Density forecasting of the Dow Jones share index

Öller, L-E and Stockhammar, P

Department of Statistics - Stockholm University

2009

Density forecasting of the Dow Jones share index

Pär Stockhammar and Lars-Erik Öller

Department of Statistics, Stockholm University
S-106 91 Stockholm, Sweden
E-mail: par.stockhammar@stat.su.se

Abstract

The distribution of differences in logarithms of the Dow Jones share index is compared to the normal (N), normal mixture (NM) and a weighted sum of a normal and an Assymetric Laplace distribution (NAL). It is found that the NAL fits best. We came to this result by studying samples with high, medium and low volatility, thus circumventing strong heteroscedasticity in the entire series. The NAL distribution also fitted economic growth, thus revealing a new analogy between financial data and real growth.

Keywords: Density forecasting, heteroscedasticity, mixed Normal- Asymmetric Laplace distribution, Method of Moments estimation, connection with economic growth.

1. Introduction

In some fields, including economic and financial practice, many series exhibit heteroscedasticity, asymmetry and leptokurtocity. Ways to account for these features have been suggested in the literature and also used in some applications. E.g. the Bank of England uses the Normal Mixture (NM) distribution when calculating interval and density forecasts of macroeconomic variables in the UK (Wallis, 1999). Another increasingly popular distribution to describe data with fatter than Normal tails is the Laplace (L) distribution. In the finance literature it has been applied to model interest rate data (Kozubowski and Podgórsky, 1999), currency exchange data (Kozubowski and Podgórsky, 2000), stock market returns (Madan and Senata, 1990) and option pricing (Madan et al., 1998), to name a few applications. Stockhammar and Öller (2008) showed that the L distribution may be too leptokurtic for economic growth. Allowing for asymmetry a mixed Normal-Asymmetric Laplace (NAL) distribution was proposed and in ibid. it was shown that the NAL distribution is more accurately describing the GDP growth data of the US, the UK and the G7 countries than the Normal (N), the NM and the L distributions. The convoluted version of Reed and Jorgensen (2004) was also examined, but proved inferior to the weighted sum of probabilities of the NAL.

In the present study, the density of the Dow Jones Industrial Average (DJIA) is investigated. This series is significantly skewed, leptokurtic and heteroscedastic. Diebold et al. (1998) showed that a MA(1)-t-GARCH$(1, 1)$ model is suitable to forecast the density of the heteroscedastic S&P 500 return series. Here another approach is employed. Instead of modeling the conditional variance, the data are divided into parts according to local volatility (each part being roughly homoscedastic). For every part we estimate and compare the density forecasting ability of the N, NM and the NAL distributions. If the NAL distribution would fit both share index data and GDP growth, this would hint at a new analogy between the financial sphere and the real economy.

This paper is organized as follows. Section 2 provides some theoretical underpinnings. The data are presented in Section 3 and a distributional discussion in Section 4. Section 5 contains the estimation set-up and a density forecasting accuracy comparison. Section 6 contains an illustrative example and Section 7 concludes.

2. Density forecast evaluation

The key tool in recent literature on density forecast evaluation is the probability integral transform (PIT). The PIT goes back at least to Rosenblatt (1952), with contributions by eg. Shepard (1994) and Diebold et al. (1998). The PIT is defined as

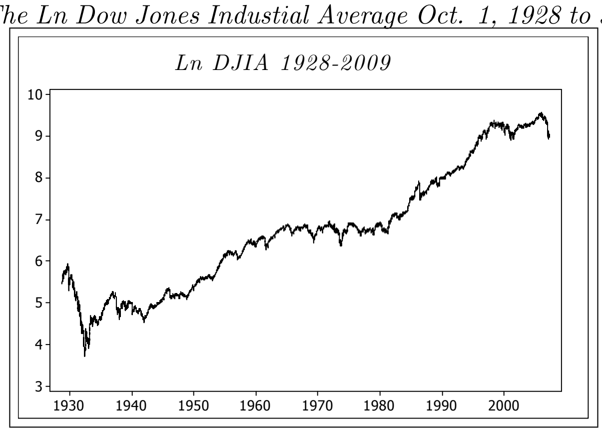
$$z_t = \int_{-\infty}^{y_t} p_t(u)du, \qquad (2.1)$$

where y_t is the realization of the process and $p_t(u)$ is the forecast density. If $p_t(u)$ equals the true density, $f_t(u)$, then z_t is simply the $U(0, 1)$ density. This suggests that we can evaluate density forecasts by assessing whether z_t, is i.i.d. $U(0, 1)$. This enables joint testing of both uniformity and independence in Section 4.

3. The data

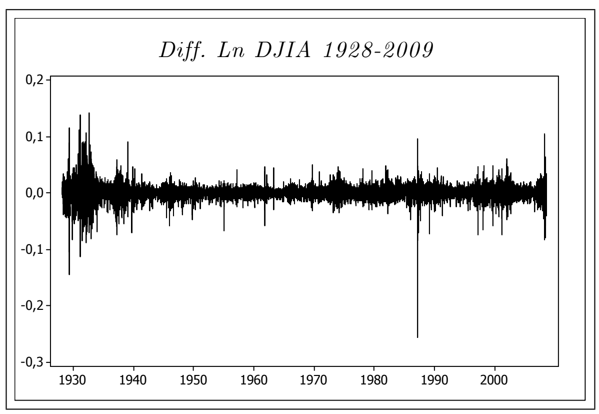
In this paper the Dow Jones Industrial average (daily closing prices) Oct. 1, 1928 to Jan. 31, 2009 (20 172 observations) is studied as appearing on the website www.finance.yahoo.com.

Figure 3.1: The Ln Dow Jones Industial Average Oct. 1, 1928 to Jan. 31, 2009



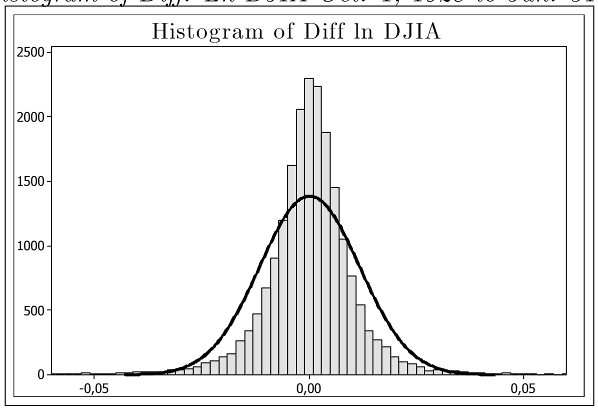
Taking the first difference of the logarithmic data reveals heteroscedasticity.

Figure 3.2: Diff. Ln Dow Jones Industial Average Oct. 1, 1928 to Jan. 31, 2009



This series is also significantly (negatively) skewed, leptokurtic and non-normal as indicated by Figure 3.3:

Figure 3.3: Histogram of Diff. Ln DJIA Oct. 1, 1928 to Jan. 31, 2009



where the solid line is the Normal distribution using the same mean and variance as in the series. The heteroscedasticity is even more evident in Figure 3.4, which shows moving standard deviations, smoothed with the Hodrick-Prescott (1980) filter (using smoothing parameter $\lambda = 1.6 \times 10^7$).

Figure 3.4: Moving standard deviations using window k=45 and $\lambda = 1.6 \times 10^7$

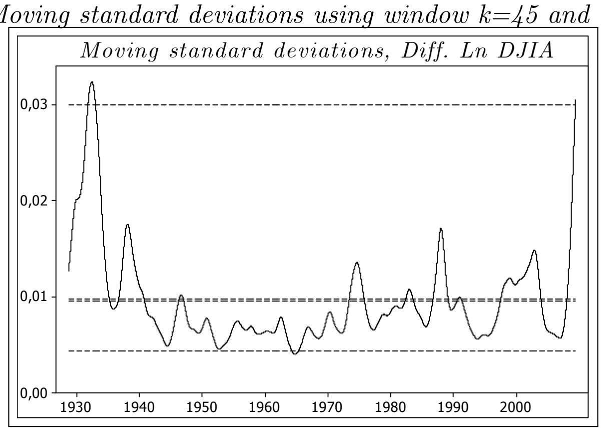


Figure 3.5 shows the distributions of the high (H), medium (M) and low (L) volatility observations, $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$. The periods of high, medium and low volatility are defined as times when the moving standard deviations, $\widehat{\sigma}_t$, (see Figure 3.4) are larger than 0.03,between 0.0095 and 0.0097, and smaller than 0.0044, respectively. These limits were chosen so as to get approximately equally-sized samples, for which in-sample variance is fairly constant. The three periods consist of 308, 267 and 277 observations, respectively.

Figure 3.5: The distributions of $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$

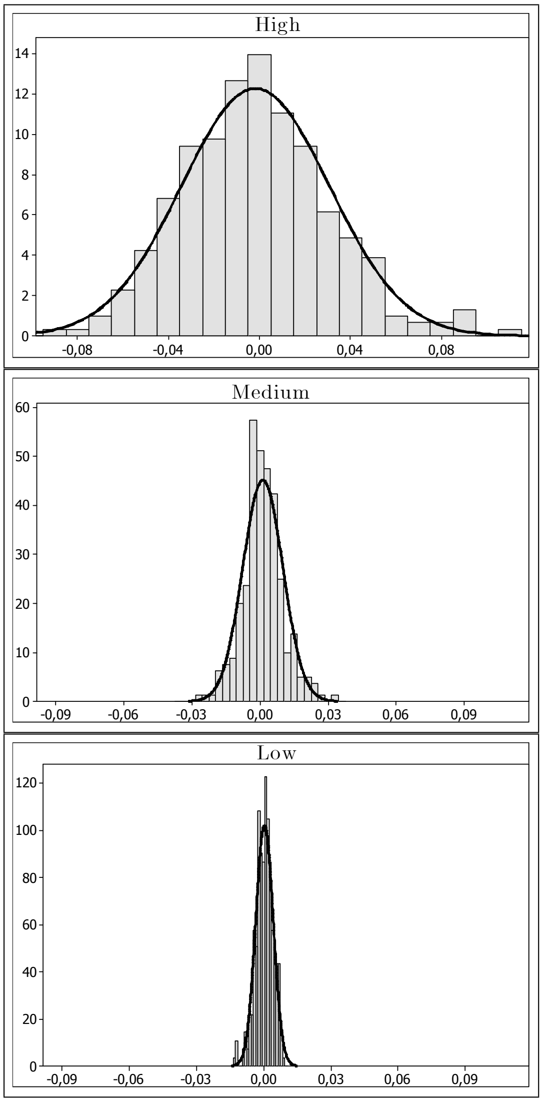


Table 3.1: The sample central moments of $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$ *and test of equality of the estimates*

				p-values ($\boldsymbol{\theta}=(\mu,\sigma,\tau,\kappa)$)		
	$y_{t,H}$	$y_{t,M}$	$y_{t,L}$	$H_{0,1}$:$\boldsymbol{\theta}_H=\boldsymbol{\theta}_M$	$H_{0,2}$:$\boldsymbol{\theta}_H=\boldsymbol{\theta}_L$	$H_{0,3}$:$\boldsymbol{\theta}_M=\boldsymbol{\theta}_L$
$\widehat{\mu}$	-0.00184	0.00121	0.00043	0.115	0.226	0.185
$\widehat{\sigma}$	0.03251	0.00884	0.00391	0.000($***$)	0.000($***$)	0.000($***$)
$\widehat{\tau}$	0.33	0.15	-0.47	0.107	0.000($***$)	0.000($***$)
$\widehat{\kappa}$	0.35	0.98	0.54	0.014($**$)	0.233	0.069($*$)

where $*$, $**$ and $***$ represent significance at the 10%, 5% and 1% levels, respectively

Table 3.2: The sample noncentral moments of $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$

	$y_{t,H}$	$y_{t,M}$	$y_{t,L}$
$E(y_t)$	-0.00184	0.00121	0.00043
$E(y_t^2)$	0.0010570	0.0000793	0.0000154
$E(y_t^3)$	0.0000055	0.0000004	0.0000000
$E(y_t^4)$	0.0000036	0.0000000	0.0000000

As expected the variance is significantly different in the three samples. It was also found that $\widehat{\tau}_H, \widehat{\tau}_L, \widehat{\kappa}_M$, and $\widehat{\kappa}_L$ significantly differ from zero (not shown), reflecting non-normality of the samples. Note that the mean of $y_{t,H}$ is negative, the volatility thus tends to increase when DJIA declines. Asymmetry is also supported by the results in Table 3.3.

Table 3.3: Tests of $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$

	$y_{t,H}$	$y_{t,M}$	$y_{t,L}$
ARCH-LM			
Aug. Dickey-Fuller	$***$	$***$	$***$
Anderson-Darling		$***$	$*$
Shapiro-Wilk	$**$	$***$	$***$
Kolmogorov-Smirnov			
Jarque-Bera	$*$	$**$	$***$

According to the ARCH-LM and the Dickey-Fuller test these series are both homoscedastic and stationary. We saw in Figure 3.3 that the N distribution is inappropriate to describe the shape of the entire DJIA series. The normality tests in Table 3.3 are based on very different measures and can therefore lead to different conclusions, as is the case here. All have low power, especially Kolmogorov-Smirnov, which fails to reject normality in all three samples. The remaining three tests reject normality in eight cases out of nine. For medium volatility data ($y_{t,M}$) there is little doubt about non-normality, c.f. Figure 3.5. But in order to vindicate the conclusions, we keep the Gaussian distribution as a benchmark. This distribution will be compared with the NM (as used by the bank of England) and the NAL distributions. That is the topic of the next Section.

4. Distributional discussion[1]

The use of different means and variances for the regimes enables introducing skewness and excess kurtosis in the NM distribution. The probability distribu-

[1] See Stockhammar and Öller (2008) for a more detailed description of the distributions.

tion function (pdf) of the NM distribution is:

$$f_{NM}(y_t;\boldsymbol{\theta}_1)=\frac{w}{\sigma_1\sqrt{2\pi}}\exp\left\{-\frac{(y_t-\mu_1)^2}{2\sigma_1^2}\right\}+\frac{1-w}{\sigma_2\sqrt{2\pi}}\exp\left\{-\frac{(y_t-\mu_2)^2}{2\sigma_2^2}\right\}, \tag{4.1}$$

where $\boldsymbol{\theta}_1$ consists of the parameters $(w,\mu_1,\mu_2,\sigma_1,\sigma_2)$ and where $0\le w\le 1$ is the weight parameter. Another distribution often used to describe fatter than normal tails is the L distribution. It arises as the difference between two exponential random variables with the same value on the parameter. The pdf of the L distribution is:

$$f_L(y_t;\boldsymbol{\theta}_2)=\frac{1}{2\phi}\exp\left\{-\frac{|y_t-\mu|}{\phi}\right\}, \tag{4.2}$$

where $\boldsymbol{\theta}_2=(\mu,\phi)$, $\mu\in\mathbb{R}$ is the location parameter and $\phi>0$ is the scale parameter. Again studying Figure 3.3 the L distribution seems promising. This is however misleading because of the significant skewness in the data. This is why we make use of the asymmetric Laplace (AL) distribution with pdf:

$$f_{AL}(y_t;\boldsymbol{\theta}_3)=\begin{cases}\frac{1}{2\psi}\exp\left\{\frac{y_t-\mu}{\psi}\right\} & \text{if } y_t\le\mu\\ \frac{1}{2\phi}\exp\left\{\frac{\mu-y_t}{\phi}\right\} & \text{if } y_t>\mu\end{cases}, \tag{4.3}$$

where $\boldsymbol{\theta}_3$ consists of the three parameters (μ,ϕ,ψ). The main advantage of the AL distribution is that it is skewed (except for the case $\psi=\phi$), conforming with the empirical evidence in Table 3.1. When $\psi\neq\phi$, this distribution has a discontinuity at μ. Another property of the AL distribution is that, unlike the pure L distribution, the kurtosis is not fixed. To further improve flexibility, Gaussian noise is added. To the author's best knowledge this distribution has not been used before for financial time series data. We assume that the probability density distribution of the diff Ln. Dow Jones series (y_t) can be described as a weighted sum of Normal and AL random shocks, i.e:

$$f_{NAL}(y_t;\boldsymbol{\theta}_4)=\frac{w}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(y_t-\mu)^2}{2\sigma^2}\right\}+(1-w)\begin{cases}\frac{1}{2\psi}\exp\left\{\frac{y_t-\mu}{\psi}\right\} & \text{if } y_t\le\mu\\ \frac{1}{2\phi}\exp\left\{\frac{\mu-y_t}{\phi}\right\} & \text{if } y_t>\mu\end{cases}, \tag{4.4}$$

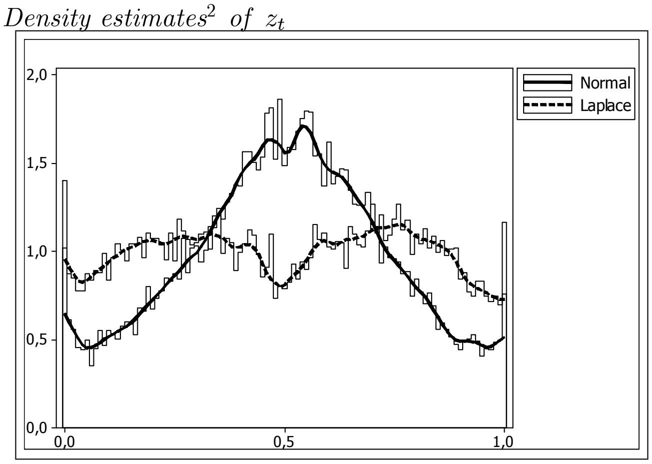
where $\boldsymbol{\theta}_4=(w,\mu,\sigma,\phi,\psi)$. Distribution (4.4) is referred to as the mixed Normal-asymmetric Laplace (NAL) distribution. Note that equal means but unequal variances are assumed for the components.

A graphical examination of the PIT histograms (see Section 2) might serve as a first guide when determining the density forecasting accuracy of the above distributions. One intuitive way to assess uniformity is to test whether the empirical cumulative distribution function (cdf) of $\{z_t\}$ is significantly different from the 45° line (the theoretical cdf). This is done using eg. the Kolmogorov-Smirnov (K-S) statistic or χ^2-tests.

Assessing whether z_t is i.i.d. can be made visually by examining the correlogram of $\{z_t - \overline{z}\}^i$ and the Bartlett confidence intervals. We examine not only the correlogram of $\{z_t - \overline{z}\}$ but also check for autocorrelations in higher moments. Here $i = 1, 2, 3$ and 4, which will reveal dependence in the (conditional) mean, variance, skewness and kurtosis. This way to evaluate density forecasts was advocated by Diebold et al. (1998).

In order to illustrate why the NAL distribution (4.4) is a plausible choice we once more study the entire series. Figure 4.1 shows the contours of calculated PIT histograms together with Kernel estimates for the L and the cumulative benchmark N distribution.

Figure 4.1 Density estimates[2] *of* z_t

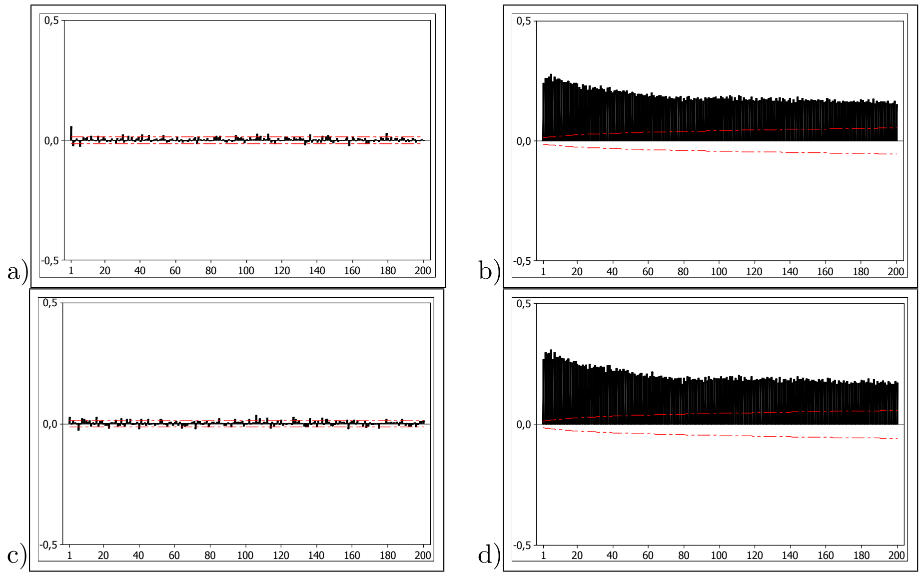


The N histogram has a distinct non-uniform "moustache" shape – a hump in the middle and upturns on both sides. This indicates that too many of the realizations fall in the middle and in the tails, relative to what we would expect if the data were N. The "seagull" shape of the L histogram is flatter than that of N, but is nevertheless non-uniform. The L histogram is the complete opposite of the N histogram with too few observations in the middle and in the tails.

Neither of the two distibutions is appropriate to use as forecast density function, but it may be possible to find a suitable weighted average of them as defined in (4.4). However assessing whether z_t is i.i.d. shows the disadvantages with the above models. Neither of them is particularly suitable to describe heteroscedastic data (such as the entire Diff. Ln series), see Figures 4.2 a-d) of the correlograms of $\{z_t - \overline{z}\}^i$ using the N distribution as forecast density.

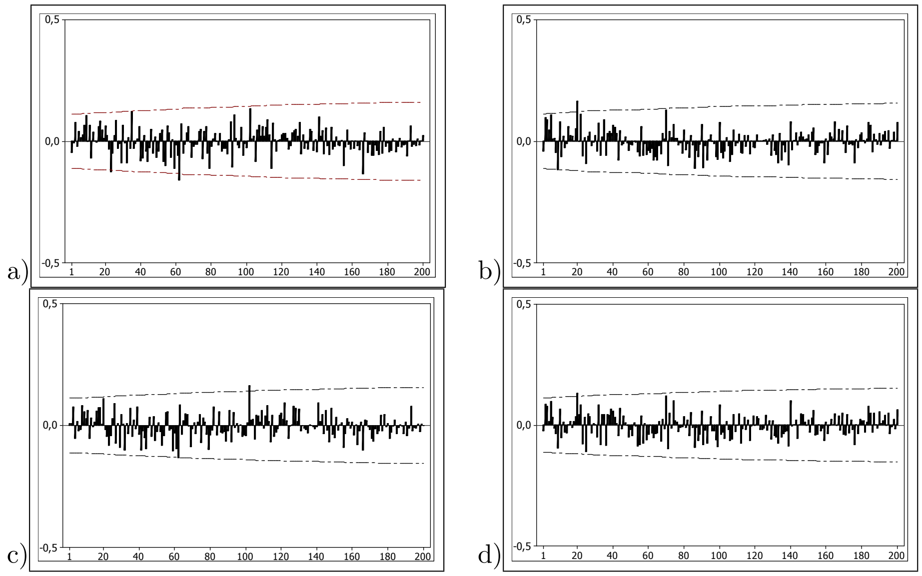
[2] 100 bins were used. If the forecast density were true we would expect one percent of the observations in each of the 100 classes, with a standard error of 0.0295 percent.

Figure 4.2: Estimates of the acf of $\{z_t - \overline{z}\}^i$, $i = 1, 2, 3$ and 4, for y_t *assuming normality*



The strong serial correlation in $\{z_t - \overline{z}\}^2$ and $\{z_t - \overline{z}\}^4$ (panels b and d) shows another key deficiency of using the N density – it fails to capture the volatility dynamics in the process. Also, the L correlograms indicate neglected volatility dynamics. This was expected. Neither single (N, L), nor mixed distributions (NM, NAL) are able to capture the volatility dynamics in the process. One could model the conditional variance using e.g. GARCH type models (as in Diebold et al., 1998), or State Space exponential smoothing methods, see Hyndman et.al (2008). Here we are more interested in finding an appropriate distribution to describe the data. Instead of modeling the conditional variance, the data are divided into three parts according to their local volatility (each of which is homoscedastic, see Table 3.2). Figure 4.3 further supports the homoscedasticity assumption in the high volatility data ($y_{t,H}$).

Figure 4.3: Estimates of the acf of $\{z_t - \overline{z}\}^i$, $i = 1, 2, 3$ and 4, for $y_{t,H}$ assuming normality



The series of medium and low volatility assuming the L, NM and NAL distributions give similar ACF:s. Standard tests do not signal autocorrelation in these series assuming any of the distributions. This means that our demand for independence is satisfied, and finding the most suitable distribution for density forecasts is a matter of finding the distribution with the most uniform PIT histogram. This is done using the K-S and χ^2 tests for $y_{t,H}$, $y_{t,M}$ and $y_{t,L}$ separately, when the parameters have first been estimated. These are issues of the next Section.

5. Estimation

The parameters are here estimated for the three periods of high, medium and low volatility respectively. For each part, the five parameters in the NM and NAL distributions (4.1 and 4.4) will be estimated using the method of moments (MM) for the first four moments. The noncentral and central moments and the cumulative distribution function (cdf) of (4.1) and (4.4) were derived in Stockhammar and Öller (2008). Equating the theoretical and the observed first four moments using the five parameters yields infinitely many solutions. A way around this dilemma is to fix μ_1 in the NM to be equal to the observed mode, which is here approximated by the maximum value of Kernel function of the empirical distribution ($\max f_K(y_i^*)$). Here $\widehat{\mu}_{1,H}$, $\widehat{\mu}_{1,M}$ and $\widehat{\mu}_{1,L}$ are subtituted for $\max f_K(y_{t,H}) = -0.0025$, $\max f_K(y_{t,M}) = -0.0001$ and $\max f_K(y_{t,L}) = 0.0011$. In the NAL μ is fixed to be equal to the MLE with respect to μ in the AL distribution, that is the observed median, $\widehat{md}$. Here $\widehat{\mu}_H = \widehat{md}_H = -0.00359$,

$\widehat{\mu}_M = \widehat{md}_M = 0.00081$ and $\widehat{\mu}_L = \widehat{md}_L = 0.00070$. Fixing one of the parameter in each distribution makes it easier to give guidelines to forecasters concerning which parameter values to use, and when. With the above parameters fixed, the NM and NAL parameter values that satisfy the moment conditions are:

Table 5.1: Parameter estimates

	$\mathrm{NM_H}$	$\mathrm{NM_M}$	$\mathrm{NM_L}$		$\mathrm{NAL_H}$	$\mathrm{NAL_M}$	$\mathrm{NAL_L}$
$\widehat{w}$	0.8312	0.7803	0.7898	$\widehat{w}$	0.8447	0.7651	0.7994
$\widehat{\mu}_2$	0.0141	0.0059	−0.0021	$\widehat{\sigma}$	0.0292	0.0091	0.0041
$\widehat{\sigma}_1$	0.0229	0.0081	0.0041	$\widehat{\psi}$	0.0365	0.0036	0.0042
$\widehat{\sigma}_2$	0.0604	0.0098	0.0011	$\widehat{\phi}$	0.0563	0.0070	0.0015

Note that the estimated weights in all cases are close to 0.8. To further improve user-friendliness, it is tempting to also fix the weights to that value. If this can be done without losing too much in accuracy it is worth further consideration. With $w = 0.8$ (and the μ´s fixed as above), the remaining three MM estimates are:

Table 5.2: Parameter estimates

	$\mathrm{NM_H}$	$\mathrm{NM_M}$	$\mathrm{NM_L}$		$\mathrm{NAL_H}$	$\mathrm{NAL_M}$	$\mathrm{NAL_L}$
$\widehat{\mu}_2$	0.0008	0.0065	−0.0023	$\widehat{\sigma}$	0.0321	0.0088	0.0041
$\widehat{\sigma}_1$	0.0217	0.0081	0.0040	$\widehat{\psi}$	0.0137	0.0040	0.0042
$\widehat{\sigma}_2$	0.0582	0.0097	0.0018	$\widehat{\phi}$	0.0312	0.0079	0.0015

Table 5.2 shows that not much happens if we fix w. The exception is for the NAL estimates of high volatility data, where both the magnitude and the ratio of $\widehat{\psi}$ to $\widehat{\phi}$ changes dramatically. Giving less weight to the N distribution is compensated for by a larger $\widehat{\sigma}$ and decreasing $\widehat{\psi}$ and $\widehat{\phi}$ and vice versa. Because of the strong positive skewness in $y_{t,H}$, $y_{t,M}$, $\widehat{\psi}_H < \widehat{\phi}_H$ and $\widehat{\psi}_M < \widehat{\phi}_M$. That $\widehat{\psi}_L > \widehat{\phi}_L$ accords well with the results in Table 3.1. Note that $y_{t,H}$ and $y_{t,L}$ have completely opposite properties in Table 3.1, $y_{t,H}$ having a mean below zero and positive skewness and the other way around for $y_{t,L}$. The relative difference between $\widehat{\psi}$ and $\widehat{\phi}$ is approximately the same in Table 5.2. $y_{t,M}$ shows yet another pattern with above zero mean and positive skewness ($\widehat{\psi}$ about half the value of $\widehat{\phi}$).

In order to compare the distributional accuracy of the above empirical distributions we make use of the K-S test. Because of the low power of this test, as with all goodness of fit tests, this is supplemented with χ^2 tests. The K-S test statistic (D) is defined as

$$D = \sup \left| F_E\left(x\right) - F_H(x) \right|,$$

where $F_E(x)$ and $F_H(x)$ are the empirical and hypothetical or theoretical distribution functions, respectively. Note that $F_E(x)$ is a step function that takes a step of height $\frac{1}{n}$ at each observation. The D statistic can be computed as

$$D = \max_i \left(\frac{i}{n} - F(x_i),\ F(x_i) - \frac{i-1}{n} \right),$$

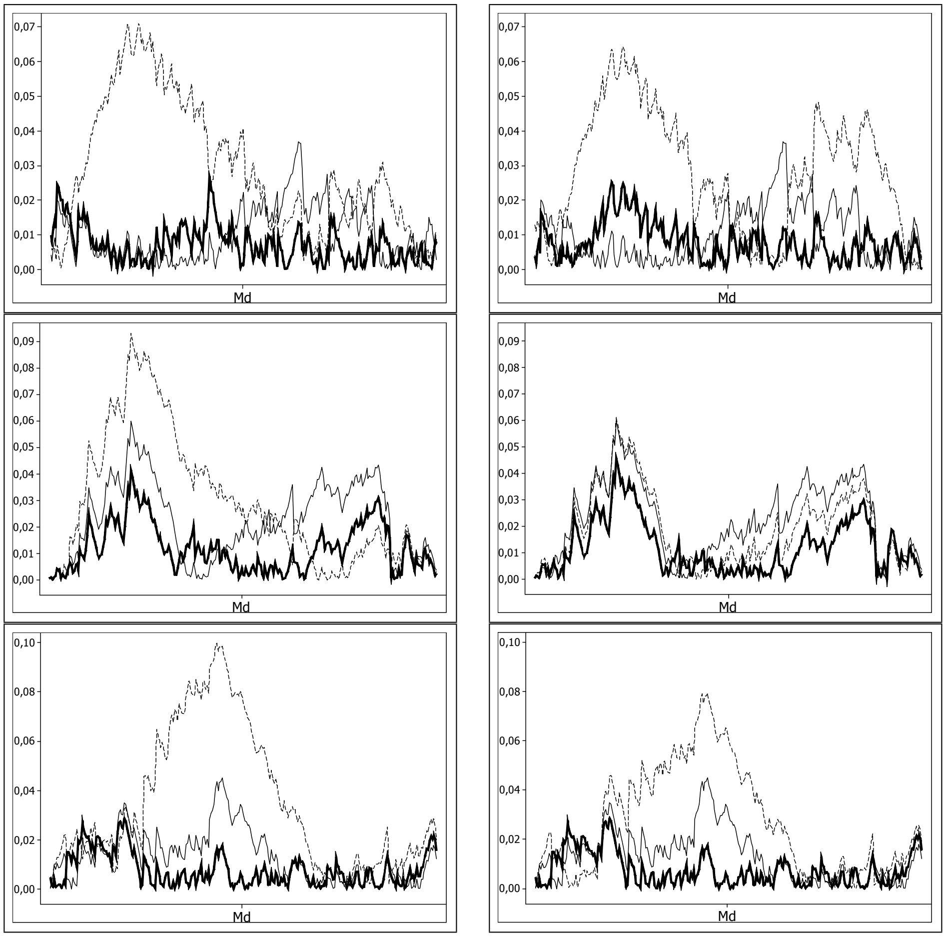
where we have made use of the PIT (2.1) and ordered the values in increasing order to get $F(x_i)$. If $F_E(x)$ is the true distribution function, the random variable $F(x_i)$ is $U(0,1)$ distributed. Table 5.3 reports the approximate p-values of the K-S test together with the value of the D statistics (in paranthesis), and also the p-values of the χ^2 test using 10 and 20 bins when testing $H_{0,1}$: $y_{t,k}$ ~N, $H_{0,2}$: $y_{t,k}$ ~NM$^{(1)}$, $H_{0,3}$: $y_{t,k}$ ~NM$^{(2)}$, $H_{0,4}$: $y_{t,k}$ ~NAL$^{(1)}$ and $H_{0,5}$: $y_{t,k}$ ~NAL$^{(2)}$ (k =High, Mid and Low). NM$^{(1)}$ and NAL$^{(1)}$ are based on the parameter estimates in Table 5.1 while NM$^{(2)}$ and NAL$^{(2)}$ are based on the estimates in Table 5.2.

Table 5.3: Goodness of fit tests

		High	Medium	Low
$H_{0,1} : y_{t,k} \sim$ N	K-S	$> 0.2(0.040)$	$> 0.2(0.060)$	$> 0.2(0.045)$
	$\chi^2(9)$	0.76	0.09	0.91
	$\chi^2(19)$	0.83	0.31	0.94
$H_{0,2} : y_{t,k} \sim$ NM$^{(1)}$	K-S	0.09(0.071)	0.02(0.092)	0.01(0.099)
	$\chi^2(9)$	0.06	0.04	0.01
	$\chi^2(19)$	0.26	0.13	0.03
$H_{0,3} : y_{t,k} \sim$ NM$^{(2)}$	K-S	0.16(0.064)	$> 0.2(0.062)$	0.06(0.079)
	$\chi^2(9)$	0.04	0.21	0.13
	$\chi^2(19)$	0.18	0.49	0.26
$H_{0,4} : y_{t,k} \sim$NAL$^{(1)}$	K-S	$> 0.2(0.026)$	$> 0.2(0.041)$	$> 0.2(0.028)$
	$\chi^2(9)$	0.98	0.74	0.92
	$\chi^2(19)$	1.00	0.47	0.90
$H_{0,5} : y_{t,k} \sim$NAL$^{(2)}$	K-S	$> 0.2(0.025)$	$> 0.2(0.046)$	$> 0.2(0.028)$
	$\chi^2(9)$	1.00	0.77	0.92
	$\chi^2(19)$	1.00	0.62	0.93

Table 5.3 shows that the NAL distributions are superior to the N and NM in every respect. Also, there is no great loss of information by fixing the weight parameter. In fact the NM fit was improved after fixing w, but the fit was nevertheless inferior to both the NAL and (surprisingly) the N distribution. The NM distributions (as employed by the Bank of England) thus have a relatively poor fit to the extreme volatility parts of diff. Ln DJIA. In general the N fit is, contrary to earlier results, quite good, particularly for the high and low volatility observations but, because of the significant skewness, the NAL fits even better. Figure 5.1 shows the absolute deviations of the empirical distribution functions of the probability integral transforms ($F(x_i)$) from the theoretical 45° lines (the measure the K-S test is based on).

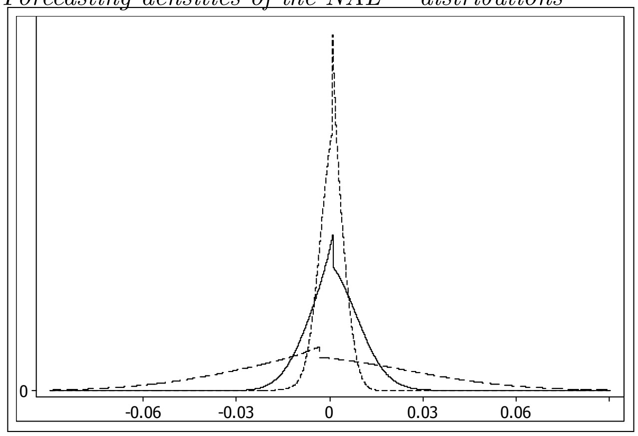
Figure 5.1: Absolute deviations of the N, $NM^{(1)}$,$NAL^{(1)}$ and N, $NM^{(2)}$, $NAL^{(2)}$ from the theoretical distributions



The N, NM and NAL distributions are marked with thin solid, dashed and thick solid lines, respectively, and the upper, centre and lower panels are the high, Mid and Low parts of the series. The panels to the left and right hand side are the distibutions in Table 5.1 and 5.2, respectively.

Figure 5.1 adds further information of the fit. The left tail fit is inferior to the right tail fit. This is particularly prominent for the NM. This conforms well with Bao and Lee (2006) who came to the same conclusion using various nonlinear models for the S&P daily closing returns. Except for the low volatility part the fit close to the median is generally rather good. Because of the similarity in distributional accuracy between the NAL$^{(1)}$ and NAL$^{(2)}$ the latter distribution is the obvious choice. With both μ and w fixed it is easier to interpret the remaining parameters. Figure 5.2 shows the forecast densities of the NAL$^{(2)}$ distribution for $y_{t,H}$ (dashed), $y_{t,M}$ (solid) and $y_{t,L}$ (dotted), respectively.

Figure 5.2: Forecasting densities of the NAL$^{(2)}$ distributions



Here a jump at the median of each distribution is evident.[3] The negative median in $y_{t,H}$ means that for high volatility data we expect a negative trend, but due to skewness, with large positive shocks being more frequent than large negative.

In a situation of a very large local variance, here defined as $\widehat{\sigma}_t > 0.03$ for the last 45 days, we propose the use the high volatility NAL distribution and the corresponding estimates in Table 5.2. Similarly we suggest to use the $\mathrm{NAL}_{\mathrm{M}}^{(2)}$ and $\mathrm{NAL}_{\mathrm{L}}^{(2)}$ estimates in Table 5.2 if the local variance falls between 0.0095 and 0.0097, or fall below 0.0044. For the intervening values a subjective choice is encouraged using the estimates in Table 5.2 as guidelines. During the world wide financial crises of 2008 and 2009 we would most often use the $\mathrm{NAL}_{\mathrm{H}}$ estimates (or values close to them). On the contrary we suggest the use of the $\mathrm{NAL}_{\mathrm{L}}$ estimates during calm, or "business as usual" periods. This is exemplified in the following Section.

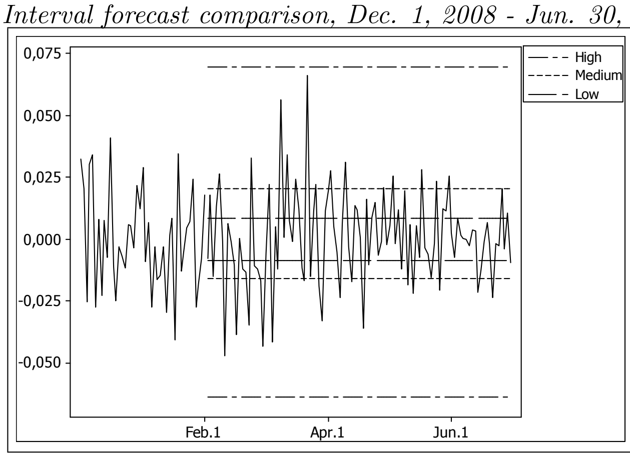
6. Application

The proposed density forecast method is here applied on the diff. Ln DJIA series Feb. 1, 2009 to Jun. 30, 2009, thus showing a realistic forecast scenario. According to Figure 3.4 the local volatility at the end of Jan 2009 is very large ($\widehat{\sigma}_t \approx 0.03$). Following the earlier discussion we should in this situation choose the $\mathrm{NAL}_{\mathrm{H}}^{(2)}$ distribution when calculating density forecasts, but to serve as comparisons we will also include the density forecasts made using the $\mathrm{NAL}_{\mathrm{M}}^{(2)}$ and $\mathrm{NAL}_{\mathrm{L}}^{(2)}$ distributions. We have used the (neutral) median in each distribution as point forecasts. Other models for the point forecasts could, and probably should, be used in real life practice. Figure 6.1 shows the original diff. Ln series

[3] The discontinuity at the median can be avoided using eg. the convoluted NAL version of Reed and Jorgensen (2004). Since this approach did not prove promising in Stockhammar and Öller (2008), we do not pursue it here.

Dec. 1, 2008 to Jun. 30, 2009 together with the 95 per cent confidence intervals for the point forecasts using the $NAL_H^{(2)}$, $NAL_M^{(2)}$ and $NAL_L^{(2)}$ distributions, calculated from Feb. 1, 2009.

Figure 6.1: Interval forecast comparison, Dec. 1, 2008 - Jun. 30, 2009



The forecasting horizon (5 months) in the above example is too long to be classified as a high volatility period. The corresponding distribution works best only for the first half of the period. For the later half it is probably better to use parameter values closer to the $NAL_M^{(2)}$ distribution.

7. Conclusions

In this paper we have looked at a way to deal with the asymmetric and heteroscedastic features of the DJIA. The heteroscedasticity problem is solved by dividing the data into volatility groups. A mixed Normal- Asymmetric Laplace (NAL) distribution is proposed to describe the data in each group. Comparing with the Normal and the Normal mixture distributions the NAL distributional fit is superior, making it a good choice for density forecasting Dow Jones share index data. On top of good fit of this distribution its simplicity is particularly desirable since it enables easy-to-use guidelines for the forecaster. Subjective choices of the parameter values is encouraged, using the given parameter values for scaling. The fact that the same distribution fits both share index data and GDP growth indicates a analogy between financial and growth data not known before. The NAL distribution was derived as a reduced form of a Schumpeterian model of growth, the driving mechanism for which was Poisson (Aghion and Howitt, 1992) distributed innovations plus Gaussian noise. Interestingly the same mechanisms seem to work with share index data.

References

Aghion, P. and Howitt, P. (1992) A model of growth through creative destruction. *Econometrica,* 60, 323-351.

Bao, Y. and Lee, T-H. (2006) Asymmetric predictive abilities of nonlinear models for stock returns: evidence from density forecast comparison. Econometric Analysis of Financial and Economic Time Series / Part B, Advances in Econometrics, 20, 41-62.

Diebold, F. X., Gunther, T. A. and Tay. A. S. (1998) Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39, 863-883.

Hodrick, R. J. and Prescott, E. C. (1980) Postwar U.S. business cycles: An empirical investigation. Discussion paper 451, Carnegie-Mellon University.

Hyndman, R. J., Koehler, A. B., Ord, J. K. and Snyder, R. D. (2008) *Forecasting with exponential smoothing.* Springer Verlag, Berlin.

Kozubowski, T. J. and Podgorski, K. (1999) A class of asymmetric distributions. *Actuarial Research Clideraring House*,1, 113-134.

Kozubowski, T. J. and Podgorski, K. (2000) Asymmetric Laplace distributions. *The Mathematical Scientist*, 25, 37-46.

Madan, D. B. and Senata, E. (1990) The Variance Gamma (V.G.) model for share market returns. *Journal of Business*, 63, 511-524.

Madan, D. B., Carr, P. and Chang, E. C. (1998) The variance gamma process and option pricing. *European Finance Review*, 2, 74-105.

Reed, W. J. and Jorgensen, M. A. (2004) The double Pareto-lognormal distribution - A new parametric model for size distributions. *Communications in Statistics: Theory and Methods,* 33(8), 1733-1753.

Rosenblatt, M. (1952) Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23, 470-472.

Shepard, N. (1994) Partial non-gaussian state space. *Biometrika*, 81, 115-131.

Stockhammar, P. and Öller, L-E. (2008) On the probability distribution of economic growth. *Research Report 2008:5*, Department of Statistics, Stockholm University.

Wallis, K. F. (1999) Asymmetric density forecasts of inflation and the Bank of England´s fan chart. *National Institute Economic Review*, 167, 106-112.