Do Stock Returns Follow a Finite Variance Distribution?[*]

Qi-Man Shao

Department of Mathematics
University of Oregon Eugene, Oregon, USA

Hao Yu

Department of Statistical and Actuarial Sciences
The University of Western Ontario London, Canada

and

Jun Yu

Department of Economics
The University of Auckland
Private Bag 92019 Auckland, New Zealand

In this paper we propose a test statistic to discriminate between models with finite variance and models with infinite variance. The test statistic is the ratio of the sample standard deviation and the sample interquartile range. Both asymptotic and finite sample properties of the test statistic are discussed. We show that the test has good power properties against infinite-variance distributions and has small size distortions in finite samples. The statistic is applied to compare the competing models for S&P 500 index returns. Our test cannot reject most distributions with finite variance for both a pre-crash sample and a post-crash sample, and hence supports the literature. However, for a sample including crash days, our test suggests that the finite-variance distributions must be rejected. © 2001 Peking University Press

Key Words: Stock Returns; Infinite Variance; Interquartile Range.
JEL Classification Numbers: C12, C14, C15.

1. INTRODUCTION

Modelling stock returns has been a very interesting topic for a long time. One reason is that some important models in financial theory critically rely on the distribution form for the returns of underlying stocks, such as mean-variance portfolio theory, capital asset pricing models, and prices of derivative securities. In the search for satisfactory descriptive models of stock returns, many distributions have been tried and some new distributions have been created over past several decades. Tucker (1992) categorizes the candidate models by independent and linear processes and time-dependent processes where both independence and linearity are relaxed. Time-dependent processes have been successful in the modelling of financial time series including daily stock returns. However, as Tucker (1992) claims, the descriptive validity of competing time-independent model still remains unresolved and there has been much debate among them. All time-independent models can be divided by two families. One family has finite-variance. Examples include the normal distribution proposed by Osborne (1959), the Student t distribution by Blattberg and Gonedes (1974), the mixture of normals (MN) by Kon (1984), the compound log-normal and normal (LN) distribution by Clark (1973), the mixed diffusion-jump (MDJ) model by Press (1967) and more recent one, the Weibull distribution by Mittnik and Rachev (1993). The other family has infinite-variance, such as the stable distribution proposed by Mandlebrot (1963) and Fama (1965).

The stable distribution has been appreciated as a possible alternative to describe the stock returns for both statistical and economic reasons. Interesting statistical properties include: (1) only stable laws have domains of attraction (generalized central limit theorem); (2) stable distributions belong to their domain of attraction (stability). Economically speaking, the stable distribution has unbounded variation, and hence is consistent with continuous-time equilibrium in competitive markets (see McCulloch (1978)).

Despite these appealing properties, the stable distribution is less commonly used today. It has fallen out of favor, partly because of the difficulties involved in theoretical modelling; standard financial theory, such as the option theory, almost always requires finite variance of returns. Furthermore, evidence has been found against the stable distribution. Firstly, using the conventional likelihood ratio test, Blattberg and Gonedes (1974) found that the Student t distribution has greater descriptive validity than the symmetric stable distribution, and Tucker (1992) found that finite-variance models outperform the asymmetric stable distribution. Using the Komogorov-Smirnov test, Mittnik and Rachev (1993) found that the Weibull distribution is the most suitable candidate. Secondly, when the tail behavior

was investigated, Akgiray and Booth (1987) found that the tails of stable distribution are too thick to fit the empirical data. Thirdly, Lau, Lau and Wingender (1990) found that as the sample size gets big the sample high moments seems to converge while the stable distribution implies that sample high moments should blow up rapidly. Finally, the evidence provided by Blattberg and Gonedes (1974) indicates that the distribution of monthly returns conforms well to the normal distribution, while the stable distribution implies that long horizon (for example, monthly) returns will be just as non-normal as short-horizon (for example, daily) returns.

The purpose of this paper is to re-examine the descriptive power of the finite-variance distribution family and the infinite-variance distribution family as models of daily stock returns. However, instead of using overall goodness of fit testing methodology or model selection criteria, we concentrate on studying the variance behavior for chosen distribution families. To be more specific, we propose a test statistic to distinguish finite-variance families against infinite-variance families for stock returns. Particular attention is paid to the variance due to two reasons. Firstly, as far as the variance is concerned, an infinite-variance model is fundamentally riskier than a finite-variance model. Secondly and more importantly, many financial models critically depend on the assumption on the second moment. Examples include the capital asset pricing model (CAPM) and the Black-Scholes option price model. As a result, finite variance and infinite variance could have very different implications for theoretical and empirical analysis. Unfortunately, testing for finite variance or infinite variance based on a sample without choosing specific distribution families will probably never be possible since such a test could have no power. Instead of directing the test on variance itself, we test a specific finite-variance model against a specific infinite-variance model.

The paper is organized as follows. The next section introduces the test statistic, motivates the intuition behind it, and obtains the statistical properties of it. Section 3 briefly summarizes the candidate models of the stock returns, including finite-variance family and infinite-variance family. The proposed statistic is used to discriminate between these two families. Section 4 tests the finite-variance models proposed in the literature and examines the finite sample properties of the test statistic in Monte Carlo studies. Section 5 describes how the basic framework can be extended to time-dependent processes, such as ARCH-type models and stochastic volatility models. Section 6 concludes. All the proofs are collected in Appendix.

2. PROPOSED STATISTIC AND ITS PROPERTIES

DuMoucher (1973) states that if a sample has a standard deviation many times as large as the interquartile range, the Data Generating Process (DGP) could have an infinite variance. However, he does not give a statistical analysis to indicate when the DGP has an infinite variance. Despite this we find that his statement is quite intuitive and study along this line serves our purpose to distinguish finite variance models and infinite variance models. In other words, the statistical properties of the relative magnitude of the sample standard deviation with the sample interquartile range should be investigated.

Suppose $\{X_i\}_{i=1}^n$ be an iid sequence of observations with common distribution function $F(x)$, common density function $f(x)$, mean μ and variance σ^2. Let

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \hat{\mu})^2$$

be the sample variance, where $\hat{\mu}$ is the sample mean. Denote the quantile process by $Q_n(t)$ (see Chapter 6, Csörgő and Horváth (1993)), that is,

$$Q_n(t) = \begin{cases} X_{1,n}, & \text{if } t = 0 \\ X_{k,n}, & \text{if } (k-1)/n < t \le k/n, \ 1 \le k \le n \end{cases}$$

where $X_{1,n} \le X_{2,n} \le \cdots \le X_{n,n}$ are order statistic. The proposed test statistic is then defined as,

$$T_n(\theta_0) = \frac{s_n}{Q_n(1-\theta_0) - Q_n(\theta_0)}, \tag{1}$$

where $0 < \theta_0 < 0.5$. Hence the denominator is the θ_0-quartile range and indeed the interquartile range when $\theta_0 = 0.25$. Therefore, $T_n(0.25)$ is basically the ratio of the sample standard deviation and sample interquartile range.

It seems natural to use sample variance or sample standard deviation to discriminate between finite-variance distributions and infinite-variance distributions. Unfortunately, the power based on the sample variance or sample standard deviation may not be good since a finite-variance distribution can generate a larger sample variance than a distribution with an infinite variance can even when the sample size is large. By taking the ratio of two dispersion parameters, however, the proposed test can be standardized or at least reduced the dispersion of any finite-variance distribution. This is because when the true DGP has a finite variance, less observations come from the tails and hence $s_n \to \sigma$. Since both $Q_n(0.75)$ and $Q_n(0.25)$ are finite for any n, T_n converges to a finite number as $n \to \infty$. Consequently, it is reasonable to believe that a large T_n comes from a DGP with

infinite variance rather than a DGP with finite variance. On the other hand, if the true DGP has an infinite variance, more observations must be from the tails and $s_n \to \infty$ as $n \to \infty$. This implies the unboundness of $T_n(0.25)$. Thus we set up the hypothesis as the following,

$$\begin{cases} H_0 : \text{DGP is a } certain \text{ finite variance distribution,} \\ H_1 : \text{DGP has an infinite variance.} \end{cases} \tag{2}$$

If H_0 is rejected, the model in H_0 should not be used as a candidate model.

In this subsection we assume $X_1, X_2, ..., X_n$ to be iid random variables. The properties of T_n are established in this section. Their proofs are found in Appendix.

PROPOSITION 2.1. *T_n is invariant for a scale-location family.*

This is an indeed appealing property. For a scale-location family, no matter how big the scale is, the expectation of the statistic always takes the same value. In other words, if we think of T_n as a measure of risk, the risk associated with a scale-location family is a constant. Because of this property, any scale-location family can be treated as one model.

PROPOSITION 2.2. *If $\sigma^2 < +\infty$, and $Q(t)$ is continuous at θ_0 and $1 - \theta_0$, then*

$$T_n \to T = \frac{\sigma}{q_1 - q_0} < \infty \quad a.s., \tag{3}$$

where $q_1 = Q(1 - \theta_0)$, $q_0 = Q(\theta_0)$ with $Q(t) = inf\{x : F(x) \geq t\}$.

This result is very intuitive since it says that T_n converges almost surely to its population counterpart. According to this proposition, if the model in H_0 has good descriptive power, it must yield a value of T which is close to the empirical T_n.

PROPOSITION 2.3. *Assume that*
(i) $f(q_1) > 0$, $f(q_0) > 0$.
(ii) $f(x)$ is continuous in a neighborhood of q_1 and q_0.
If $E|X_1|^4 < \infty$, then,

$$\sqrt{n}(T_n - T) \xrightarrow{d.} N(0, \Sigma^2), \tag{4}$$

that is,

$$T_n \overset{a}{\sim} N(T, \frac{\Sigma^2}{n}), \tag{5}$$

where

$$\begin{aligned}
\Sigma^2 &= (q_1 - q_0)^{-4} E \left\{ \frac{(q_1 - q_0)}{2\sigma}((X_1 - \mu)^2 - \sigma^2) - \frac{\sigma(I\{X_1 > q_1\} - \theta_0)}{f(q_1)} \right. \\
&\left. - \frac{\sigma(I\{X_1 \leq q_0\} - \theta_0)}{f(q_0)} \right\}^2,
\end{aligned} \tag{6}$$

The asymptotic distribution in Proposition (2.3) is the main result of the paper since it provides the basis for a large sample test procedure. Although T_n is invariant for a scale-location family, it is important to note that in general both T and Σ^2 depend on f and hence H_0. Therefore in general our statistic cannot be used to test the following hypothesis,

$$\left\{ \begin{array}{ll} H_0 : & \text{DGP is } any \text{ finite variance distribution,} \\ H_1 : & \text{DGP has an infinite variance.} \end{array} \right. \tag{7}$$

Instead T_n can be used as a non-nested test of a specific finite variance distribution against an infinite variance distribution, where the distribution along with the parameters in H_0 have to be specified except for the scale and location parameters. Hence the test bears some resemblance to misspecification test statistics.

PROPOSITION 2.4. *Under assumptions of the proposition (2.3), if f is symmetric, then*

$$\Sigma^2 = \frac{K - 1}{4a^2} + \frac{2\theta_0(1 - 2\theta_0)}{a^4 b^2} + \frac{\theta_0 c_1 - (1 - \theta_0 c_2)}{2a^3 b}, \tag{8}$$

where K is the kurtosis of X, $a = \frac{1}{T}$, $b = \sigma f(q)$, $c_1 = \int_{-\infty}^{q_1} (\frac{x-\mu}{\sigma})^2 f(x)\,dx$, and $c_2 = \int_{q_1}^{\infty} (\frac{x-\mu}{\sigma})^2 f(x)\,dx$.

3. CANDIDATE MODELS FOR DAILY STOCK RETURNS

In this section we introduce the most well-known time-independent models for daily stock returns, briefly review the properties of the candidate models, and discuss the relevant estimation method and numerical algorithm if necessary. In the finite-variance family, the normal distribution, the Student t distribution, the mixture of normals, mixed diffusion-jump model, the compound log-normal and normal model, and the Weibull distribution are presented, while the stable distribution represents infinite-variance family. Note that we are not able to cover all the candidates in

the literature since some new distributions are still being created. However, we believe that our test can be used in the same way for these distributions.

3.1. Normal Distribution

The first model used in the literature to describe daily stock returns is the normal distribution proposed by Bachelier (1900) and extended by Osborne (1959). Black and Scholes (1973) provide a formula to price an option assuming the normality of underlying asset. Although the assumption of normality greatly simplified the theoretical modelling, many empirical studies have shown evidence against it (see Blattberg and Gonedes (1974), Clark (1973), Kon (1984) and Niederhoffer and Osborne (1966)). For example, empirical daily stock returns exhibit fatter tails and greater kurtosis than the normal distribution. Despite this evidence, in this paper we still choose it as a competing model because we want to check the validity of this assumption by using our test statistic. Observe that all moments for the normal distribution exist and the kurtosis for the normal family is three. Furthermore, since the normal distribution belongs to a scale-location family, T_n is invariant to both μ and σ^2 and hence parameter estimation is not necessary.

3.2. Student Distribution

The Student distribution is first proposed to model the stock returns by Blattberg and Gonedes (1974). Its density is,

$$g(x) = \frac{\Gamma[(1+\nu)/2]\nu^{\nu/2}\sqrt{H}}{\Gamma(1/2)\Gamma(\nu/2)}[\nu + H(x-m)^2]^{-(\nu+1)/2}, \qquad (9)$$

where $\nu \geq 2$, and H, m, ν are the scale parameter, location parameter, and degrees-of-freedom parameter. Therefore, T_n is invariant to both H and m, but depends on ν. Furthermore, when $\nu > 4$ the Student distribution has a finite fourth moment and hence the C.L.T. in Section 2 can be applied. The model is estimated by the maximum likelihood method using a Quasi-Newton algorithm.[1]

3.3. Mixture of Normals

Kon (1984) proposes to use the mixture of normals to model stock returns, i.e., the stock return X_i come from $N(\mu_j, \sigma_j^2)$ with probability α_j and $\alpha_1 + \cdots + \alpha_k = 1$. A characteristic of this model is that it can capture

---

[1]With little effort, we can show that $\hat{m}_{ML} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Therefore, only parameters H and ν are considered in the numerical algorithm.

the structural change. The density function is,

$$g(x) = \sum_{j=1}^{k} \alpha_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left\{-\frac{(x-\mu_j)^2}{2\sigma_j^2}\right\}. \qquad (10)$$

All moments exist for the mixture of normals. However, in this paper we only consider the mixture of two normals due to two reasons. Firstly, Tucker (1992) found the mixture of two normals has the greatest descriptive power among the family of the mixture of normals. Secondly, we want to avoid a model with too many parameters. The parameters for the mixture of two normals are $\alpha, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and T_n depends on all of them. The maximum likelihood method is employed using a Newton-Raphson algorithm.

3.4. Mixed Diffusion-Jump Process

Press (1967) and Merton (1976) propose a process which mixes Brownian motion and a compound Poisson process to model the movement of stock prices,

$$dP(t) = \alpha P(t)dt + \sigma_D P(t)dB(t) + P(t)(\exp(Q)-1)dN(t). \qquad (11)$$

where $B(t)$ is a standard Brownian motion (BM). $N(t)$ is a homogeneous Poisson process with parameter λ. Q is a normal variate with mean μ_Q and variance σ_Q^2.

Using Ito's Lemma, we can solve the stochastic differential equation (11) for the stock return $X(t)(= \log(P(t)/P(t-1)))$,

$$X(t) = \mu_D + \sigma_D B(1) + \sum_{n=1}^{\Delta N(t)} Q_n, \qquad (12)$$

where $\mu_D = \alpha - \frac{\sigma_D^2}{2}$. The density function for the process is,

$$g(x) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \exp\left(-\frac{(x-\mu_D-n\mu_Q)^2}{2(\sigma_D^2+n\sigma_Q^2)}\right) \frac{1}{\sqrt{2\pi(\sigma_D^2+n\sigma_Q^2)}}. \qquad (13)$$

All moments are finite for this density and T_n depends on all five parameters $\mu_D, \sigma_D^2, \mu_Q, \sigma_Q^2$ and λ. The maximum likelihood estimates are found by using a Quasi-Newton algorithm. However, to numerically maximize the likelihood, we have to truncate the infinite sum in the equation (13) after some value of N. In practice, we choose $N = 11$ which provides satisfactory accuracy.

3.5. Compound Log-normal and Normal

This model is first proposed by Clark (1973). Instead of modelling returns as drawn from a single distribution or a mixture of two distributions, Clark (1973) assumes the returns to be conditional normal, conditional on a variance parameter which is itself stochastic. To be more specific, he assumes $X_i|Z \sim N(0, Z\sigma_1^2)$ and $\log(Z) \sim N(\alpha, \sigma_2^2)$. The density is then,

$$g(x) = \int_0^\infty \{\frac{1}{\sqrt{2\pi z\sigma_1^2}} \exp(-\frac{x^2}{2z\sigma_1^2})\}\{\frac{1}{z\sqrt{2\pi\sigma_2^2}} \exp(-\frac{(\log z - \alpha)^2}{2\sigma_2^2})\} \, dz.$$
(14)

It is easy to show that α and σ_1^2 can be only identified jointly. Consequently, we assume $X_i|Z \sim N(0, Z\sigma_1^2)$ and $\log(Z) \sim N(0, \sigma_2^2)$. The density is then,

$$g(x) = \int_0^\infty \{\frac{1}{\sqrt{2\pi z\sigma_1^2}} \exp(-\frac{x^2}{2z\sigma_1^2})\}\{\frac{1}{z\sqrt{2\pi\sigma_2^2}} \exp(-\frac{(\log z)^2}{2\sigma_2^2})\} \, dz. \quad (15)$$

All moments exist for this density and T_n is invariant to σ_1^2. The estimates are obtained by the maximum likelihood method using a Quasi-Newton algorithm.

3.6. Weibull Distribution

Mittnik and Rachev (1993) first propose to use the Weibull distribution to model stock daily returns. The Weibull distribution is attractive since it is one type of min-stable distribution. More specifically, suppose $m_n = \min\{X_1, \cdots, X_n\}$, where $X_1, \cdots, X_n$ are iid. If, for some constants $c_n > 0$ and $d_n \in R$, $c_n m_n + d_n \overset{d}{\to} Z$, where Z is a random variable with non-degenerate distribution function m, then m could be a Weibull distribution.

The density function for the Weibull distribution is,

$$f(x) = \begin{cases} 0 & \text{if } x < b \\ \frac{\alpha}{a}(\frac{x-b}{a})^{\alpha-1} \exp\{-(\frac{x-b}{a})^\alpha\} & \text{if } x \geq b \end{cases},$$

where α is the index parameter, b is the location parameter and a is the scale parameter and thus T_n is invariant to both a and b. Furthermore, the density has finite all order of moments, for example, $E(X) = a\Gamma(\frac{1}{\alpha} + 1) + b$, $Var(X) = a^2\left\{\Gamma(\frac{2}{\alpha}+1) - \left(\Gamma(\frac{1}{\alpha}+1)\right)^2\right\}$. The estimates are obtained by the maximum likelihood method using a Quasi-Newton algorithm.

3.7. Stable Distribution

Mandlebrot (1963) is the first person who proposes the stable distribution to model stock returns. The stable distribution is usually characterized by the characteristic function. The characteristic function of the general stable

TABLE 1.

T_n in the Empirical Samples

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
$T_n(\theta_0 = 0.25)$	0.8406	0.9694	1.0174

distribution is given by,

$$c(t) = \begin{cases} \exp\{i\mu t - \sigma|t|^\alpha[1 - i\beta sign(t)\tan(\frac{\pi\alpha}{2})]\} & \text{if } \alpha \neq 1 \\ \exp\{i\mu t - \sigma|t|^\alpha[1 + i\beta\frac{2}{\pi}sign(t)\ln(|t|)]\} & \text{if } \alpha = 1 \end{cases}. \qquad (16)$$

where index(α), skewness(β), scale(c), and location(a) are parameters. Obviously T_n is invariant to both c and a. If $1 < \alpha < 2$, which is the case for almost every financial series, the tails of the stable are fatter than those of the normal and the variance is infinite. Unfortunately, the density function has no closed form for $1 < \alpha < 2$. A fast program to do maximum likelihood estimation of stable distributions is provided in Nolan (1999). It is an improved program over the one first proposed in Nolan (1997) and uses a fast 3-dimensional cubic spline interpolation. In this paper the maximum likelihood estimates of all four parameters are found using Nolan's program.

4. IMPLEMENTATION, SIMULATION AND APPLICATION

The dataset we use is daily returns for the Standard and Poor 500 (S&P500) stock market composite raw index. We consider three different periods. The first one is pre-crash sample covering the period from January 1976 to March 1985 with 2,400 observations. The second one also has 2,400 observations but covers the period after the crash from May 1988 to July 1997. The entire sample from January 1976 to July 1997 with 5,614 observations is also examined. Table 1 reports T_n with $\theta_0 = 0.25$ for these three samples. We note that the post-crash sample shows a larger value of T_n than the pre-crash sample. Furthermore, since the entire sample includes October, 1987 — stock-market crash days, it is not surprising that the associated T_n is largest.

As we argued before, the hypothesis we are going to test is the one given by (2). Since all the competing models except the stable distribution have finite variance, we set H_0 to be one of finite-variance models. When T_n is parameter free under H_0, we can choose H_0 to be one distribution family, such as the normal family. Unfortunately, in most cases T_n is not completely parameter free. Consequently, H_0 has to be a certain model with parameters specified. In this paper we specify the parameters to be

TABLE 2.

Estimates of the Competing Models

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
student	$\nu = 6.3879$	$\nu = 3.9942$	$\nu = 3.9382$
MN	$\mu_1 = -2.5873 \times 10^{-4}$	$\mu_1 = 6.2079 \times 10^{-4}$	$\mu_1 = 5.0964 \times 10^{-4}$
	$\mu_2 = 1.3990 \times 10^{-3}$	$\mu_2 = 3.758 \times 10^{-4}$	$\mu_2 = -1.3989 \times 10^{-3}$
	$\sigma_1 = 6.1653 \times 10^{-3}$	$\sigma_1 = 4.2105 \times 10^{-3}$	$\sigma_1 = 7.2094 \times 10^{-3}$
	$\sigma_2 = 1.1670 \times 10^{-2}$	$\sigma_2 = 1.0351 \times 10^{-2}$	$\sigma_2 = 2.6493 \times 10^{-2}$
	$\alpha = 0.6736$	$\alpha = 0.5674$	$\alpha = 0.9528$
LN	$\sigma_2^2 = 0.4576$	$\sigma_2^2 = 0.8811$	$\sigma_2^2 = 0.9063$
MDJ	$\mu = -3.732 \times 10^{-4}$	$\mu = 4.7576 \times 10^{-4}$	$\mu = 5.168 \times 10^{-4}$
	$\mu_Q = 7.06 \times 10^{-4}$	$\mu_Q = 3.05 \times 10^{-5}$	$\mu_Q = -2.047 \times 10^{-4}$
	$\sigma^2 = 2.59 \times 10^{-5}$	$\sigma^2 = 7.42 \times 10^{-6}$	$\sigma^2 = 2.527 \times 10^{-5}$
	$\sigma_Q^2 = 4.73 \times 10^{-5}$	$\sigma_Q^2 = 3.72 \times 10^{-5}$	$\sigma_Q^2 = 9.25 \times 10^{-5}$
	$\lambda = 0.92847$	$\lambda = 1.2796$	$\lambda = 0.5157$
Weibull	$\alpha = 5.0693$	$\alpha = 9.0062$	$\alpha = 20.3287$
stable	$\alpha = 1.6932$	$\alpha = 1.4991$	$\alpha = 1.5460$
	$\beta = 0.2013$	$\beta = -0.0166$	$\beta = 0.0580$

the ML estimates.[2] Therefore, to implement the test, we have to first fit the models in H_0 to the data sets. The relevant estimation method for each candidate model was presented in Section 3. After setting up the hypothesis, we can obtain the asymptotic means and asymptotic variances for T_n based on Proposition (2.3). The p-values are then calculated.

In Table 2 we report the ML estimates of all finite variance models for each data set. We also report the ML estimates of the stable model. Since T_n is invariant to μ, σ^2 in the normal distribution, H, m in the Student distribution, σ_1^2 in the compound log-normal and normal model, a, b in the Weibull distribution, and a, c in the stable distribution, the estimates of these parameters are not reported. Moreover, the estimates of ν in the Student model are less than 4 for both the post-crash sample and the entire sample, the asymptotic distribution of T_n is not applicable in either situation. Table 3 reports the asymptotic distributions and Table 4 shows the associated p-values.

It is also interesting to consider the finite sample performance of the test, such as finite sample distributions and size distortions. To do that we have to specify a distribution for H_1. In this paper we use the stable distribution. In all the Monte Carlo studies 3,000 replications are generated under H_0 and H_1 respectively according to the ML estimates reported in Table 2. T_n is calculated for each replication and thus the finite sample

---

[2]Actually only those parameters on which T_n depends are needed to be specified.

TABLE 3.

Asymptotic Distribution of T_n under H_0

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
Normal	$N(0.7413, 1.97 \times 10^{-4})$	$N(0.7413, 1.97 \times 10^{-4})$	$N(0.7413, 8.42 \times 10^{-5})$
Student	$N(0.8440, 4.50 \times 10^{-4})$	Not Applicable	Not Applicable
MN	$N(0.8402, 3.25 \times 10^{-4})$	$N(0.9802, 5.57 \times 10^{-4})$	$N(0.8953, 4.18 \times 10^{-4})$
LN	$N(0.8543, 4.11 \times 10^{-4})$	$N(0.9688, 8.01 \times 10^{-4})$	$N(0.9754, 3.34 \times 10^{-4})$
MDJ	$N(0.8511, 2.75 \times 10^{-4})$	$N(0.9648, 4.32 \times 10^{-4})$	$N(0.9425, 1.57 \times 10^{-4})$
Weibull	$N(0.7307, 1.86 \times 10^{-4})$	$N(0.7569, 2.35 \times 10^{-4})$	$N(0.7854, 1.31 \times 10^{-4})$

TABLE 4.

p-values of the Test

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
Normal	0	0	0
Student	0.5628	Not Applicable	Not Applicable
LN	0.7512	0.4915	0.0116
MN	0.4907	0.6757	0
MDJ	0.7368	0.4120	0
Weibull	0	0	0

distributions of T_n under H_0 and H_1 are obtained. Using the finite sample distributions, we calculate critical values and powers of the test. In Table 5 we present the finite sample means and sample variances of T_n under H_0 for all three samples. We report the 95% critical value in Table 6 and the power of the test in Table 7. We also perform a Monte Carlo study to obtain the real sizes of the test in finite samples and compare them with the nominal sizes. 3,000 replications are generated under H_0 according to the estimates reported in the second column of Table 2 and each replication has 2,400 observations. The nominal sizes are chosen to be $0.1\%, 0.5\%, 1\%, 5\%, 10\%, 20\%$ and 50%. The sizes are reported on Table 8 and plotted in Figure 1.

A detailed examination of Table 3 and Table 5 reveals that the asymptotic distribution of T_n is very close to the finite sample distribution of T_n across all three samples and all finite-variance distributions. Not surprisingly, therefore, we end up the same conclusions from Table 4 and Table 6. Table 4 indicates that, for all three samples, the normal distribution can be easily rejected by the proposed test statistic, consistent with empirical results when some other test statistics, such as the sample kurtosis, are used. Furthermore, for both the pre-crash sample and the post-crash sample, most finite variance distributions cannot be rejected. For example, for the pre-crash sample the Student distribution, the mixture of normals,

TABLE 5.

Finite Sample Distribution of T_n under H_0

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
Normal	$(0.7420, 1.96 \times 10^{-4})$	$(0.7420, 1.96 \times 10^{-4})$	$(0.7418, 8.98 \times 10^{-5})$
Student	$(0.8442, 4.44 \times 10^{-4})$	Not Applicable	Not Applicable
MN	$(0.8316, 3.46 \times 10^{-4})$	$(0.9760, 5.76 \times 10^{-4})$	$(0.8990, 4.30 \times 10^{-4})$
LN	$(0.8551, 4.10 \times 10^{-4})$	$(0.9691, 8.07 \times 10^{-4})$	$(0.9759, 3.52 \times 10^{-4})$
MDJ	$(0.8520, 3.92 \times 10^{-4})$	$(0.9663, 6.18 \times 10^{-4})$	$(0.9434, 2.38 \times 10^{-4})$
Weibull	$(0.7313, 1.92 \times 10^{-4})$	$(0.7575, 2.40 \times 10^{-4})$	$(0.7855, 1.26 \times 10^{-4})$

TABLE 6.

Finite Sample Distribution of T_n under H_0

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
Normal	0.7653	0.7653	0.7577
Student	0.8814	Not Applicable	Not Applicable
MN	0.8622	1.0155	0.9331
LN	0.8905	1.0182	1.0065
MDJ	0.8837	1.0071	0.9692
Weibull	0.7543	0.7837	0.8035

TABLE 7.

Power of the Test

	Sample 1: 76-85	Sample 2: 88-97	Sample 3: 76-97
Normal	1	1	1
Student	1	Not Applicable	Not Applicable
MN	1	1	1
LN	1	1	1
MDJ	1	1	1
Weibull	1	1	1

TABLE 8.

Size of the Test

Nominal size	0.001	0.005	0.01	0.05	0.1	0.2	0.5
Normal	0.003	0.009	0.012	0.056	0.114	0.223	0.507
Student	0.002	0.0067	0.0110	0.065	0.107	0.192	0.484
MN	0.001	0.0077	0.0117	0.0517	0.096	0.196	0.490
LN	0.0	0.0087	0.0197	0.0637	0.112	0.193	0.501
MDJ	0.007	0.0217	0.0250	0.0937	0.1450	0.257	0.513
Weibull	0.001	0.006	0.0137	0.062	0.114	0.218	0.519

Figure 1: Size of the Test
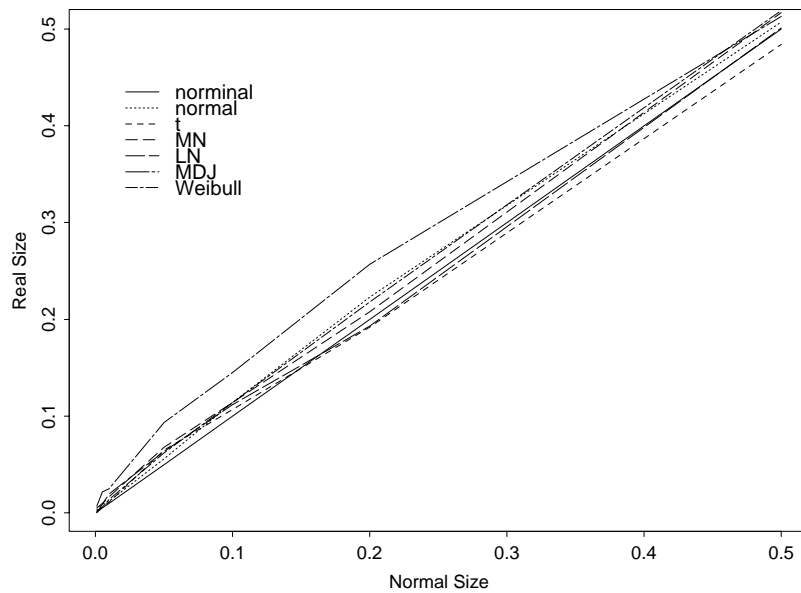
the mixed diffusion-jump process and the compound log-normal and normal model cannot be rejected at 5% significant level. For the post-crash sample the mixture of normals, the mixed diffusion-jump process and the compound log-normal and normal model cannot be rejected at 5% significant level. This finding is consistent with what is normally found in most of the recent literature; see Tucker (1992), Kon (1984), Blattberg and Gonedes (1974). However, for the entire sample all the finite-variance models are rejected at 5% or even smaller significant levels. The finding is very interesting and suggests that when the value of T_n gets bigger and bigger, it is harder and harder for the data to be modeled by the existing finite-variance models. The result is not surprising since a finite-variance model is prone to generate a value of T which is not large enough to match the empirical T_n. If we interpret T_n as a measure of risk, the above finding means that the existing finite-variance models have difficulties to explore the high risk that the actual stock markets have. Finally, Table 7 provides the evidence that in finite samples our test has very good power. From Table 6 and Figure 1, we note that in terms of the size of the test, it works quite well for the normal distribution, the Student t distribution, the mixture of normal distribution, the compound log-normal and normal model, and the

Weibull distribution. Although the size distortions are larger for the mixed diffusion jump model, the biases suggest under-rejection of the model and hence support our finding of rejection of all finite-variance distributions in the above empirical study.

5. EXTENSIONS

Recent literature has found success of using the time-dependent processes for the stock returns. Among them, two families have attracted most attentions, namely, the ARCH-type model (Bollerslev, Chou and Kroner (1992)) and the SV model (Ghysels, Harvey, and Renault (1996)). Both families allow the returns to follow a martingale difference sequence satisfying the α-mixing condition.[3] By assuming the conditional volatility to be correlated, however, the processes are not iid. Under some regular conditions, the marginal distribution of both the ARCH and SV models allows excess kurtosis and has finite variance (see, for example, Bollerslev (1986) and Ghysels, Harvey, and Renault (1996) for the analytical expression of unconditional moments).

The proposed statistic can be used to test an α-mixing sequence. To do that, we obtain the asymptotic distribution of T_n in the following proposition.

PROPOSITION 5.5. *Suppose that* $\{X_1, X_2, \cdots, X_n, \cdots\}$ *is a stationary* α-*mixing sequence with marginal distribution function* $f(x)$. *We further assume*

(i) $f(q_1) > 0$, $f(q_0) > 0$.
(ii) $f(x)$ *is continuous in a neighborhood of* q_1 *and* q_0.
(iii) $E|X_1|^{2r} < \infty$ *for some* $r > 2$ *and* $\sum_{n=1}^{\infty} \alpha^{1-2/r}(n) < \infty$.
Then,

$$T_n \overset{a}{\sim} N(T, \frac{\Sigma^2}{n}), \tag{17}$$

where

$$\Sigma^2 = (q_1 - q_0)^{-4}(E\xi_0^2 + 2\sum_{i=1}^{\infty} E\xi_0\xi_i) \tag{18}$$

with

$$\xi_i = \frac{q_1 - q_0}{2\sigma}((X_i - \mu)^2 - \sigma^2) - \frac{\sigma(I\{X_i > q_1\} - \theta_0)}{f(q_1)} - \frac{\sigma(I\{X_i \leq q_0\} - \theta_0)}{f(q_0)}. \tag{19}$$

---

[3]See Lin and Lu (1996) for the detail on the definitions of various mixing conditions.

It is known that under some conditions most ARCH-type models and SV models are strictly stationary α-mixing processes. Hence (17) can be used to test them against an infinite variance distribution.

6. CONCLUSIONS

This paper has considered a test for the competing models for daily stock returns with particular concern about the variance behavior. In the recent literature, the likelihood ratio test and the Komogorov-Smirnov test are used to compare the descriptive power of the competing models. Both tests suggest that distributions with finite variance outperform the distribution with infinite variance. A common feature for these two tests is that all the observations receive the same weight. Model selection criterion, such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), shares the same spirit although they are used infrequently in this context. Our test statistic, however, assigns different observations different weights. Obviously in our test statistic more extreme observations receive larger weight than less extreme ones. Consequently, our test statistic prefers a distribution whose tail behavior is closer to the empirical distribution to a distribution whose near-origin behavior is closer to the empirical distribution. Although some finite variance models have good descriptive power for both pre-crash sample and after-crash sample, they do not perform well for the entire sample. Therefore, our empirical results suggest either direction. This finding is different from what has been discovered in the recent literature where the finite-variance distributions are found to dominate the stable distribution (see Tucker (1992), Akgiray and Booth (1987), Lau, Lau and Wingender (1990)).

It is important to stress that the purpose of the proposed test is not to choose one out of a fixed set of models as the "best" one and hence different from model selection criteria. Instead our test could serve as pretest or diagnostic checking in order to decide not to use models which appear to be incompatible with the data.

Our test is directly motivated from modelling stock returns, however, it can be also used under other circumstances. One possible area to use the test is the noise behavior in regression models. While classical estimation procedures such as OLS usually perform well and conventional test procedures such as Durbin-Waston test are valid under some moment conditions, serious problems may be encountered in the cases where the variance of the noise is infinite. Our test can be used in this context to check the validity of the finite-variance distribution of the disturbance, and hence serve to select appropriate tools of estimation and inference.

APPENDIX

Proof of Proposition 2.1 It is easy to see that both numerator and the denominator are invariant for a scale-location family up to σ. Hence T_n is invariant.

Proof of Proposition 2.2 The proposition follows immediately from the strong law of large number, since $S_n \to \sigma \quad a.s.$ and $Q_n(1 - \theta_0) - Q_n(\theta_0) \to q_1 - q_0 \quad a.s..$

Proof of Proposition 2.3 Letting $q_{1,0} = q_1 - q_0$, note that

$$\frac{s_n}{Q_n(1 - \theta_0) - Q_n(\theta_0)} - \frac{\sigma}{q_1 - q_0} = \frac{q_{1,0}s_n - \sigma(Q_n(1 - \theta_0) - Q_n(\theta_0))}{(Q_n(1 - \theta_0) - Q_n(\theta_0))q_{1,0}} \quad \text{(A.1)}$$

and

$$s_n = \left(\frac{1}{n-1} \{ \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \} \right)^{1/2} \quad \text{(A.2)}$$

$$= \sigma \left\{ 1 + \frac{1}{(n-1)\sigma^2} \sum_{i=1}^{n} \{(X_i - \mu)^2 - \sigma^2\} - \frac{n(\bar{X} - \mu)^2 - \sigma^2}{(n-1)\sigma^2} \right\}^{1/2}$$

$$= \sigma \left\{ 1 + \frac{1}{2(n-1)\sigma^2} \sum_{i=1}^{n} \{(X_i - \mu)^2 - \sigma^2\} - \frac{n(\bar{X} - \mu)^2 - \sigma^2}{2(n-1)\sigma^2} \right\}$$

$$+ O_P \left(\left(\frac{1}{n} \sum_{i=1}^{n} \{(X_i - \mu)^2 - \sigma^2\} \right)^2 + \left(\frac{n(\bar{X} - \mu)^2 - \sigma^2}{n} \right)^2 \right)$$

$$= \sigma + \frac{1}{2n\sigma} \sum_{i=1}^{n} \{(X_i - \mu)^2 - \sigma^2\} + O_P(1/n).$$

Therefore,

$$q_{1,0}s_n - \sigma(Q_n(1 - \theta_0) - Q_n(\theta_0)) \quad \text{(A.3)}$$

$$= \frac{q_{1,0}}{2n\sigma} \sum_{i=1}^{n} \{(X_i - \mu)^2 - \sigma^2\} - \sigma \{Q_n(1 - \theta_0) - Q_n(\theta_0) - q_{1,0}\} + O_P(1/n)$$

According to the Bahadur representation (see Chapter 3, Csörgő and Horváth (1993)), we have

$$Q_n(1 - \theta_0) - q_1 = -\frac{1}{nf(q_1)} \sum_{i=1}^{n} \{I\{X_i \le q_1\} - (1 - \theta_0)\} + o_P(n^{-1/2})$$

and

$$Q_n(\theta_0) - q_0 = -\frac{1}{nf(q_0)}\sum_{i=1}^{n}\{I\{X_i \leq q_0\} - \theta_0)\} + o_P(n^{-1/2})$$

Putting the above statements together yields

$$\sqrt{n}\Big(q_{1,0}s_n - \sigma(Q_n(1-\theta_0) - Q_n(\theta_0))\Big) \qquad\qquad\qquad (A.4)$$

$$=n^{-1/2}\Bigg\{\frac{q_{1,0}}{2\sigma}\sum_{i=1}^{n}\{(X_i-\mu)^2 - \sigma^2\} + \frac{\sigma}{f(q_1)}\sum_{i=1}^{n}\{I\{X_i \leq q_1\}$$

$$- (1-\theta_0)\} - \frac{\sigma}{f(q_0)}\sum_{i=1}^{n}\{I\{X_i \leq q_0\} - \theta_0)\}\Bigg\} + o_P(1)$$

$$=n^{-1/2}\sum_{i=1}^{n}\Bigg\{\frac{q_{1,0}}{2\sigma}((X_i-\mu)^2 - \sigma^2) - \frac{\sigma(I\{X_i > q_1\} - \theta_0)}{f(q_1)}$$

$$- \frac{\sigma(I\{X_i \leq q_0\} - \theta_0)}{f(q_0)}\Bigg\} + o_P(1)$$

This proves (4). (5) simply follows (4).

Proof of Proposition 2.4

Expanding the expression for Σ^2, we have,

$$\begin{aligned}
\Sigma^2 &= \left(\frac{\sigma}{2(q_1-q_0)}\right)^2 E\left(\left(\frac{X_1-\mu}{\sigma}\right)^4 - 1\right) \\
&+ \frac{\sigma^2\theta_0(1-\theta_0)}{(q_1-q_0)^4}\left(\frac{1}{f^2(q_0)} + \frac{1}{f^2(q_1)}\right) \\
&- \frac{\sigma^2}{(q_1-q_0)^3 f(q_1)}E\left\{\left(\frac{X_1-\mu}{\sigma}\right)^2(I\{X_1 > q_1\} - \theta_0)\right\} \\
&- \frac{\sigma^2}{(q_1-q_0)^3 f(q_0)}E\left\{\left(\frac{X_1-\mu}{\sigma}\right)^2(I\{X_1 \leq q_0\} - \theta_0)\right\} \\
&- \frac{2\sigma^2\theta_0^2}{(q_1-q_0)^4 f(q_0)f(q_1)}.
\end{aligned}$$

Since f is symmetric about μ, we have $f(q_0) = f(q_1)$. A simplification of above expression gives us (8).

Proof of Proposition 5.1

By the strong law of large numbers and the central limit theorem for stationary α-mixing sequence (see, e.g., Lin and Lu (1996)), $(A.4)$ remains valid and so does (5.17).

REFERENCES

Akgiray, V. and G. Booth, 1987, Compound distribution of models of stock returns: An empirical comparison. *Journal of Financial Research* **10**, 269–280.

Bachelier, L., 1900. *Theorie de la Speculation.* Gauthier-Villars, Paris.

Black, F. and M. Scholes, 1973, The pricing of options and corporate liabilities. *Journal of Political Economics* **81**, 637–653.

Blattberg, R. and N. Gonedes, 1974, A comparison of stable and student distribution. *Journal of Business.* **47**, 244–280.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

Bollerslev, T., R. Y. Chou, and K. F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* **52**, 5–59.

Clark, P.K., 1973, A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**, 135–155.

Csörgő, M. and L. Horváth, 1993, *Weighted approximations in probability and statistics.* Wiley, New York.

DuMouchel, W.H., 1973, On the asymptotic normality of the maximum-likelihood estimate when sampling from a stable distribution. *The Annals of Statistics* **1**, 948–957.

Fama, E., 1965, The behavior of stock prices. *Journal of Business* **47**, 244–280.

Ghysels, E., A. C. Harvey, and E. Renault, 1996, Stochastic volatility. In: G.S. Maddala and C.R. Rao, editors, *Handbook of Statistics, Vol 14, Statistical Methods in Finance.* North-Holland, Amsterdam.

Kon, S., 1984, Model of stock returns–a comparison. *Journal of Finance* **39**, 147–165.

Lau, A., H. Lau, and J. Wingender, 1990, The distribution of stock returns: New evidence against the stable model. *Journal of Business and Economic Statistics* **8**, 217–223.

Lin, Z. Y. and C.R. Lu, 1996, *Limit theory for mixing dependent random variables.* Kluwer Academic, New York.

Mandelbrot, B., 1963, The variations of certain speculative prices. *Journal of Business* **36**, 394–419.

McCulloch, J. H., 1978, Continuous time processes with stable increments. *Journal of Business* **51**, 601–619.

Merton, R., 1976, Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* **3**, 125–144.

Mittnik, S. and S.T. Rachev, 1993, Modeling asset returns with alternative stable distributions (disc: P331-389). *Econometric Reviews* **12**, 261–330.

Niederhoffer, V. and M.F.M. Osborne, 1966, Market making and reversal on the stock exchange. *Journal of the American Statistical Association* **61**, 897–916.

Nolan, J., 1997, Numerical calculation of stable densities and distribution functions. *Communication Statistics: Stochastic Models* **13**, 759–774.

Nolan, J., 1999, Maximum likelihood estimation and diagnostics for stable. Manuscript.

Osborne, M., 1959, Brownian motions in the stock market. *Operation Research* **7**, 145–173.

Press, S.J., 1967, A compound events models of security prices. *Journal of Business* **40**, 317–335.

Tucker, A., 1992, A reexamination of finite- and infinite-variance distributions as models of daily stock returns. *Journal of Business and Economic Statistics* **10**, 73–81.