Normal Log-normal Mixture: Leptokurtosis, Skewness and Applications

Minxian Yang

School of Economics

The University of New South Wales

Abstract

The properties and applications of the normal log-normal (NLN) mixture are considered. The moment of the NLN mixture is shown to be finite for any positive order. The expectations of exponential functions of a NLN mixture variable are also investigated. The kurtosis and skewness of the NLN mixture are explicitly shown to be determined by the variance of the log-normal and the correlation between the normal and log-normal. The issue of testing the NLN mixture is discussed. The NLN mixture is fitted to a set of cross-sectional data and a set of time-series data to demonstrate its applications. In the time series application, the ARCH-M effect and leverage effect are separately estimated and both appear to be supported by the data.

JEL Classification: C22, C51, G12

Key words: GARCH, stochastic volatility, ARCH-M, maximum likelihood

Correspondence: `m.yang@unsw.edu.au`, Tel: 61-2-9385-3353, Fax: 61-2-9313-6337

*Financial support from the Australian Research Council is gratefully acknowledged.

1 Introduction

The normal log-normal (NLN) mixture in this paper is defined as the distribution of the product of a normal random variable of a log-normal random variable that are generally correlated. The NLN mixture has long been recognized as a useful distribution for describing speculative price changes or returns. Clark (1973) showed that the marginal distribution of price changes should be the NLN mixture rather than a member of the stable family. Tauchen and Pitts (1983) introduced a bi-variate model for price changes and trading volumes, where the marginal distribution of the price changes was the NLN mixture. Empirical work of Hsieh (1989) demonstrated the usefulness of the NLN mixture in generalized autoregressive conditional heteroskedasticity (GARCH) type models (Engle (1982) and Bollerslev (1986)). The assumption of zero correlation between the normal and log-normal was maintained in the above articles.

More recently, in the literature of stochastic volatility (SV) models (see Ghysels *et al* (1996)), the distribution of shocks to returns is also the NLN mixture (when normality is assumed for both the mean and the log variance processes) [1]. The SV models generally allow for non-zero correlation between the normal and log-normal, which is labelled as the leverage effect of Black (1976). Ghysels *et al* (1996) showed that the absolute moment of a SV process is finite for any positive order under the assumption of zero correlation. We also refer to Koopman and Uspensky (2002) for applications and references of the SV models.

In this paper, we investigate the moment properties of the NLN mixture with non-zero correlation between the normal and log-normal. Similar to the

[1] It is occasionally assumed in the SV models that the log-normal shock (to the log variance) is one-lag behind the normal shock (to the return)

result of Ghysels *et al* (1996), the moment of the NLN mixture is shown to be finite for any positive order. By deriving explicitly the first four centered-moments, we also show that the skewness and kurtosis are determined by the variance of the log-normal and the correlation and that the NLN mixture is generally skewed and leptokurtotic.

In exponential GARCH (EGARCH) models, it is desirable to determine the existence of expectations of exponential functions of a NLN mixture variable in order to assert the stationarity of data generating processes. Similar to a result of Nelson (1992), we find that $E\exp\{au\}$ does not exist for any constant $a \neq 0$, where u is the NLN mixture random variable. For a function $\tau(u)$ that is linear for small $|u|$ and logarithmic for large $|u|$, we show that $E\exp\{a\tau(u)\}$ is finite for any a.

The NLN mixture density function is reduced to the normal density when the log-normal variance approaches to zero. This implies that, in testing the null hypothesis of the normal distribution against the alternative of the NLN mixture, the correlation parameter is unidentified under the null. Along the line of Andrews and Ploberger (1994,1995), a strategy for testing the NLN mixture is suggested, which may ease the computational burden of the mixture test in certain situations.

We argue that the NLN mixture is useful in a cross-sectional context where the error terms in a regression model possess idiosyncratic variances. To demonstrate the cross-sectional applications, a set of annual cross-sectional stock returns from the Australian Stock Exchange is fitted to the NLN mixture, using the maximum likelihood (ML) method. The NLN mixture is compared with the normal, t and skewed t distributions for this data set. The NLN mixture appears to be able to capture the heterogeneity in the error term's variance.

As a time-series application, a general SV model is considered as a starting point for modelling speculative return series. Following the ARCH literature, we allow the conditional log variance process to depend directly on past shocks such that volatility clustering can be captured. However, we maintain the SV specification that the log variance is the sum of the conditional log variance and a contemporaneous shock. The resulting model turns out to be an EGARCH model with ARCH in mean (ARCH-M) effect, where the iid disturbance term has the NLN mixture distribution. An interesting feature of this model is that the (positive) ARCH-M effect is separated from the (negative) leverage effect, making the model useful in quantifying these effects separately. A filtering function (linear for small shocks and logarithmic for large shocks) is introduced in the log variance process for two purposes. First, it ensures the stationarity of the model's data generating process. Second, it reduces, to the extent of an estimable parameter, the impact of extremely large shocks on the conditional variance and makes the model robust to outliers. The NLN based model is estimated, using the ML method, for a SP500 return series of Koopman and Uspensky (2002). The model's fit to the data appears reasonably good. The estimation results lend some support to the positive ARCH-M effect and the negative leverage effect.

Since the NLN mixture density function can only be expressed as an integral, the density evaluation required by the ML estimation is carried out using Romberg's numerical integration method. A subroutine in Fortran-90 for computing the density function is available upon request.

Section 2 contains some properties of the NLN mixture. Section 3 is a brief discussion on mixture test. Sections 4 and 5 are respectively examples for cross-sectional and time-series applications. Sections 6 concludes and Section 7 collects proofs. Throughout the paper, the (natural) exponential function are expressed either by $\exp(x)$ or simply e^x.

2 Normal Log-normal Mixture

Consider the random variable u given by

$$u = e^{\frac{1}{2}\eta}\varepsilon, \tag{1}$$

where ε and η are random variables satisfying

$$\begin{bmatrix} \varepsilon \\ \eta \end{bmatrix} \sim \mathrm{N}\Big(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix} \Big), \qquad -1 < \rho < 1 \tag{2}$$

with σ and ρ being constant parameters. The random variable u will be labelled as the normal log-normal (NLN) mixture. In the context of time series, with the time index t attached to (ε, η), the mixture $e^{\frac{1}{2}\eta_t}\varepsilon_t$ (with a correction in mean) can be viewed as the simplest stochastic volatility model for speculative return series, where ε and η are respectively the shocks to the mean and log-variance of the return.

The distribution of u is a mixture of normals and its conditional distribution is $\mathrm{N}\big((\rho/\sigma)\eta e^{\frac{1}{2}\eta},\, (1-\rho^2)e^{\eta}\big)$ for each given η. The joint density function of $[u, \eta]'$ can be written as

$$\begin{aligned} \mathrm{pdf}_{u,\eta}(u, \eta) &= \mathrm{pdf}_{u|\eta}(u|\eta) \times \mathrm{pdf}_{\eta}(\eta) \\ &= [2\pi(1-\rho^2)e^{\eta}]^{-\frac{1}{2}} \exp\{-\frac{1}{2(1-\rho^2)e^{\eta}}[u - (\rho/\sigma)\eta e^{\frac{1}{2}\eta}]^2\} \\ &\quad \times (2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2}\eta^2\} \end{aligned} \tag{3}$$

where $\mathrm{pdf}_{\eta}(\cdot)$ is the marginal density function of η and $\mathrm{pdf}_{u|\eta}(\cdot|\eta)$ is the conditional density function of u for given η. The function $\mathrm{pdf}_{u|\eta}(\cdot|\eta)$ is not defined for $\sigma = 0$ (when η degenerates to zero).

The marginal density function of u is given by

$$\mathrm{pdf}_u(u|\sigma,\rho) = \int_{-\infty}^{\infty} \mathrm{pdf}_{u|\eta}(u|\eta)\mathrm{pdf}_\eta(\eta)\mathrm{d}\eta.$$

By the transformation $\eta = \sigma y$, the above integral becomes

$$\mathrm{pdf}_u(u|\sigma,\rho) = \int_{-\infty}^{\infty} f(u,y|\sigma,\rho)\phi(y)\mathrm{d}y, \tag{4}$$

where $\phi(\cdot)$ is the standard normal density and

$$f(u,y|\sigma,\rho) = [2\pi(1-\rho^2)e^{\sigma y}]^{-\frac{1}{2}} \exp\{-\frac{[u-\rho y e^{\frac{1}{2}\sigma y}]^2}{2(1-\rho^2)e^{\sigma y}}\}.$$

Although the analytical form of $\mathrm{pdf}_u(u|\sigma,\rho)$ is unknown, it can be readily evaluated for given (u,σ,ρ) either by simulation or by numerical integration.

We note that $f(u,y|\sigma,\rho) = f(u,-y|-\sigma,-\rho)$ and the distribution of y is symmetric about zero. Hence, $\mathrm{pdf}_u(u|\sigma,\rho) = \mathrm{pdf}_u(u|-\sigma,-\rho)$. The implication of this, which is used in Section3, is that the usual restriction "$\sigma \geq 0$" can be ignored in estimating and testing the NLN mixture.

Let $\Psi = \{(\sigma,\rho) : \sigma \in (-\bar{\sigma},\bar{\sigma}), \rho \in (-\bar{\rho},\bar{\rho})\}$ be a space of (σ,ρ), where $\bar{\sigma} > 0$ and $0 < \bar{\rho} < 1$. The function $f(u,y|\sigma,\rho)$ in (4) satisfies

$$f(u,y|\sigma,\rho) \leq \bar{f}(y) = \begin{cases} (2\pi(1-\bar{\rho}^2)e^{\bar{\sigma}y})^{-\frac{1}{2}}, & \text{if } y < 0 \\ (2\pi(1-\bar{\rho}^2)e^{-\bar{\sigma}y})^{-\frac{1}{2}}, & \text{if } y \geq 0 \end{cases}$$

for all u and all $(\sigma,\rho) \in \Psi$. As $\int_{-\infty}^{\infty} \bar{f}(y)\phi(y)dy$ is finite, by the dominated convergence theorem, the density $\mathrm{pdf}_u(u|\sigma,\rho)$ is continuous at $\sigma = 0$. Further, it can easily be shown that $\mathrm{pdf}_u(u|0,\rho) = \phi(u)$ is the standard normal density. Therefore, ρ is unidentified when $\sigma = 0$.

It can be verified that the first four centered moments of u are given by

$$\begin{aligned} c_1 &= E(u) = \frac{1}{2}\rho\sigma e^{\frac{1}{8}\sigma^2}, \\ c_2 &= E(u-c_1)^2 = e^{\frac{1}{2}\sigma^2}[1+\rho^2\sigma^2(1-\frac{1}{4}e^{-\frac{1}{4}\sigma^2})], \end{aligned}$$

$$
\begin{aligned}
c_3 &= E(u-c_1)^3 = \rho\sigma e^{\frac{9}{8}\sigma^2}\{(1-\rho^2)[\frac{9}{2}-\frac{3}{2}e^{-\frac{1}{2}\sigma^2}] \\
&\quad +\rho^2[(\frac{9}{2}+\frac{27}{8}\sigma^2)-\frac{3}{2}(1+\sigma^2)e^{-\frac{1}{2}\sigma^2}+\frac{1}{4}\sigma^2 e^{-\frac{3}{4}\sigma^2}]\}\,, \\
c_4 &= E(u-c_1)^4 = e^{2\sigma^2}\{3(1-\rho^2)^2 \\
&\quad +6\rho^2(1-\rho^2)[(1+4\sigma^2)-\frac{3}{2}\sigma^2 e^{-\frac{3}{4}\sigma^2}+\frac{1}{4}\sigma^2 e^{-\frac{5}{4}\sigma^2}] \\
&\quad +\rho^4[(3+24\sigma^2+16\sigma^4)-(9+\frac{27}{4}\sigma^2)e^{-\frac{3}{4}\sigma^2} \\
&\qquad +\frac{3}{2}(1+\sigma^2)\sigma^2 e^{-\frac{5}{4}\sigma^2}-\frac{3}{16}\sigma^4 e^{-\frac{3}{2}\sigma^2}]\}\,. \qquad (5)
\end{aligned}
$$

If ρ is small such that the terms associated with ρ^2 can be ignored, then the skewness and kurtosis of u are given by

$$
\begin{aligned}
\text{Skewness} &\approx \frac{1}{2}\rho\sigma e^{\frac{3}{8}\sigma^2}(9-3e^{-\frac{1}{2}\sigma^2}), \\
\text{Kurtosis} &\approx 3e^{\sigma^2}. \qquad (6)
\end{aligned}
$$

The marginal distribution of u is skewed and thick-tailed when both σ and ρ are non-zero. The kurtosis is mainly controlled by σ^2 and the skewness by $\rho\sigma$. These properties of u appear desirable for modelling the returns of speculative prices, which are often found to have sample distributions with leptokurtosis (thick-tails and a large peak at the origin) and negative skewness. The kurtosis formula $3e^{\sigma^2}$ was given in Clark (1973) for the case that $\rho = 0$.

To compare $\mathrm{pdf}_u(u|\sigma,\rho)$ with the standard normal pdf, the density function of the standardized mixture variable $v = (u-c_1)/\sqrt{c_2}$

$$
m(v|\,\sigma,\rho) = \sqrt{c_2}\,\mathrm{pdf}_u(c_1+\sqrt{c_2}\,v|\,\sigma,\rho) \qquad (7)
$$

is plotted with $\phi(v)$ in Figure 1 for various values of σ and ρ. Evidently, $m(v|\sigma,\rho)$ is close to $\phi(v)$ for small σ and possesses prominent leptokurtosis for large σ. Further, when $\sigma > 0$ and $\rho < 0$, the distribution has a positive mode and a thick left-tail. We also note that $m(v|-\sigma,-\rho) = m(v|\,\sigma,\rho)$.

In addition to the moments given in (5), the following propositions provide further results regarding the moments of u

Proposition 1 *For any finite integer $k > 0$, $E|u|^k < \infty$.*

While all moments of u exist as indicated in the above proposition, it is the expectation of exponential functions of u that is of interest in exponential ARCH models. We provide the following propositions for this purpose, where $v^+ = \max(0, v)$, $v^- = \max(0, -v)$ and $v = (u - c_1)/\sqrt{c_2}$.

Proposition 2

(a) For any finite constant $a \neq 0$, $E(e^{au}) = \infty$.

(b) $E(e^{av^+})$ and $E(e^{av^-})$ are finite if and only if $a \leq 0$.

The above results also imply that $E(e^{d_1|v|+d_2 v})$ exists if and only if both $d_1 + d_1 < 0$ and $d_1 - d_2 < 0$. The integral $E(e^{au})$ diverges because $u = e^{\frac{1}{2}\eta}\varepsilon$ contains an exponential factor that dominates eventually. Intuitively, if a function $\tau(u)$ behaves like logarithm for large $|u|$, the expectation of $e^{\tau(u)}$ should exist. This idea is formalized as the following proposition.

Proposition 3

(a) For any continuous function $\tau(\cdot)$ that satisfies

$$\tau(u) \leq \begin{cases} a_0 + a_1 \ln(a_2 + a_3|u|), & \textit{for } \ |u| > b \\ a_4, & \textit{for } \ |u| \leq b \end{cases},$$

where $a_0, a_1, a_2, a_3, a_4, b$ are constants with $a_1 > 0$, $a_3 > 0$, $b > 0$ and b being sufficiently large, $E(e^{\tau(u)})$ is finite.

(b) For the function

$$\tau_b(x) = \begin{cases} -b - \ln(1 + |x + b|), & \textit{for } \ x < -b \\ x, & \textit{for } \ |x| \leq b \\ b + \ln(1 + |x - b|), & \textit{for } \ x > b \end{cases},$$

where $b > 0$, $E(e^{d_1|\tau_b(v)|+d_2\tau_b(v)})$ is finite for any constants d_1 and d_2. Further, $E(e^{a\tau_b(v^+)})$ and $E(e^{a\tau_b(v^-)})$ are also finite for any constant a.

The function $\tau_b(x)$ is continuous, increasing, odd, linear for small $|x|$ and logarithmic for large $|x|$. It can be verified that the first-order derivative of $\tau_b(\cdot)$ is continuous.

Below is a summary of the properties of the NLN mixture, where $v = (u - c_1)/\sqrt{c_2}$ and $m(v|\sigma, \rho)$ is the density function of v.

- The kth moment of v is finite for any finite k.
- The mean and variance of v are zero and one respectively.
- When $\sigma = 0$, $m(v|\sigma, \rho) = \phi(v)$ and ρ is unidentified.
- When $\sigma > 0$ and $\rho = 0$, $m(v|\sigma, \rho)$ is symmetric with leptokurtosis.
- When $\sigma > 0$ and $\rho < 0$, $m(v|\sigma, \rho)$ is skewed to the left.
- When $\sigma > 0$ and $\rho > 0$, $m(v|\sigma, \rho)$ is skewed to the right.
- $E(e^{au})$ does not exist for any $a \neq 0$.
- $E(e^{d_1|\tau_b(v)|+d_2\tau_b(v)})$ exists for any constants d_1 and d_2.

The appeal of the NLN mixture $\mathrm{pdf}_u(u|\sigma, \rho)$, or $m(v|\sigma, \rho)$, in modelling speculative returns has long been recognized in the literature [see Clark (1973), Tauchen and Pitts (1983) and Hsieh (1989) among others]. However, the properties of the NLN mixture given in this section appear to be new.

3 Mixture Test

In applications, an obvious question is whether or not the mixture distribution is favored over the normal distribution. The question can be answered by testing $H_0 : \sigma = 0$ (normal) against $H_1 : \sigma \neq 0$ (mixture). However, since ρ is not identified under H_0, the usual χ^2 asymptotics does not apply to the likelihood ratio (LR) statistic in this context. In general, the asymptotic null distribution of the LR statistic and its critical values need to be simulated on a case-by-case basis [see Andrews and Ploberger (1994, 1995)].

To take advantage of the χ^2 distribution, we consider the following three hypotheses

$$H_0 : \sigma = 0; \quad H_r : \sigma \neq 0, \rho = 0; \quad H_1 : \sigma \neq 0, \rho \in (-1, 1);$$

where ρ is not identified under H_0. Let $L(r)$ be the LR statistic for testing H_0 against H_r. The χ^2 critical values are approximately correct for $L(r)$ because testing H_0 against H_r has no complications. Therefore, H_0 is rejected if $L(r)$ is greater than the critical value (3.84 at 5%, say). In this case, the hypotheses about ρ can be further tested on the models with $\sigma \neq 0$ in the usual manner.

However, if $L(r)$ is not greater than the critical value (3.84), we cannot infer the validity of H_0 because H_r is only a restricted version of H_1. In this case, H_0 has to be tested directly against H_1, where the likelihood is maximized over all feasible ρ, and corresponding critical values need to be simulated.

Because simulating critical values in non-linear models is difficult (it involves thousands of non-linear maximizing operations), the above description of the mixture test is particularly useful in the case that H_0 can be rejected in favor of H_r (which is true for the applications in Sections 4 and 5 where simulating critical values is avoided).

4 Example for Cross-sectional Data

We demonstrate that (1) may be used to describe cross-sectional data. It is useful for modelling a heterogeneous population where the variance of the error term in a regression model varies from one individual to another. In such cases, treating the variance as random may be desirable.

In particular, for cross-sectional data $\{y_1, \ldots, y_n\}$, we consider the simplest regression model $y_i = \mu_0 + e^{\frac{1}{2}h_i}\varepsilon_i$ $(i = 1, \ldots, n)$, where the variance of disturbance is determined by $h_i = \lambda + \eta_i$, μ_0 and λ are constant parameters, (ε_i, η_i) are iid and follow the distribution in (2). The specification $h_i = \lambda + \eta_i$ captures idiosyncratic variations in the variance of the error term $e^{\frac{1}{2}h_i}\varepsilon_i$ (subject to a mean correction). The correlation between η_i and ε_i quantifies the joint behavior of variations in mean and variance. Apparently, this model can easily be extended to the multiple regression model $y_i = \mu_0 + x_i\beta + e^{\frac{1}{2}h_i}\varepsilon_i$, where x_i is a vector of regressors and β a vector of parameters.

The above model can be rewritten as

$$y_i = \mu + e^{\frac{1}{2}\omega} v_i, \qquad i = 1, \ldots, n, \tag{8}$$

where v_i are independent errors with the distribution (7), $\mu = \mu_0 + c_1 e^{\frac{1}{2}\lambda}$ $e^{\frac{1}{2}\omega} = \sqrt{c_2} e^{\frac{1}{2}\lambda}$, c_1 and c_2 are given in (5). We recall that the parameter σ and ρ in (7) control the skewness and kurtosis of the NLN distribution.

We consider the cross-sectional returns from the top 200 stocks listed in Australian Stock Exchange (ASX200). The annual log returns, for the year from 15 January 2002 to 15 January 2003, of the ASX200 stocks are used. As some stocks were not listed on 15 January 2002, there are only 192 useful observations. A uniform distribution is used to obtain a random sample of size 100 from the 192 observations. The descriptive statistics of the sample are presented in Table 1, which exhibit skewness and excess kurtosis.

Table 1. Descriptive Statistics for 100 Annual Returns

Mean	Standard Deviation	Skewness	Excess Kurtosis	Minimum	Maximum
-0.0738	0.3578	-0.9779	1.7778	-1.3556	.7792

The data are fitted to the five models listed below

- M0: standard normal distribution,
- M1: normal log-normal mixture distribution with $\rho = 0$,
- M2: normal log-normal mixture distribution,
- MT: Student's t distribution (v_i follows t),
- MS: Skewed t distribution (v_i follows skewed t).

The Skewed t distribution, introduced by Hansen (1994), has the density function

$$\kappa(z|\nu,\varphi) = \begin{cases} bc\Big(1+\frac{1}{\nu-2}(\frac{bz+a}{1-\varphi})^2\Big)^{-(\nu+1)/2}, & z < -a/b \\ \\ bc\Big(1+\frac{1}{\nu-2}(\frac{bz+a}{1+\varphi})^2\Big)^{-(\nu+1)/2}, & z \geq -a/b \end{cases}, \tag{9}$$

where $\nu \in (2,\infty)$, $\varphi \in (-1,1)$, $a = 4\varphi c(\nu-2)/(\nu-1)$, $b^2 = 1+3\varphi^2-a^2$, and $c = [\Gamma(\frac{\nu+1}{2})/\Gamma(\frac{\nu}{2})]/\sqrt{\pi(\nu-2)}$. The mean, mode and variance of the distribution are 0, $-a/b$ and 1 respectively. When $\varphi = 0$, the distribution reduces to the Student's t (symmetric). When $\varphi < 0$ or $\varphi > 0$, the distribution is skewed to the left or right respectively.

The estimation results for the five models are reported in Table 2. Clearly, M1 is a restricted version of M2 ($\rho = 0$) and and M0 is a restricted version of M1 ($\sigma = 0$). A direct comparison of M0 and M2 is not straightforward because ρ is not identified under the null model M0. However, using the strategy in Section 3, we infer that M0 is rejected in favor of M1 and M1 is rejected in favor of M2 (the asymptotic p-values for the likelihood ratios are less than 1%). Similarly, as MT is nested in MS, the likelihood ratio indicates that MT should be rejected in favor of MS.

Table 2. Estimation Results for ASX200 Annual Returns

Parameter	M0	M1	M2	MT	MS
μ	-.0738	-.0009	-.0775	-.0130	-.0912
	[.036]	[.052]	[.036]	[.042]	[.038]
ω	-2.0658	-1.9914	-1.8020	-1.8346	-1.5580
	[.191]	[.253]	[.295]	[.437]	[.678]
σ, ν		1.0390	1.0886	3.1537	2.8332
		[.342]	[.233]	[1.231]	[.929]
ρ, φ			-.5009		-.4552
			[.119]		[.115]
LogL	-38.606	-33.327	-26.998	-33.543	-26.938

An advantage of M2 over MS is that the parameters (σ, ρ) have a natural interpretation. For M2, the estimate of the correlation between the mean error and variance error (ρ) is significantly negative and may be interpreted as a piece of evidence for the leverage effect (the tendency that a decrease in the stock price increases the stock's riskiness or volatility).

The coefficient of variation

$$\mathrm{CV} = \sqrt{\mathrm{Var}(e^{h_t})}/\mathrm{E}(e^{h_t}) = \sqrt{e^{\sigma^2} - 1}$$

may be used to compare the variation in the variance to the variation in the mean, as the denominator of CV is proportional to the *average* conditional variance (given η_t) of the error term to the mean equation [2]. From the point

[2] The conditional variance is $c_2(1 - \rho^2)e^{h_t}$.

estimates for M2, CV is estimated as 1.507 with the standard error being 0.551, signifying a great deal of variations in the variance. Presumably, extending the model to the multiple regression and including firm-specific factors (e.g. market capitalization and sector dummy) would reduce the variation in the variance.

The density and cumulative distribution functions of M2 and MS (evaluated at the point parameter estimates) are presented in Figure 2, together with the standard normal and standardized data distributions. These graphs present some visual evidence that M2 and MS fit the data better than M0 (standard normal), in the sense that the M2 and MS curves are generally closer to the data curve. The graphs also indicate that M2 fits data slightly better than MS.

Both M2 and MS capture the kurtosis and skewness of the data but neither matches the sample kurtosis and skewness in Table 1. For M2, the kurtosis and skewness implied by the point estimates of (σ, ρ) are 19.08 and -2.74 respectively, far different from the simple sample estimates in Table 1. Similarly, for MS, the point estimate for the degree of freedom (ν) imply that the third and fourth moments do not exist. It is likely that the mean function (a constant) here is too simply to adequately explain all features in the data by either M2 or MS. However, given that the simple sample skewness and kurtosis measures are not robust (Kim and White, 2004), it is not clear whether matching the sample skewness and kurtosis is a desirable criterion to judge a model.

5 Example for Time Series Data

The empirical properties of asset return series have been well documented in the literature (e.g. Bollerslev *et al* (1994) and Ghysels *et al* (1996)). The main characteristics include thick-tailed distribution (large excess kurtosis), volatility clustering (strong serial correlation in squared return series) and leverage effects (negative correlation between shocks to return and to volatility). Various volatility models have been proposed to capture these characteristics. In the discrete-time context, the volatility models may be divided into two classes: the autoregressive conditional heteroskedasticiy (ARCH) models and the stochastic volatility (SV) models.

We show that a change in the variance specification in a SV model leads to an exponential GARCH (EGARCH) model with the normal log-normal (NLN) mixture distribution. This model has the following merits. It allows both leptokurtosis and skewness in the shocks to a return series. It can separate the ARCH-M effect from the leverage effect. It is relatively robust to outliers in the sense that the impact of extreme shocks on the return's conditional variance is reduced.

A number of thick-tailed distributions have been suggested for ARCH-type models, including the Student-t distribution, generalized-error distribution and skewed-t distribution (Hansen, 1996). Of these, only the skewed-t distribution accommodates skewness. An advantage of the NLN distribution in (7) is the useful interpretations for its parameters (σ, ρ).

The NLN-based volatility model is estimated for a SP500 return series of Koopman and Uspensky (2002).. In this section, symbols without a time subscript represent constant parameter.

5.1 Model

Let the return of an asset in the period t be $x_t = \ln(S_t/S_{t-1})$, where S_t the spot price of the asset at the end of period t. Suppose that the return x_t can be modelled as

$$x_t = \mu_t + e^{\frac{1}{2}h_t}\varepsilon_t \qquad h_t = \lambda_t + \eta_t, \tag{10}$$

where the 2-dimensional disturbance series $[\varepsilon_t, \eta_t]'$ is iid with the distribution (2), μ_t and λ_t depend only on information prior to t. When λ_t is a function of h_{t-1} only, the above model is the SV model. When λ_t is a function of $(\varepsilon_{t-1}, \lambda_{t-1})$ and $\sigma = 0$ (or $\eta_t \equiv 0$), the above model is the ARCH-type model.

From a practical point of view, it is reasonable to favor a model that allows ε_{t-1} to directly impact h_t without ruling out the contemporaneous shock η_t to h_t. Taking the middle way between the SV model and the ARCH-type model, we specify λ_t as a function of (v_{t-1}, λ_{t-1}),

$$\lambda_t = \omega + \alpha g(v_{t-1}) + \beta\lambda_{t-1}, \tag{11}$$

where $g(\cdot)$ is a function, $v_t = (u_t - c_1)/\sqrt{c_2}$ with $u_t = e^{\frac{1}{2}\eta_t}\varepsilon_t$, $c_1 = E(u_t)$ and $c_2 = \mathrm{Var}(u_t)$ (as given in (5)). Under this specification, $x_t = \mu_t + e^{\frac{1}{2}\lambda_t}u_t$.

If $g(v_{t-1}) = |v_{t-1}| + \gamma v_{t-1}$, (11) becomes the EGARCH formulation of Nelson (1991) that allows asymmetric effects of positive and negative shocks. We note that $g(v_{t-1}) = (1+\gamma)v_{t-1}^+ + (1-\gamma)v_{t-1}^-$ with $v_{t-1}^+ = \max(0, v_{t-1})$ and $v_{t-1}^- = \max(0, -v_{t-1})$. When $|\beta| < 1$, according to Proposition 2, the (unconditional) expectation of $e^{\frac{1}{2}\lambda_t}$ exist if and only if $\alpha(1+\gamma) \le 0$ and $\alpha(1-\gamma) \le 0$, a result similar to that of Theorem A1.2 in Nelson (1991). This condition is unlikely met in practice because the conditional variance is typically positively related to past shocks (clustering).

Instead, we propose to use

$$g(v_{t-1}) = |\tau_b(v_{t-1})| + \gamma\,\tau_b(v_{t-1}), \tag{12}$$

where the function $\tau_b(\cdot)$ is defined in Proposition 3. There are a number of advantages for choosing this function. First, the asymmetric effects of past shocks on λ_t are carried over from the original Nelson specification. Second, the the data-generating process for x_t is covariance stationary for all possible values of α and γ. Finally, the function $\tau_b(\cdot)$ dampers the impact of large shocks, making λ_t robust to outliers. The threshold b here is estimable.

We specify the mean function μ_t as

$$\mu_t = \mu + \delta e^{\frac{1}{2}\lambda_t}, \tag{13}$$

where the ARCH-M effect of Engle *et al* (1987) is accommodated. Under this specification, the mean equation for x_t becomes

$$x_t \;=\; \mu + (\delta + c_1)e^{\frac{1}{2}\lambda_t} + \sqrt{c_2}e^{\frac{1}{2}\lambda_t}v_t. \tag{14}$$

Let $\mathcal{F}_t$ be the set of observable information at t (the sigma-field generated by $\{\lambda_0, x_1, \ldots, x_t\}$). The conditional mean and variance becomes $E(x_t|\mathcal{F}_{t-1}) = \mu + (\delta + c_1)e^{\frac{1}{2}\lambda_t}$ and $\mathrm{Var}(x_t|\mathcal{F}_{t-1}) = c_2 e^{\lambda_t}$ respectively. The iid shock to x_t, $v_t = (u_t - c_1)/\sqrt{c_2}$, follows the NLN mixture distribution in (7).

The model defined in (10-14), while being derived from a SV model, is clearly an EGARCH model of Nelson (1991) with the ARCH-M effect. It can obviously be extended to higher orders. By Theorem 2.1 of Nelson (1991), x_t is strictly stationary and ergodic and λ_t is covariance stationary if and only if $|\beta| < 1$. For higher order cases, the stationarity condition becomes that the roots of the autoregressive polynomial (with respect to λ_t and its lags) are outside the unit circle. Under the same condition, for any finite $b > 0$ in $\tau_b(\cdot)$, x_t is also covariance stationary by Proposition 3.

An interesting feature of the model is that the total ARCH-M effect is characterized by $(\delta + c_1)$. Here, δ may be viewed as the asset's risk premium, which should be positive according to the capital-asset-pricing model

(CAPM). On the other hand, $c_1 = \frac{1}{2}\rho\sigma e^{\frac{1}{8}\sigma^2}$ (not a free parameter) is determined by ρ, the leverage effect or the correlation between the contemporaneous shocks ε_t and η_t, which is theoretically negative as argued by Black (1976). Therefore, the sign of the total ARCH-M effect $(\delta + c_1)$ is generally undetermined and dependent on which effect dominates in a given return series. Indeed, the empirical evidence on the relationship between the conditional mean and the conditional variance of speculative return series has been mixed and inconclusive (see Koopman and Uspensky (2002) for a concise discussion and references). Towards resolving this issue, our model makes it possible that the risk premium effect (δ) and the leverage effect (ρ) can be quantified separately.

For a given sample $\{x_1, \ldots, x_n\}$, the parameter vector

$$\theta = (\mu, \delta, \omega, d, \gamma, \alpha, \beta, \sigma, \rho),$$

where $d = 1/\sqrt{b}$ (or $b = 1/d^2$) with b being defined in Proposition 3, can be estimated by the maximum likelihood (ML) method. The log likelihood function is given by

$$\ell(\theta|\lambda_0, v_0) = \sum_{t=1}^{n} \Big\{ \ln \Big[m\Big(\frac{x_t - \mu - (\delta + c_1)e^{\frac{1}{2}\lambda_t}}{\sqrt{c_2}e^{\frac{1}{2}\lambda_t}} \Big| \sigma, \rho \Big) \Big] - \frac{1}{2}(\lambda_t + \ln c_2) \Big\}, \quad (15)$$

where the density function $m(v|\sigma, \rho)$ is given in (7) and (4), which is evaluated by numerical integration using Romberg's method. The initial values λ_0 and v_0 are treated as known. In estimation, λ_0, v_0 and $|v_0|$ are replaced by the log sample variance, zero and the sample absolute deviation respectively.

The consistency and asymptotic normality of the quasi-ML estimator of GARCH models have been established by Lee and Hansen (1994), Lumsdaine (1996) and Ling and McAleer (2003) among others. Their results are obtained under the assumption that the density function used for the quasi-ML estimation is the standard normal density, which is not applicable to the ML estimation in this paper. Since establishing asymptotic properties for the NLN-based EGARCH model here is beyond the scope of this paper, we will assume that the consistency and asymptotic normality hold for the ML estimator of (15). The interpretations of the estimation results in this section are based on this assumption.

5.2 Data and Estimation Results

A daily excess return series of the SP500 price index, taken from Koopman and Uspensky (2002), is used to demonstrate the model suggested in the previous section. The series covers the period from 1/Jan/88 to 31/Dec/98, consisting of 2869 observations. Table 3 lists some descriptive statistics.

Table 3. Descriptive Statistics for Daily SP500 Return Series

Mean	Standard Deviation	Skewness	Excess Kurtosis	Minimum	Maximum
0.0415	0.8643	-0.6643	7.9538	-7.1262	4.9748

The data are fitted to the five models that mainly differ in the shock's distribution. These are

- M0: standard normal distribution,
- M1: normal log-normal mixture distribution with $\rho = 0$,
- M2: normal log-normal mixture distribution,
- MT: Student's t distribution,
- MS: Skewed t distribution.

We note that M2, given by (10-14), nests M0 and M1. The models MT and MS are defined by (11), (12) and (14) with $c_1 = 0$, $c_2 = 1$, and v_t being the t distribution and the skewed t distribution respectively.

Table 4. Estimation Results for SP500 Return Series

Parameter	M0	M1	M2	MT	MS
μ	-.0293	-.0454	-.0458	-.0349	-.0366
	[.039]	[.063]	[.041]	[.043]	[.043]
δ	.0784	.1410	.1461	.1078	.1026
	[.053]	[.099]	[.065]	[.057]	[.057]
ω	-.0729	-.0843	-.0832	-.0767	-.0765
	[.019]	[.020]	[.019]	[.016]	[.016]
d	.4111	.5387	.5269	.3859	.3843
	[.088]	[.120]	[.103]	[.093]	[.092]
γ	-.5729	-.6151	-.6214	-.5834	-.5891
	[.131]	[.132]	[.126]	[.128]	[.128]
α	.0905	.0982	.0977	.0955	.0956
	[.022]	[.020]	[.019]	[.019]	[.019]
β	.9852	.9852	.9851	.9860	.9857
	[.008]	[.008]	[.007]	[.007]	[.007]
σ, ν		.8432	.8315	4.7701	4.8140
		[.056]	[.055]	[.474]	[.478]
ρ, φ			-.0483		-.0204
			[.027]		[.022]
LogL	-3391.1	-3231.3	-3230.0	-3235.2	-3234.9
Skewness	-0.8545	-0.7734	-0.7753	-0.8677	-0.8667
Kurtosis	6.5005	5.7888	5.7917	6.6671	6.6544
Q_{12}	18.6	18.1	18.2	18.5	18.5

The estimation results for these models are presented in Table 4. The standard errors, using the robust formula of White (1982), are given in the brackets.

First, the evidence for thick-tailed distributions appears strong. The large likelihood ratio of M1 against M0 (319.6) unambiguously favors the NLN mixture that is thick-tailed. The point estimate for σ implies that the estimate of CV for v_t, $\sqrt{\mathrm{Var}(e^{\eta_t})}/\mathrm{E}(e^{\eta_t}) = \sqrt{e^{\sigma^2} - 1}$, is 0.998 with standard error 0.091. The estimated degrees of freedom for MT and MS (approximately 4.8) also clearly indicate leptokurtosis in the shock's distribution.

Second, there is mild evidence supporting the leverage effect (or negative skewness in the shocks). The p-value of the one-tailed t-ratio for $\rho < 0$ is 0.037. Comparing M1 and M2, we note that the presence of ρ leads to smaller the standard errors. However, the skewness parameter (φ) in MS appears insignificant.

Third, there is some evidence for a positive ARCH-M effect ($\delta > 0$). For M0, MT and MS, the positive ARCH-M effect is significant at the 10% level if tested on one tail and significant at the 15% level if tested on two tails. For M1, the one-sided and two-sided p-values for δ are 0.077 and 0.154 respectively. Remarkably, for M2, the ARCH-M effect δ is significantly positive at the 1.5% level. The sharper result for the ARCH-M effect in M2 comes from the fact that the positive ARCH-M effect is separated from the negative leverage effect c_1 in the model. The implied point estimate for c_1 is $\hat{c}_1 = -0.0219$ with a standard error of 0.013 (calculated using "delta" method), which is significant at the 5% level on one-tail and at 10% on two-tails.

Fourth, the filtering function $\tau_b(\cdot)$ is effective in all models, in the sense that the parameter $d = 1/\sqrt{b}$ is significantly different from zero in all models.

As mentioned in the previous sub-section, a finite b (or non-zero d) not only ensures the covariance-stationarity of x_t but also makes the λ_t robust to outliers. For example, the point estimates for d in M2 is 0.5269, implying that the impact of shocks outside 3.602 (point estimates for b) on the log-variance process in M2 is reduced.

Fifth, the stylized GARCH modelling results in the literature are also observed in Table 4. The β estimates are all close to unity, signalling strong persistence in the variance process. The asymmetric effects of shocks (negative γ) are also significantly visible in all models. Further, the positive estimates of $1 \pm \gamma$ and α capture the clustering effect in volatilities.

All models fit the data reasonably well, in the sense that the serial correlation in the standardized residuals ($\hat{v}_t$) appears insignificant. In all models, none of the estimated autocorrelations (up to 50 lags) in $\hat{v}_t$ is significant. The Ljung-Box Q statistics also indicate the absence of autocorrelation in $\hat{v}_t$ if the degrees of freedom for the χ^2 distribution are set as the number of squared autocorrelations in Q. However, if the degrees of freedom are adjusted for the number of estimated parameters, the Q statistics reject at 5% the null that the first k autocorrelations are zero for $k \leq 25$ but cannot reject the null for $k > 25$. The Q_{12} statistics are given in Table 4. Adding the lagged return x_{t-1} in the mean function (its coefficient is not significant for all models) does not qualitatively alter the results in Table 4. The above exercises are also carried out for $\hat{v}_t^2$, $\hat{v}_t^3$ and $\hat{v}_t^4$ and no evidence for autocorrelation can be found, and Q statistics are below the 5% critical values for all k in all models with the adjusted degrees of freedom.

The NLN distribution (at the point estimates of σ and ρ) and the sample distribution of the standardized residuals from M2 are plotted in Figure 3. The fit of M2 appears reasonable in Figure 3.

The sample skewness and excess kurtosis for the standardized residuals in Table 4 show similar magnitudes for all models. The skewness and excess kurtosis implied by the point estimates in the NLN mixture model (M2) are -0.1788 and 3.0362 respectively, which do not match their sample counterparts. Again, it is not clear whether matching these is a desirable criterion, given that the sample skewness and kurtosis are sensitive to outliers (Kim and White, 2004).

The NLN mixture model (M2) were also estimated with the EGARCH(1,2) and EGARCH(2,1) lag-specifications in the variance process. Judged by the SBIC criterion, the EGARCH(1,1) specification (as reported in Table 4) is preferable. Further, the results from these different lag structures are qualitatively identical to those in Table 4.

Koopman and Uspensky (2002) analyzed this data set, using the stochastic volatility in mean (SVM) model and the GARCH-M model. Since their SVM model markedly differs from the model presented in this section (for example, the contemporaneous variance of x_t was included in their mean equation), their results are not directly comparable with those in Table 4. However, the positive ARCH-M effect in Table 4 is consistent with their findings in the GARCH-M model.

6 Conclusion

We investigate the properties of the normal log-normal (NLN) mixture and demonstrate its applications. The moment properties given here are useful for establishing covariance stationarity for EGARCH type processes. The NLN mixture is particularly useful in the cases where data are skewed and leptokurtic and in the cases where data are heterogeneous in variance. In the cross-sectional example, the NLN mixture model appears able to capture the extent of variations in the variance of a heterogeneous data set. In the time-series context, the NLN mixture based model draws merits from both ARCH type models and SV type models. The resulting model is able to separate the positive ARCH-M effect from the negative leverage effect. We find some evidence, in a SP500 return series, supporting the two opposite effects. The NLN based model is easier than the SV models to estimate, but is harder than the ARCH type models to estimate because a uni-variate integration is required for each density evaluation. While the numerical integration method (used in this paper) appears effective, it is desirable to find more efficient methods for evaluating the NLN mixture density function.

7 Proofs

Proof of Proposition 1

This can be shown by writing $\varepsilon = (\rho/\sigma)\eta + \xi$ with $\xi \sim \mathrm{N}(0, (1-\rho^2))$ being independent of η. Then, $E|u|^k = E|(\rho/\sigma)\eta e^{\frac{1}{2}\eta} + \xi e^{\frac{1}{2}\eta}|^k < \infty$ because $E|\xi^i \eta^j e^{\frac{1}{2}(i+j)\eta}| = E|\xi|^i E|\eta^j e^{\frac{1}{2}(i+j)\eta}| < \infty$ for any finite $i, j \geq 0$. $\square$

Proof of Proposition 2

(a) Because $\varepsilon = (\rho/\sigma)\eta + \xi$, $E(e^{au}|\eta) = \exp\{[\frac{a\rho}{\sigma}\eta e^{-\frac{1}{2}\eta} + \frac{1}{2}(1-\rho^2)a^2]e^{\eta}\}$. We note that the function $A(\eta) = \eta^k e^{-\delta\eta}$ has a unique maximum at $\eta^* = k/\delta$, $A(\eta^*) = (k/\delta)^k e^{-k}$, for $k \geq 0$ and $\delta > 0$. Then,

$$
\begin{aligned}
E(e^{au}) &= \int_{-\infty}^{\infty} E(e^{au}|\eta)(2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-\frac{\eta^2}{2\sigma^2}\} d\eta \\
&\geq \int_{0}^{\infty} E(e^{au}|\eta)(2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-\frac{\eta^2}{2\sigma^2}\} d\eta \\
&= (2\pi\sigma^2)^{-\frac{1}{2}} \int_0^{\infty} \exp\{[-\frac{\eta^2}{2\sigma^2}e^{-\eta} + \frac{a\rho}{\sigma}\eta e^{-\frac{1}{2}\eta} + \frac{1}{2}(1-\rho^2)a^2]e^{\eta}\} d\eta
\end{aligned}
$$

$$\geq \quad (2\pi\sigma^2)^{-\frac{1}{2}} \int_0^\infty \exp\{[-\frac{1}{2\sigma^2} - \frac{|a\rho|}{\sigma} + \frac{1}{2}(1-\rho^2)a^2]e^\eta\}d\eta,$$

where the final expression diverges to ∞.

(b) The proof of "if" part is trivial because both e^{av^+} and e^{av^-} are monotone and bounded between 0 and 1 if $a \leq 0$. The "only if" part is shown by noticing that, if $a > 0$,

$$\begin{aligned} E(e^{av^+}) &\geq E(e^{av}) = E(e^{a(u-c_1)/\sqrt{c_2}}), \\ E(e^{av^-}) &\geq E(e^{av}) = E(e^{a(u-c_1)/\sqrt{c_2}}), \end{aligned}$$

which diverge to infinity according to (a). □

Proof of Proposition 3

(a) Clearly,

$$E(e^{\tau(u)}) \quad = \quad \int_{|u|\leq b} e^{\tau(u)}\mathrm{pdf}_u du + \int_{|u|>b} e^{\tau(u)}\mathrm{pdf}_u du,$$

where the first term on the right-hand side (RHS)is finite. We claim that the second term on the RHS is also finite because

$$\ln(a_2 + a_3|u|) < \ln(\frac{|a_2|}{b} + a_3) + \ln|u|, \quad \text{for} \quad |u| > b$$

and the second term is bounded by

$$\int_{|u|>b} \exp\{a_0 + a_1 \ln(\frac{|a_2|}{b} + a_3)\}|u|^{a_1}\mathrm{pdf}_u du,$$

which by Proposition 1 is finite.

(b) We note that $d_1|\tau_b(v)| + d_1\tau_b(v) \leq (|d_1| + |d_2|)b$ for $|v| \leq b$ and

$$d_1|\tau_b(v)| + d_1\tau_b(v) \leq (|d_1| + |d_2|)[b + \ln(1 + |c_1|/\sqrt{c_2} + b + |u|/\sqrt{c_2})]$$

for $|v| > b$. As a function of u, $d_1|\tau_b(v)| + d_1\tau_b(v)$ satisfies the condition in (a) and the claimed finiteness follows. The existence of $E(e^{a\tau_b(v^+)})$ and $E(e^{a\tau_b(v^-)})$ can similarly be shown. □

8 References

Andrews, D.W.K. and W. Ploberger (1994), "Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica*, 62, 1383-1414

Andrews, D.W.K. and W. Ploberger (1995), "Admissibility of the likelihood ratio test when a nuisance parameter is present only under the alternative," *Annals of Statistics*, 23, 1609-1629

Black, F. (1976), "Studies of stock price volatility changes," *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, 177-181

Bollerslev, Tim (1986), "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 47, 244-280

Bollerslev, T., R.F. Engle and D.B. Nelson (1994), "ARCH models," R.F. Engle and D. McFadden (eds.), Handbook of Econometrics, vol. 4, North-Holland

Clark, P.K. (1973), "A subordinate stochastic process model with finite variance for speculative prices," *Econometrica*, 41, 135-155

Engle, Robert F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation," *Econometrica*, 50, 987-1008

Engle, R.F., D.M. Lilien and R.P. Robins (1987), "Estimating time-varying risk premia in the term structure: the ARCH-M model," *Econometrica*, 55, 391-407

Ghysels, E. A.C. Harvey and E. Renault (1996), Stochastic Volatility, *Handbook of Statistics* (Vol. 14), Statistical Methods in Finance, Maddala and Rao eds, North-Holland, Amsterdam

Hansen, Bruce E. (1994), "Autoregressive conditional density estimation," *International Economic Review*, 35, 705-730

Hsieh, D.A. (1989), "Modelling heteroskedasticity in daily foreign-exchange rates," *Journal of Business & Economic Statistics*, 7, 307-317

Kim, T.H. and H. White (2004), "On more robust estimation of skewness and kurtosis," *Finance Research Letters*, 1, 56-73

Koopman, S.J. and E.H. Uspensky (2002), "The stochastic volatility in mean model: empirical evidence from international stock markets," *Journal of Applied Econometrics*, 17, 667-689

Lee, S.W. and B. Hansen (1994), "Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator," *Econometric Theory*, 10, 29-52

Ling, S. and M. McAleer (2003), "Asymptotic theory for a vector ARMA-GARCH model," *Econometric Theory*, 19, 280-310f

Lumsdaine, R.L. (1996), "Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models," *Econometrica*, 64, 575-596

Nelson, Daniel B. (1991), "Conditional Heteroskedasticity in asset returns: a new approach," *Econometrica*, 59, 347-370

Tauchen, G. and M. Pitts (1983), "The price variability-volume relationship on speculative markets," *Econometrica*, 51, 485-505

Figure 1. Standardized Density Function $m(v|\sigma,\rho)$

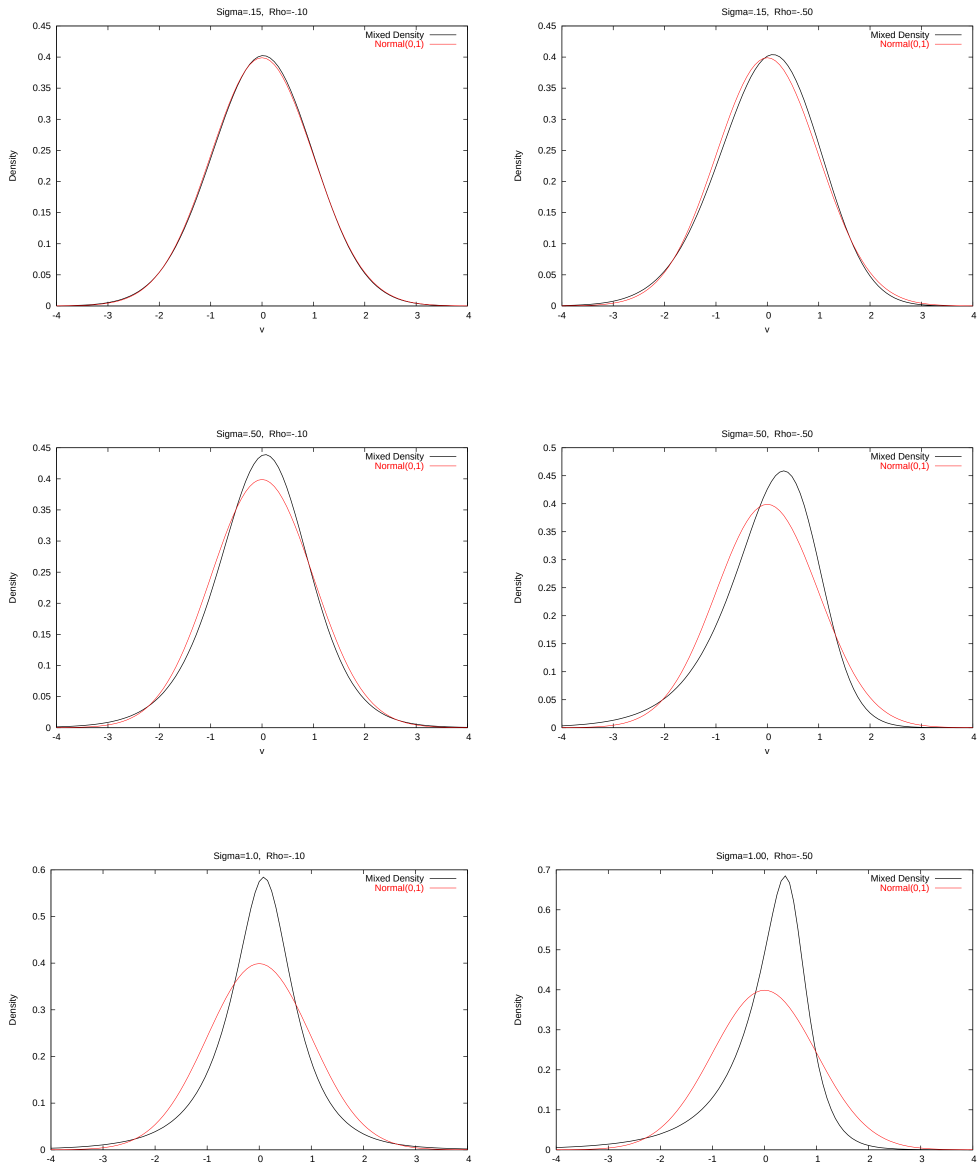
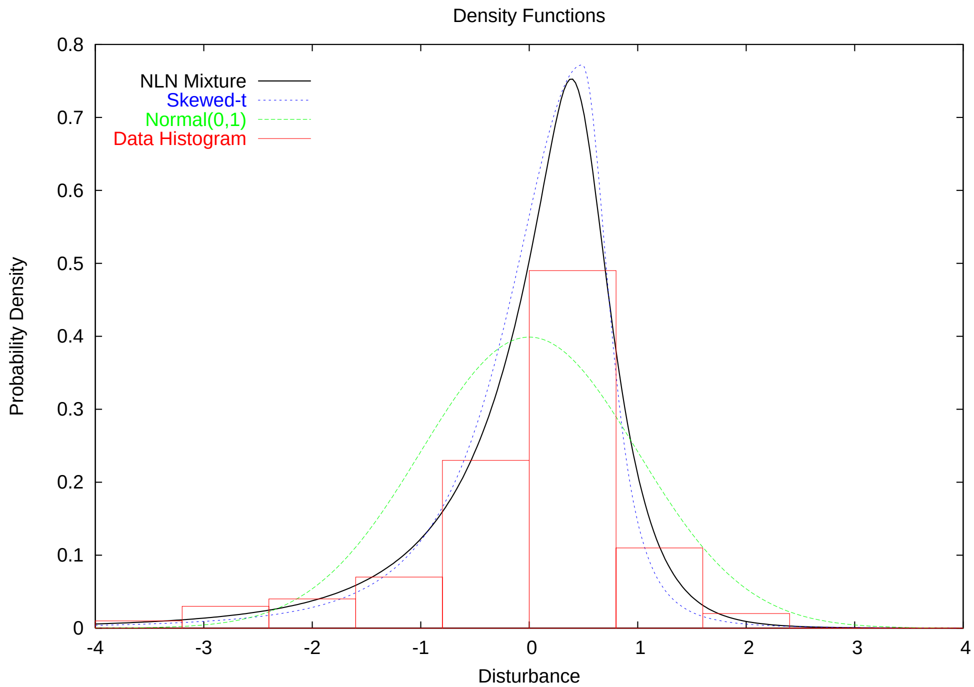


Figure 2. Distribution Comparison



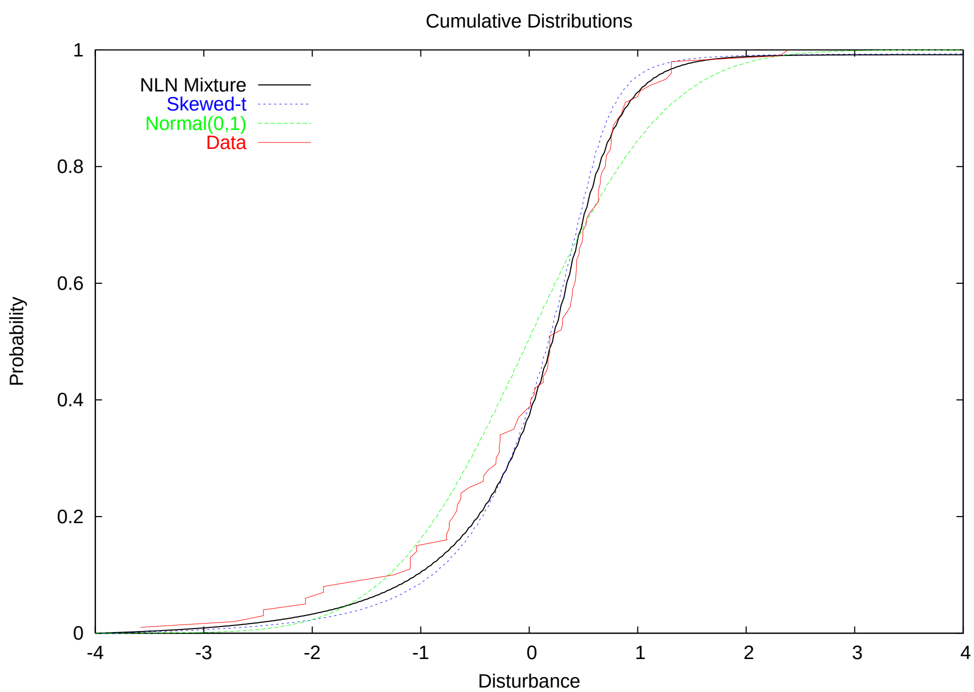


Figure 3. Distribution Comparison