. The posterior distribution for $\eta_d$ is similar, hence we sample $\eta_d \sim \Gamma(m + 1, D)$.

A.4. Sampling the drift parameter

The posterior distribution of parameter $\mu$ is comparable to the posterior distribution when the mixture model is the usual mixture of Normal distributions. In this case, terms involving $\eta_u$ and $\eta_d$ act as normalizing constants and do not contribute any information to the posterior distribution of $\mu$. The quadratic, Gaussian kernel of each component adjusts the observed $r_t$ by $Y^u$ or $Y^d$. Only one of the jump sizes will be non-zero by construction. So, we can re-write the posterior distribution to reflect this fact. The posterior distribution of $\mu$ is

$$f(\mu | \mathbf{R}, \sigma^2) \propto \prod_{t=1}^M e^{-\frac{1}{2\sigma^2} (r_t - Y^u - \mu + \frac{\sigma^2}{2})^2}$$

which is the kernel of a Normal distribution. Accordingly, we sample $\mu \sim N \left( \frac{\sum_{t=1}^M r_t - Y^u - Y^d}{M}, \frac{\sigma^2}{M} \right)$.

A.5. Sampling the volatility parameter

The posterior distribution of $\sigma^2$ is given below.

$$f(\sigma^2 | \mathbf{R}, \mu) \propto \frac{1}{(\sigma^2)^{\frac{M+n-1}{2}}} e^{-\frac{1}{2\sigma^2} \sum_{t=1}^M (r_t - Y^u - Y^d - \mu + \frac{\sigma^2}{2})^2}$$

This is an unrecognizable distribution (non-standard) and we are unable to directly sample from the posterior. Subsequently, we rely on a MH accept-reject scheme. However, the posterior closely resembles the kernel of an inverse Gamma distribution where the second parameter, located in the exponent, actually depends on $\sigma^2$. Hence, we sample $\sigma^2$ by drawing from an inverse Gamma distribution. Then, we use the following MH accept-reject algorithm

- Sample $\sigma^2 \sim \Gamma^{-1} \left( \frac{M}{2} - 1, \sum_{t=1}^M \left( r_t - Y^u - Y^d - \mu + \frac{\sigma^2}{2} \right)^2 \right)$.
- Calculate the probability of accepting the candidate sample
  $$p^* = \frac{\pi(\sigma^2 | \mu, \mathbf{R}) \pi_1(\sigma^2 | \sigma^2, \mu, \mathbf{R})}{\pi(\sigma^2 | \mu, \mathbf{R}) \pi_1(\sigma^2 | \sigma^2, \mu, \mathbf{R})},$$
  where $\pi_1(\cdot)$ is the inverse Gamma distribution using the conditioned value of $\sigma^2$ or $\sigma^2$ in the construction of the exponent. Let $p^* = \min \{p^*, 1\}$.
- Generate $p \sim \text{Uniform}(0, 1)$ and if $p^* > p$, accept $\sigma^2$.

Using simulated and real data, this MH scheme for sampling $\sigma^2$ is efficient and accepts 80%-90% of the candidate samples.

References


A.3. Sampling the jump size parameters

The parameters that govern the generating process for the size of the up- and down-jumps are $\eta_u$ and $\eta_d$ respectively. All of the information needed to develop the posterior distribution of $\eta_u$ is contained in the summary statistics of $R$. The posterior distribution of $\eta_u$ is

$$f(\eta_u | R) \propto \prod_{t=1}^{M} \eta_u e^{-\eta_u (r_t - Y_t^u - \mu + \sigma^2/2)} \propto \eta_u^{M+1} e^{-\eta_u U},$$

which is the kernel of a Gamma distribution; hence we sample $\eta_u \sim \Gamma(n+1, U)$. The posterior distribution for $\eta_d$ is similar, hence we sample $\eta_d \sim \Gamma(m+1, D)$.

A.4. Sampling the drift parameter

The posterior distribution of parameter $\mu$ is comparable to the posterior distribution when the mixture model is the usual mixture of Normal distributions. In this case, terms involving $\eta_u$ and $\eta_d$ act as normalizing constants and do not contribute any information to the posterior distribution of $\mu$. The quadratic, Gaussian kernel of each component adjusts the observed $r_t$ by $Y_t^u$ or $Y_t^d$. Only one of the jump sizes will be non-zero by construction. So, we can re-write the posterior distribution to reflect this fact. The posterior distribution of $\mu$ is

$$f(\mu | R, \sigma^2) \propto \prod_{t=1}^{M} e^{-\frac{1}{2\sigma^2} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2} \propto \prod_{t=1}^{M} e^{-\frac{1}{2\sigma^2} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2} \prod_{t=1}^{M} e^{-\frac{1}{2\sigma^2} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2} \propto e^{-\frac{1}{2\sigma^2} \sum_{t=1}^{M} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2},$$

which is the kernel of a Normal distribution. Accordingly, we sample $\mu \sim N \left( \frac{\sum_{t=1}^{M} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)}{M} \right)$.

A.5. Sampling the volatility parameter

The posterior distribution of $\sigma^2$ is given below.

$$f(\sigma^2 | R, \mu) \propto \frac{1}{(2\pi)^{(M+1)/2}} e^{-\frac{1}{2} \sum_{t=1}^{M} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2}.$$

This is an unrecognizable distribution (non-standard) and we are unable to directly sample from the posterior. Subsequently, we rely on a MH accept-reject scheme. However, the posterior closely resembles the kernel of an inverse Gamma distribution where the second parameter, located in the exponent, actually depends on $\sigma^2$. Hence, we sample $\sigma^2$ by drawing from an inverse Gamma distribution. Then, we use the following MH accept-reject algorithm

- Sample $\sigma^2 \sim \Gamma^{-1} \left( \frac{M}{2} - 1, \frac{1}{2} \sum_{t=1}^{M} (r_t - Y_t^u - Y_t^d - \mu + \sigma^2/2)^2 \right)$.

- Calculate the probability of accepting the candidate sample

$$p^* = \frac{\pi(\sigma^2 | R, \mu)}{\pi(\sigma^2 | R, \mu)} \frac{\pi(\sigma^2 | R, \mu)}{\pi(\sigma^2 | R, \mu)},$$

where $\pi(\cdot | \cdot)$ is the inverse Gamma distribution using the conditioned value of $\sigma^2$ or $\sigma^2$ in the construction of the exponent. Let $p^* = \min\{p^*, 1\}$.

- Generate $p \sim \text{Uniform}(0, 1)$ and if $p^* > p$, accept $\sigma^2$.

Using simulated and real data, this MH scheme for sampling $\sigma^2$ is efficient and accepts 80%-90% of the candidate samples.

References


