Stockholm University

Master thesis

Department of Statistics

Masteruppsats,Statistiska institutionen

Fitting probability distributions to economic growth a maximum likelihood approach

MAHMOOD UL HASSAN

Masteruppsats30högskolepoäng, vt 2014

Supervisor: Pär Stockhammar

Abstract

The growth rate of the gross domestic product (GDP) usually carries heteroscedasticity, asymmetry and fat-tails. In this study three important and significantly heteroscedastic GDP series are studied. A Normal (N), Normal-Mixture (NM), Normal-Asymmetric Laplace (NAL) distribution and a Student's t Asymmetric Laplace (TAL) distribution mixture are proposed for distributional fit comparison of GDP series after removing heteroscedasticity. The maximum likelihood method is used for estimation of the parameters of the distributions. Based on the results of different accuracy measures, goodness of fit tests and plots, we find that in the case of asymmetric, heteroscedastic and highly leptokurtic data the TAL-distribution fits better than the alternatives. In the case of asymmetric, heteroscedastic but less leptokurtic data the NM fit is superior. Further, a simulation study has been carried out to obtain standard errors for the estimated parameters.

Keywords: Mixed Normal-Asymmetric Laplace distribution, Mixed Student's t-Asymmetric Laplace distribution, Method of maximum likelihood estimation and the Nelder and Mead General Purpose optimization.

Acknowledgment

First of all I am thankful to Allah Almighty that enabled me to bring my effort in fascinating manners and the messenger of Allah, Muhammad (Peace Be Upon Him) the educator of mankind.

I would like to express my greatest gratitude to the people who have helped and supported me. I am grateful to my advisor Pär Stockholmar for the continuous support, his patience, motivation, and immense knowledge.

A special thank goes to my friends, Abrar Raza Khan and Muhammad Shoaib for all their help, motivation and appreciation.

Table of Contents

List of figures

List of tables

1 Introduction

The Gross Domestic Product(GDP), the market value of all officially recognized final goods and services produced within a country in a given period of time, has historically been considered a measure of economic growth. Later it was adjusted for the changing population, and consequently, GDP per capita (GDP divided by total population) became a widely used measure of economic growth and standard of living. Though the concept of economic growth has evolved greatly over the decades, it still contains income, or in other words GDP, as one of the most important elements, and an indicator of growth. All else equal (particularly population), a rising GDP definitely refers to increased production of goods and services for a country's population.

Economic growth shows typical irregular patterns in the short and in the long run. GDP growth rates contain fat-tails (large kurtosis) and heteroscedasticity (see, e.g., Barro 1991; Canning et al. 1998; Lee et al. 1998; Fagiolo et al. 2008; Castaldi and Dosi 2009; Bottazzi and Duenas 2012). Behavior of GDP growth series is asymmetric, which is both expected and empirically confirmed (see, e.g., Hess and Iwata 1997; Bodman 1998; Narayan 2009). The general conclusion from the above studies is that growth rate of GDP growth is heteroscedastic, and asymmetric and leptokurtic.

The quarterly increase in GDP percentage here reflects the GDP growth rate. It indicates at which pace a country's economy is growing. Thus, accurate density distributions are required to forecast economic growth known as 'density forecast'. A density forecast is an estimate of the probability distribution of the possible future outcomes of the variable. It provides us with a complete description of uncertainty, associated with the forecast. Interval forecast is the difference between two extremes, which indicates the probability that actual outcome will fall within a stated interval. Density distribution for the growth of GDP is the primary objective of this study. Density forecasting is rapidly getting more attention in the field of economic and financial time series (see, e.g., Diebold et al. 1998; Tay and Wallis 2002). Heteroscedasticity affects the estimates of parameters. In order to find the correct density distribution it is important to filter the data for heteroscedasticity. This is done by using the filter proposed by Stockhammar and Öller (2011). After the filtering, the series becomes homoscedastic, but the asymmetry and leptokurtic still remains.

The Normal Mixture (NM) distribution is widely used in empirical finance and has a long history of its application in various fields, which include astronomy, biology, economics, finance and engineering. Applications of the NM distribution in different fields are well documented in Everitt and Hand (1981), Titterington, Smith and Makov (1985), McLachlan and Peel (2000), Schlattmann (2009) and Mengersen et al. (2011). The NM distribution is able to capture the leptokurtic, asymmetric and multimodal characteristics of any time series data. Newcomb (1963) was first to use a NM distribution to handle a fat tail. Gridgeman (1970) proved that when the regimes had the same mean, NM would be leptokurtic. The NM distribution has a long history in the modelling of asset returns (See, e.g., Press 1967; Praetz 1972; Clark 1973; Blattberg and Gonedes 1974; Kon 1984). The finding that skewness and leptokurtosis can be introduced by varying the parameters, was used as early as the late nineteenth century by e.g. Pearson (1895). Kamaruzzaman et al. (2012) used two components in the NM distribution for financial time series, and found that the NM distribution captures the leptokurtic as well as skewness in the data. So the NM distribution could be used to model the growth data.

The Laplace (L) distribution is symmetric around its mean and it is a well-known and widely used symmetric distribution for modeling data with heavier tails than the normal distribution. The L distribution is not appropriate for modeling asymmetric data. For such cases, a skewed generalization of L distribution is considered appropriate. In the last several decades, several forms of asymmetric Laplace (AL) distribution has been introduced (for more details on AL distributions see Kozubowski and Nadaarjah (2010). AL distributions have been applied in analyses of currency exchange rates, stock price changes, interest rates, daily financial market series, economics and marketing data etc. (see Kozubowski and Podgorski 1999 and 2000; Linden 2001; Kozubowski and Nadaarjah 2010).

It was found that the excess kurtosis in AL models is too large for the filtered (and unfiltered) growth series. Stockhammar and Öller (2011) added Gaussian noise to the AL-distributionand introduced the Normal-AL (NAL) distribution. It was partly based on a Schumpeterian theory of economic growth. According to Schumpeter, modern economies share certain internal factors that determine their growth. He holds the opinion that the R & D investments, aimed at creating new and better products, are the main factors that lead endogenously to economic growth. The NAL distribution is capable of capturing a wide range of skewness and kurtosis.

Here student’s t distributed noise is added to the AL-distribution to account for the excess kurtosis of AL. The standard student’s t and normal distributions are special cases of student’s t distribution. The AL distribution is combined with Student’s t distribution leading to the weighted mixed Student’s t-AL (TAL) distribution. The TAL distribution is capable of generating a wide range of skewness and kurtosis, making the model very flexible. A mixture distribution is a suitable for data that are divided into natural groups. Introduction to mixture distributions, as well as further detail on the theory, parameter estimation methods and applications can be found in Everitt and Hand (1981), Titterington et al. (1985), McLachlan and Basford (1988), Lindsay (1995), McLachlan and Peel (2000), Frühwirth-Schnatter (2006), Schlattmann (2009) and Mengersen et al. (2011). The Mixture distribution parameters are estimated using the maximum likelihood (ML) method.

This thesis is structured as follows. The data is presented in section 2. Section 3 is about data preparation. In section 4 a model discussion is presented along with the proposed model. Section 5 contains the estimation set-up with maximum likelihood estimates (MLE) and a distributional accuracy comparison. Section 6 concludes the thesis.

2 The data

Quarterly and seasonally adjusted GDP series of three countries US (1947-2012), UK (1955-2012) and CA (Canada) (1961-2012) are studied in this thesis. Data have been taken from the websites of Bureau of Economic Analysis (www.bea.gov), UK National Statistics (www.statistics.gov.uk) and of Statistics Canada (www.statcan.gc.ca), respectively.

Several time series variables like GDP per capita, population size, total consumption, etc. have shown a tendency to grow exponentially. We remove the trend by taking differences of logarithms of these GDP series which represents growth rate of GDP. Long series are required for accurate estimation of the N, NM, NAL and TAL parameters. The above series are the most important and longest quarterly GDP series available. The first logarithmic differences of series and their frequency distributions are presented in figure 2.1. Moreover, an estimation of a [1]Kernel density and Normal distribution with mean and variance of the first logarithmic differences of series are shown.

---

[1]The Kernel estimate is defined as

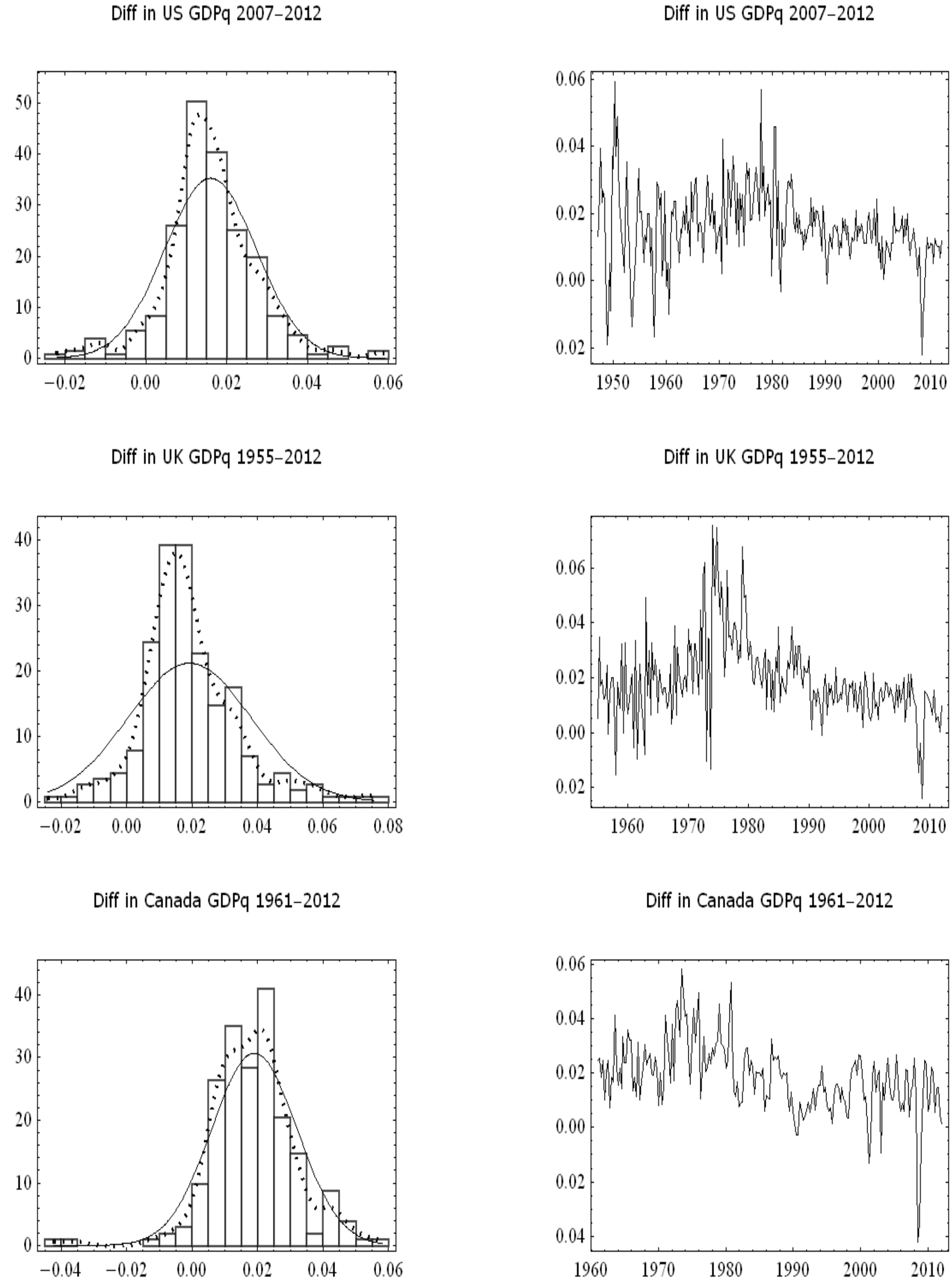
$$\hat{f}(y,h) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{y-y_i}{h}\right)$$

where k(.) is the kernel function and h is the bandwidth parameter. In this study we have used the Gaussian Kernel, $K(\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\mu^2}{2}}$, and the Silverman (1986) "Rule of Thumb" bandwidth

$$\hat{h} = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{\frac{1}{5}} \approx 1.059\hat{\sigma} n^{-\frac{1}{5}}$$

which is considered to be optimal when data are close to normal as the case here.

Figure 2.1: Shows the first difference of the logarithmic GDP series. The panels on the left hand side show the distribution of the data. The solid line represents the normal distribution and the dotted line is the Kernel distribution.



The first difference of the log GDP series appears to be leptokurtic. This is also confirmed in Table 2.1. The excess kurtosis exceeds zero. The results of ARCH-LM test for heteroscedasticity and Augmented Dickey Fuller test for stationarity are also presented in the table.

Table 2.1: The moments and the results of unit root and homocedasticity tests of the difference of the logrithemic GDP series of US, UK and CA

Statistics	US	UK	CA
Mean	0.0160	0.0192	0.0185
STD	0.0113	0.0188	0.0130
Skewness	0.1484	0.8520	-0.4000
Kurtosis	2.1297	2.1664	3.3226
ADF (p-value)	0.0000	0.0059	0.0000
ARCH-LM (p-value)	0.0000	0.0000	0.0000

The skewness seems to be non-zero in the UK and CA series. High kurtosis appears in all series as excess kurtosis in all cases exceeds zero. The ARCH-LM test rejected the null hypothesis of homoscedasticity in all series with a p-value of 0.0000. Heteroscedasticity implies an unequal weighting of the observations leading to inefficient parameter estimates. The Augmented Dickey-Fuller test is also rejecting the null hypothesis for a unit root in the difference logrithemic GDP series with p-value of 0.0000. Heteroscedasticity affects the estimates of parameters and most time series models require stationarity. The heteroscedasticity must be removed to compare the distributions of data. In order to make a fair comparison between the frequency distributions of the three series and various probability distributions, the filter proposed by Stockhammar and Öller (2011) is used. This filter enables us to work with mean and variance stationary time series.

3 Data preparation

The Hodrick and Prescott (HP) filter (1997) is a popular tool which decomposes a given macroeconomic time series into a non-stationary growth component and a stationary cyclical component. The HP filter was designed to analyze postwar US business cycles, as opposed to the smoothing methods used for inventory and production data. Let x_t be a seasonally adjusted time series, and let the decomposition of x_t into an unobserved trend component g_t and an unobserved cyclical component c_t at time t be

$$x_t = g_t + c_t \qquad for \quad t = 1, 2, 3, ..., T$$

The HP filter is defined as the solution to the following minimization problem

$$g_{\min} := \min_{[g_t]_{t=1}^{T}} \left\{ \sum_{t=1}^{T} c_t^2 + \gamma \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right\} \qquad (3.1)$$

where $c_t = x_t - g_t$, $0 \le \gamma \le \infty$, $\Delta^2 g_t = [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2$ and $g_{\min}$ is the HP filter. The first sum of (3.1) accounts for the accuracy of the estimation, while the second sum represents the smoothness of the trend. The second sum, ($\Delta^2 g_t$), is the square of the trend components, g_t second differences at time t. The smoothness parameter γ is a positive number which penalizes the variability in the growth component series. The larger the value of γ the smoother is the solution series and vice versa. Hodrick and Prescott (1997) recommended a value of $\gamma = 1600$ for quarterly data.

Stockhammar and Öller (2011) proposed a new filter for removing the heteroscedasticity from the data by the use of the HP filter. They used the HP filter in order to smooth the moving standard deviations.

The same method has been used in this study. Let $\tilde{z}_t$ be the filtered series

$$\tilde{z}_t = s_y \left[\frac{\left(z_t^{(i)}\right)^d}{HP^{(r)}\left(\sqrt{\sum_{\tau=t-\nu}^{t+\nu}\left(z_\tau^{(i)}\right)^{2d} / 2\nu}\right)} \right] + \bar{y} \qquad (3.2)$$

where $t = \max\left[k-\eta, l-\nu\right], \max\left[k-\eta+1, l-\nu+1\right], ...$ with k and l odd numbers as the window lengths in the numerator and denominator, respectively, and $\eta = (k-1)/2, \nu = (l-1)/2.$ and $i = a, b$ indexes the two detrending operations

$$\text{(a) } z_t^{(a)} = \Delta y_t - \sum_{\tau=t-\eta}^{t+\eta} \Delta y_\tau / k, \qquad t = \eta+1, \eta+2, ..., n-\eta \qquad (3.3a)$$

Note that for $\eta = (n-1)/2$, the term $\sum_{\tau=t-\eta}^{t+\eta} \Delta y_\tau / k$ equals $\overline{\Delta y}$.

and with y_τ delayed one period

$$\text{(b) } z_t^{(b)} = \Delta y_t - \sum_{\tau=t-\eta}^{t+\eta} \Delta y_{\tau-1} / k \quad t = \eta+2, \eta+3, ..., n-\eta+1 \quad (3.3b)$$

In case when k=1 then $\eta = 0$, operation (3.3b) is used. $\Delta y_{\tau-1}$ is equivalent to second order difference operation $\Delta^2 y_t$ where $\Delta y_t = y_t - y_{t-1}$, y_t is the logarithmic series at time t.

The transformations in (3.2) are generalized by raising $z_t^{(i)}$ to the power of d, which is not necessarily an integer. The best choice of η depends on the properties of the series studied. Stockhammar and Öller (2011) proposed using window length k = l = 15 (or η = ν = 7) and the standard value used for quarterly data, γ = 1 600. Stockhammar and Öller (2011) also used these values and set d=1 for the UK, US and G7 GDP series. The same filter is here used for the UK, US and CA GDP series.

Figure 3.1 shows the difference log US, UK and CA series after the heteroscedasticity filtering (3.2).

Figure 3.1: The heteroscedasticity filtered diffference in logarithmic GDP series. The left hand side panel in each row shows the frequency distribution of the filtered data. The solid line refers to the normal distribution with the same mean and variance as in the filtered series, and the dotted line is the Kernal distribution.

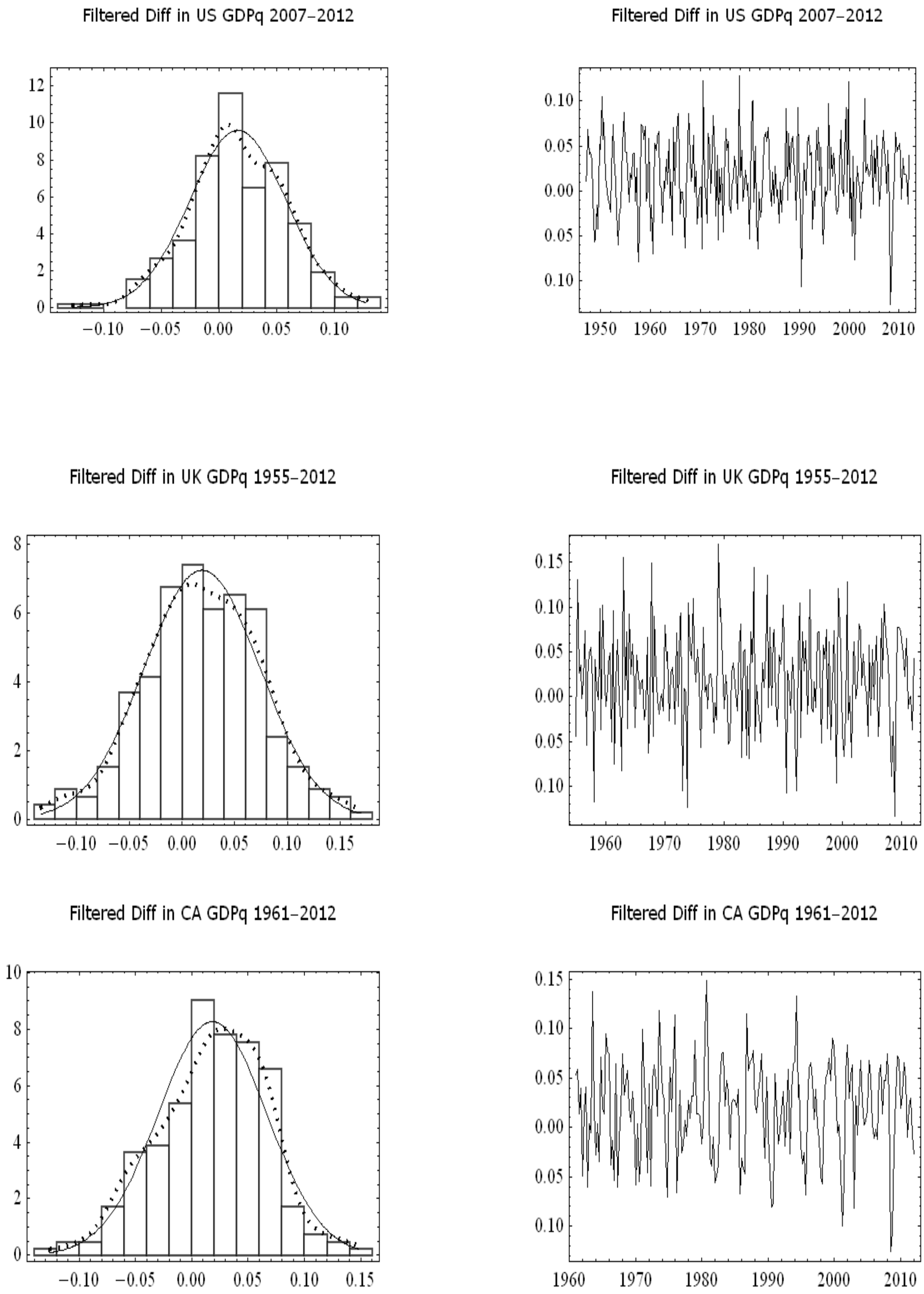


Table 3.1: The moments and homocedasticity tests of the filtered US, UK and CA GDP series.

	$\hat{\mu}_1$	$\hat{\sigma}$	$\hat{\tau}$	$\hat{K}$	ARCH-LM (p-value)
$\tilde{z}_t$,US	0.016	0.042	-0.109	0.262	0.863
$\tilde{z}_t$,UK	0.019	0.055	-0.066	0.046	0.596
$\tilde{z}_t$,CA	0.018	0.048	-0.222	0.000	0.070

skewness $(\hat{\tau})$, excess kurtosis $(\hat{K})$

Heteroscedasticity is removed from the data in the US and UK filtered series, whereas in the CA series the null hypothesis is accepted at $\alpha = 0.05$. The US series is more leptokurtic as compared to the other series. All the filtered GDP series are negatively skewed.

Table 3.2: Filter effects on the moments of the difference log US, UK and CA GDP series.

	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{K}_1$	$\hat{K}_2$
Period	1	2	1	2	1	2	1	2
y_t, US	0.019	0.014	0.013	0.008	-0.178	0.134	1.214	4.811
$\tilde{z}_t$, US	**0.016**	**0.016**	**0.041**	**0.042**	**0.081**	**-0.292**	**-0.211**	**0.758**
y_t, UK	0.025	0.014	0.017	0.009	0.474	-0.572	0.657	2.881
$\tilde{z}_t$, UK	**0.019**	**0.019**	**0.055**	**0.055**	**0.099**	**-0.224**	**0.380**	**-0.216**
y_t, CA	0.025	0.011	0.011	0.011	0.606	-1.840	0.058	6.775
$\tilde{z}_t$, CA	**0.017**	**0.019**	**0.048**	**0.048**	**0.225**	**-0.665**	**-0.262**	**0.389**

Period 1 represents 1947q1-1979q3 (US), 1955q1-1983q2 (UK) and 1961q1-1986q2 (CA). Period 2 contains 1979q4-2012q3 (US), 1983q3-2012q1 (UK) and 1986q3-2012q1.

The mean $\hat{\mu}$ and the standard deviation $\hat{\sigma}$ are stable in filtered series. Except for skewness in the US series, the estimates of skewness $(\hat{\tau})$ and excess kurtosis $(\hat{K})$ are more stable in filtered series. Stockhammar and Öller (2011) showed that this filter did not distort white noise, and thus preserved the dynamics of the time series.

The unfiltered series in Figure 2.1 do not appear to be normal. Table 3.3 shows that the filter brings them closer to normality.

Table 3.3: Filter effects on the normality of the diff log GDP series, US, UK and CA.

	AD	SW	KS	JB	χ^2	CVM	SF
y_t, US	***	***	***	***	***	***	***
$\tilde{z}_t$, US							
y_t, UK	***	***	***	***	***	***	***
$\tilde{z}_t$, UK							
y_t, CA	***	***	**	***	***	***	***
$\tilde{z}_t$, CA	*				**	*	

In Table 3.3 *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively, for the null hypothesis of normality. Seven commonly used normality tests are reported, where AD, SW, KS, JB, χ^2, CVM and SF are the Anderson-Darling, Shapiro-Wilk, Kolmogorov-Smirnov and Jarque-Bera, Pearson chi-square, Cramer-Von Mises and Shapiro-Frania test respectively. These tests are based on very different measures, and can therefore lead to different conclusions.

According to e.g. Dyer (1974), Thadewald and Buning (2007) and Razali and Wah (2011) the power of normality tests is generally low, especially in small samples. Note that the χ^2, AD and CVM statistic for the CA series reject the null hypotheses of normality at 5 and 10 percent level respectively. At least for the CA series it seems meaningful to see if there are other distributions that better fit the data. Considering the low power of the tests we will try the same for the US and the UK series. The normal distribution remains an alternative hypothesis.

4 Models for the shock distributions

A mixture distribution is a probability density function of the form

$$f(x) = \sum_{k=1}^{K} \lambda_k f(x;\theta_k)$$

Here, K is the number of components in the mixture distribution and λ_k is the mixing weights, for all $K\, \lambda_k \geq 0$, $\sum_{k=1}^{K} \lambda_k = 1$. For each K, $f(x;\theta_k)$ is the PDF of component number K.

A non-negligible risk is involved when the distribution changes over time in long time series. The data might have passed through a number of different regimes, not completely eliminated by filter (3.2). Every such regime can follow a different distribution. The filtered US and UK GDP show a small hump in the right tail while the filtered CA shows it in the left tail in Figure 3.1, which may indicate that the data are characterized by at least two regimes. Given the relatively few observations, the numbers of possible regimes we take into account are here restricted to two. Moreover, the homoscedasticity test was unable to detect non-constancy of variances which makes it hard to detect regimes with different variances.

The probability density function (PDF) of the NM distribution is:

$$f_{NM(\tilde{z}_t;\theta)} = \frac{w}{\sigma_1\sqrt{2\pi}} \exp\left\{-\frac{\left(\tilde{z}_t - \mu_1\right)^2}{2\sigma_1^2}\right\} + \frac{1-w}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{\left(\tilde{z}_t - \mu_2\right)^2}{2\sigma_2^2}\right\} \qquad (4.1)$$

where $0 \leq w \leq 1$ is the weight parameter and θ consists of the parameters (w, μ_1, μ_2, σ_1, σ_2) .

It is possible to introduce skewness and excess kurtosis in the NM distribution by introducing different means and variances for the regimes. In empirical finance, NM distributions are widely used. Wirjanto and Xu (2009) provided a selected review of recent developments and applications of the NM distribution in empirical finance. The NM distribution is able to capture the leptokurtic, skewed and multimodal characteristics, and is flexible enough to accommodate various forms of continuous distribution in time series data. Kamaruzzaman et al. (2012) found that the NM distribution captured the leptokurtic as well as skewness in the data, and they proposed a two component NM distribution for financial time series. This data included monthly rates of returns for three indices of Bursa Malaysia Index Series which had characteristics of non-normality and were asymmetric. The NM distribution is suitable to accommodate certain discontinuities in shock returns such as 'weekend effect', 'the turn-of the month effect' and 'the January effect', see Klar and Meintanis (2005).

The Laplace (L) distribution is also called the double exponential distribution. L distribution is the distribution of differences between two independent variates with identical exponential distributions. The L distribution PDF is:

$$f_{L(\tilde{z}_t;\theta)} = \frac{1}{2\phi}\exp\left\{-\frac{|\tilde{z}_t-\mu|}{\phi}\right\} \quad (4.2)$$

where $\theta=(\mu,\phi)$, $\mu\in\mathbb{R}$ is the location parameter and $\phi>0$ is the scale parameter.

The L distribution has been used in many fields like engineering, finance, electronics, etc. (see Kotz et al. 2001). The L distribution is symmetric around its mean (μ) with var(y) = $2\phi^2$ and excess kurtosis $\hat{k}$ = 3. The L distribution has fatter tails compared to the normal distribution. It is, however, hard to find a clear shape parameter which makes it rather inflexible. Also, the excess kurtosis is restricted to the constant value (3), no matter what the kurtosis in the data. Table 3.1 shows that the kurtosis in Laplace distribution is too large for the filtered growth series in this study ($\hat{k}$ = 0.262 for the US, $\hat{k}$ = 0.046 for the UK and $\hat{k}$ = 0.00002 for the CA). Clearly, the data cannot be explained by L distribution alone.

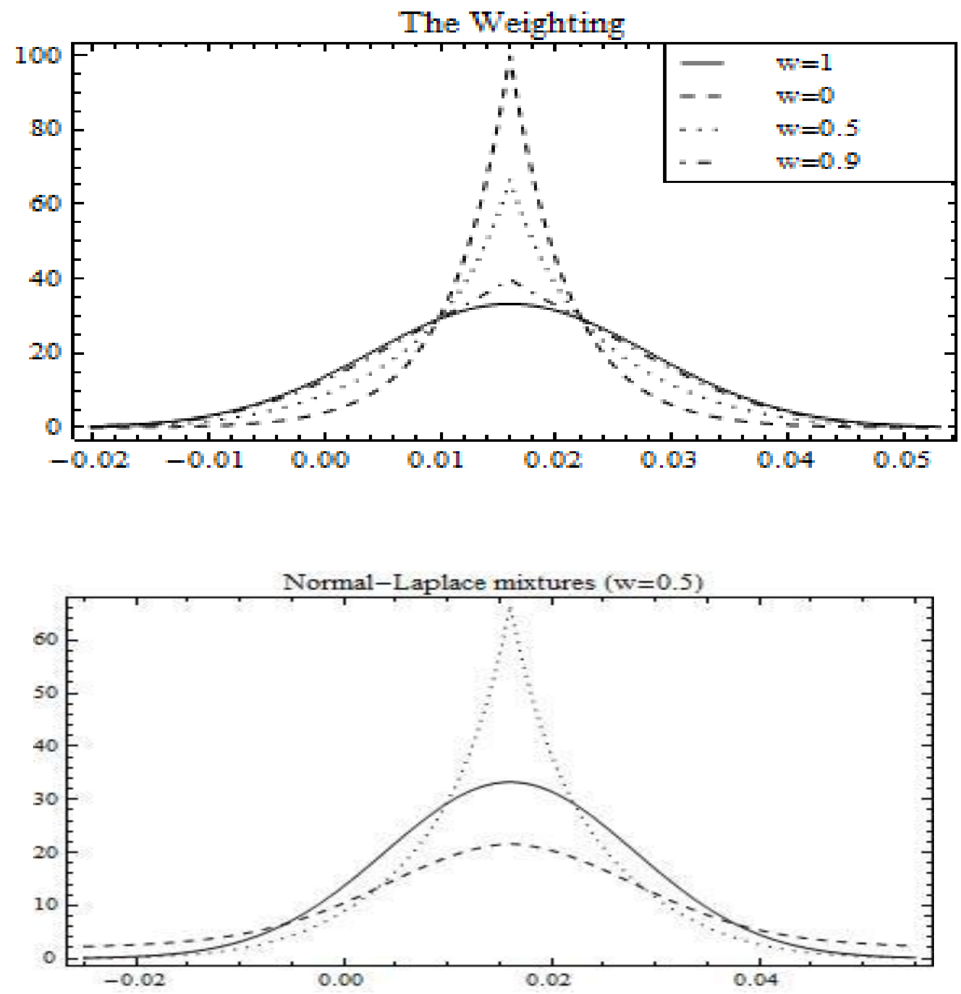
It is, however, possible to modify the L distribution by allowing it to have a second stochastic component. This means that its empirical counterpart is buried in Gaussian noise. We therefore combine (4.2) with a normal distribution with a weight parameter *w*. This mixture was introduced by Kanji (1985) to model wind shear data.

The Normal -Laplace (NL) mixture distribution specified by:

$$f_{NL(\tilde{z}_t;\theta)} = \frac{w}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\tilde{z}_t-\mu)^2}{2\sigma^2}\right\} + \frac{(1-w)}{2\sigma}\exp\left\{-\frac{|\tilde{z}_t-\mu|}{\phi}\right\} \quad (4.3)$$

for $-\infty < \tilde{z}_t < \infty$ and for the parameters: $-\infty < \mu < \infty, 0 \le w \le 1$ and $\sigma > 0$. In (4.3) the N and L distributions carry the same mean. Jones and McLachlan (1990) generalized (4.3) and demonstrated that this may lead to an even better fit than that by Kanji. Hass, Mittnik, and Paolella (2006) used NL mixture in modeling and predicting financial risk based on 25 daily stock return series. The characteristics of the NL density are shown in Figure 4.1

Figure 4.1: NL densities. The upper panel shows different weightings of the two components in the NL distribution (with μ =0.016; σ =0.012; ϕ =0.005). The solid line in the lower panel represents the pure N (0.016, 0.012) distribution together with two NL mixture distributions with w=0.5; ϕ =0.05 (dashed line) and ϕ =0.005 (dotted line), respectively.



The L and NL mixture distributions in figure 4.1 do not account for skewness in the data. McGill (1962) has proposed a suitable skewed generalization of the L distribution. He considers an asymmetric Laplace (AL_1) distribution with a PDF of the form

$$f_{AL_1(\tilde{z}_t;\theta)} = \begin{cases} \dfrac{1}{2\psi}\exp\left\{\dfrac{\tilde{z}_t-\mu}{\psi}\right\} & if\ \tilde{z}_t \le \mu \\ \dfrac{1}{2\phi}\exp\left\{\dfrac{\mu-\tilde{z}_t}{\phi}\right\} & if\ \tilde{z}_t > \mu \end{cases} \quad (4.4)$$

where again μ is the location parameter. The maximum likelihood (ML) estimate of μ is the median. The distribution has three parameters $\theta = (\mu, \phi, \psi)$. For $\psi > \phi$, the distribution is negatively skewed and vice versa for $\phi < \psi$. The L distribution is a special case of AL when $\phi = \psi$. In $AL_{1,}$ ψ is the parameter of shocks weaker than the trend and ϕ that of stronger shocks than the trend.

The AL distribution can be used for modeling currency exchange rate, interest rate, stock price changes, etc. In the last few decades, various forms and applications of AL distributions can be traced in the literature (see Kozubowski and Nadarajah 2010).

Kozubowski and Podgorski (1999, 2000) used the AL distribution for modeling interest rate and currency exchange rate. Kotz et al. (2001) studied L and AL distribution application in communication, engineering, economics and finance. Linden (2001) demonstrated highly significant ψ and ϕ using AL distribution to model the return of 20 stocks. A three-parameter AL distribution was fitted to flood data by Yu and Zhang (2005). Jayakumar and Kuttykrishnan (2007) developed autoregressive model with AL distribution to apply it on time series data. Julia and Rego (2008) used AL distribution in the field of microbiology to fit flow cytometric scatter data. Kozubowski and Nadarajah (2010) reviewed sixteen known variations in the Laplace distribution. They provided the basic mathematical properties, including its moment and ML estimator and for each particular case, and discussed the area of application with references. Harandi and Alamatsaz (2013) introduced a new class of Alpha–Skew–Laplace distribution with flexible hazard rate behavior and demonstrated that such distributions were more flexible, and fitted better to some real data sets.

An advantage of the AL distribution is that, unlike the L distribution, the kurtosis is not fixed. The AL distribution becomes even more leptokurtic compared to the L distribution with an excess kurtosis that varies between 3 and 6 (the smallest value for the L distribution and the largest value for the exponential distribution). Secondly, AL_1 distribution is skewed (for $\phi \neq \psi$) which is another advantage. An enhanced flexibility of AL distributions can be achieved by changing the asymmetry and kurtosis.

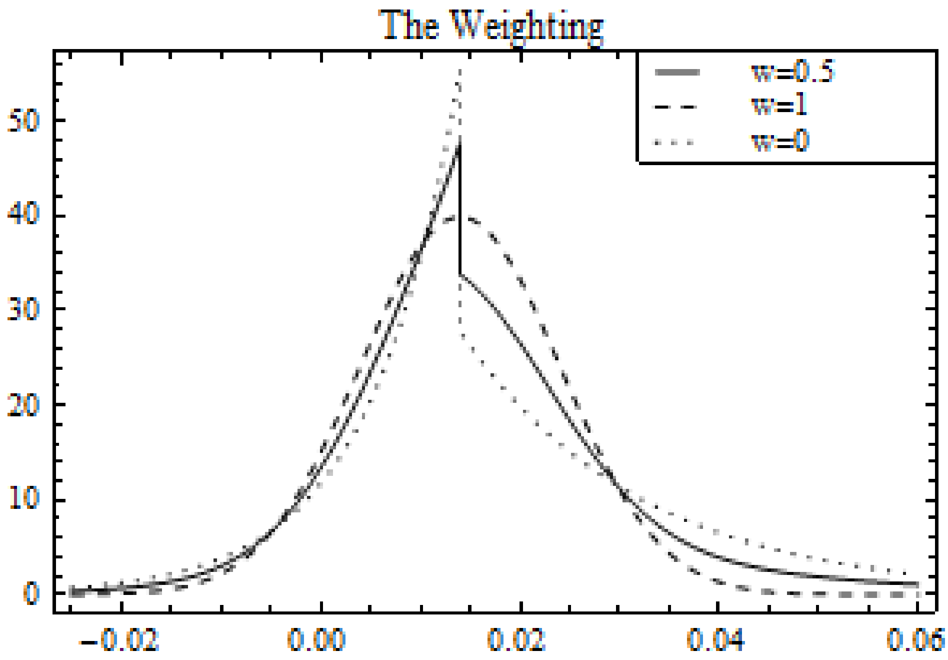
Because of the large leptokurtosis of the AL_1 distribution, Stockhammar and Öller (2011) added Gaussian noise and used this mixture of distribution first time on macroeconomic time series data. Basic assumption was that each shock was an independent drawing from either a N or AL distribution. The probability density distribution of the filtered growth series ($\tilde{z}_t$) was described by a weighted sum of N and AL_1 random shocks, i.e.

$$f_{NAL_1(\tilde{z}_t;\theta)} = \frac{w}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\tilde{z}_t-\mu)^2}{2\sigma^2}\right\} + (1-w)\begin{cases} \frac{1}{2\psi}\exp\left\{\frac{\tilde{z}_t-\mu}{\psi}\right\} & if\ \tilde{z}_t \leq \mu \\ \frac{1}{2\phi}\exp\left\{\frac{\mu-\tilde{z}_t}{\phi}\right\} & f\ \tilde{z}_t > \mu \end{cases} \tag{4.5}$$

where θ consisted of the five parameters (w, μ, σ, ψ, ϕ).

Equation (4.5) is referred to as the mixed Normal-Asymmetric Laplace-1 (NAL_1) distribution. Like Jones and McLachlan (1990), Stockhammar and Öller (2011) assumed equal medians but unequal variances for the components in the mixture distribution. It had a jump discontinuity at μ when $\phi \neq \psi$ see Figure 4.2. Looking at the smoothed empirical distributions in Figure 3.1, the discontinuity seemed counterintuitive. However, the histograms in Figure 3.1 lent some support to a jump close to μ. Figure 4.2 shows NAL_1 densities for three different values of the weight parameter *w*.

Figure 4.2: NAL_1 densities. The Figure shows a pure N (0.014, 0.04), an AL_1 (ψ =0.02, ϕ =0.01 and μ =0.014) distribution (w=1 and w=0, respectively) and a compound of these two components with w=0.5. Note the discontinuity at μ .



The PDF of Student's t distribution with location parameter μ, scale parameter σ and shape parameter (degrees of freedom) ν is defined as

$$f_{t(\tilde{Z}_t;\theta)} = \frac{\nu^{-\frac{1}{2}+\frac{1+\nu}{2}}\left[\dfrac{1}{\nu+\dfrac{(\tilde{Z}_t-\mu)^2}{\sigma^2}}\right]^{\frac{1+\nu}{2}}}{\sigma Beta\left[\dfrac{\nu}{2},\dfrac{1}{2}\right]}$$

Where θ consists of three parameters (μ, σ, ν,). If x is independent standard student's t distributed variable with degree of freedom ν while σ and μ are overall scaling and the location of the distribution then the variable

$$z = x * \sigma + \mu$$

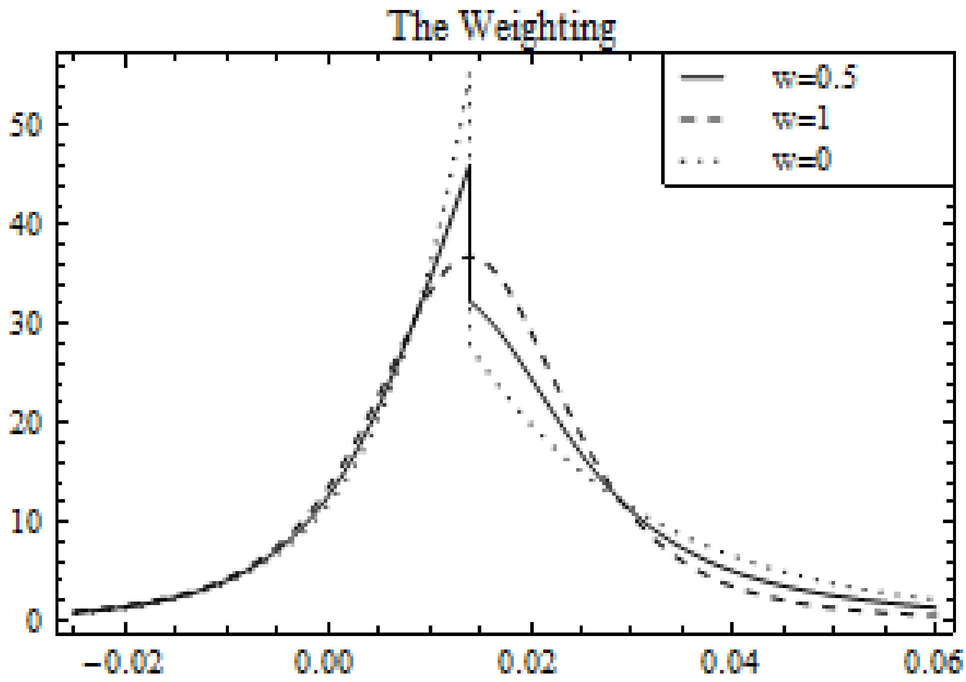
is said to have a student t distribution with three parameter μ, σ and ν. It converges to normal distribution with mean μ and standard deviation σ as degree of freedom ν becomes large.

We have introduced a new mixture by adding Student's t distribution with AL_1 to decrease the excess kurtosis in AL_1. To the author's best knowledge this distribution has not been used before for microeconomic time series data. Student's *t* distributions are symmetric, uni-modal, bell-shaped and leptokurtic distributions. In the case where $\mu = 0$ and $\sigma = 1$ we have Standard student's t distribution. The shape parameter determines the fatness of the tails; excess kurtosis will decrease as the degree of freedom ν increases. We assume that each shock is an independent drawing from either a student's t or an AL_1 distribution. The probability density distribution of the filtered growth series ($\tilde{z}_t$) can then be described by a weighted sum of student's t and AL_1 random shocks, i.e.

$$f_{TAL_1(\tilde{z}_t;\theta)} = \frac{w\ \nu^{-\frac{1}{2}+\frac{1+\nu}{2}} \left[\dfrac{1}{\nu + \dfrac{(\tilde{z}_t - \mu)^2}{\sigma^2}} \right]^{\frac{1+\nu}{2}}}{\sigma Beta\left[\dfrac{\nu}{2}, \dfrac{1}{2}\right]} + (1-w) \begin{cases} \dfrac{1}{2\psi} \exp\left\{\dfrac{\tilde{z}_t - \mu}{\psi}\right\} & if\ \tilde{z}_t \le \mu \\ \dfrac{1}{2\phi} \exp\left\{\dfrac{\mu - \tilde{z}_t}{\phi}\right\} & f\ \tilde{z}_t > \mu \end{cases} \qquad (4.6)$$

where θ consists of the five parameters (w, μ, ν, σ, ψ, ϕ). Equation (4.6) is referred to as the mixed Student's t Asymmetric Laplace-1 (TAL_1) distribution. Equal medians, but unequal variances, are assumed for the components in the proposed distribution. It has a jump discontinuity at μ when $\phi \neq \psi$ see Figure 4.2. Figure 4.3 shows TAL_1 densities for three different values of the weight parameter *w*.

Figure 4.3: TAL$_1$ densities. The Figure shows a pure T (0.014, 0.04, 3), an AL$_1$ (ψ =0.02, ϕ =0.01 and μ =0.014) distribution (w=1 and w=0 respectively) and a compound of these two components with w=0.5. Note the discontinuity at μ .



Stockhammar and Öller (2011) used convoluted version suggested by Reed and Jorgensen (2004) for Mixture of N and AL$_2$ distributions. For convolution, instead of using the AL$_1$ parameterization in (4.4) they used:

$$f_{AL_2(\tilde{z}_t;\theta)} = \begin{cases} \dfrac{\alpha\beta}{\alpha+\beta}\exp\{\beta\tilde{z}_t\} & if\ \tilde{z}_t \le 0 \\ \dfrac{\alpha\beta}{\alpha+\beta}\exp\{-\alpha\tilde{z}_t\} & if\ \tilde{z}_t > 0 \end{cases} \qquad (4.7)$$

We have used AL$_2$ to make the weighted mixture of AL$_2$ with normal and Student's t distribution. We assume that each shock is an independent drawing from either N or AL$_2$ distribution.

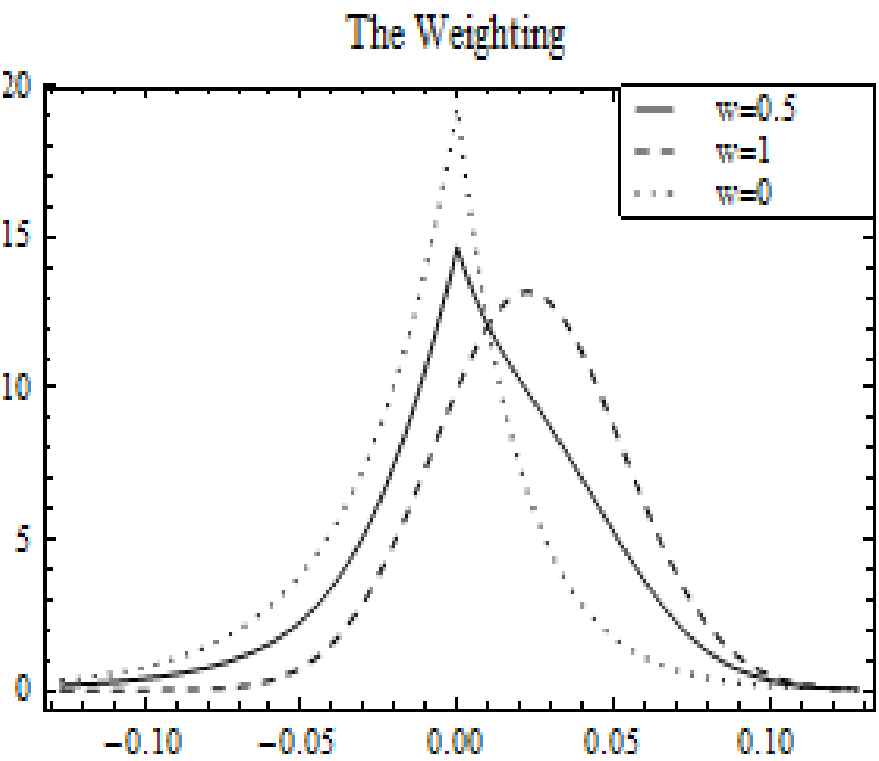
The probability density distribution of the filtered growth series ($\tilde{z}_t$) can then be described by a weighted sum of N and AL$_2$ as

$$f_{NAL_2}(\tilde{z}_t;\theta) = \frac{w}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\tilde{z}_t-\mu)^2}{2\sigma^2}\right\} + (1-w)\begin{cases} \frac{\alpha\beta}{\alpha+\beta}\exp\{\beta\tilde{z}_t\} & if\ \tilde{z}_t \le 0 \\ \frac{\alpha\beta}{\alpha+\beta}\exp\{-\alpha\tilde{z}_t\} & if\ \tilde{z}_t > 0 \end{cases} \quad (4.8)$$

Where θ consists of the five parameters (w, μ, σ, α, β).

Equation (4.8) is referred to as the mixed Normal Asymmetric Laplace-2 (NAL_2) distribution.

Figure 4.4: NAL_2 densities. The Figure exhibits a pure N (0.022, 0.04), an AL_2 (α =48.17, β =32.35) distribution (w=1 and w=0 respectively) and a compound of these two components with w=0.5. Note there is no discontinuity at μ.



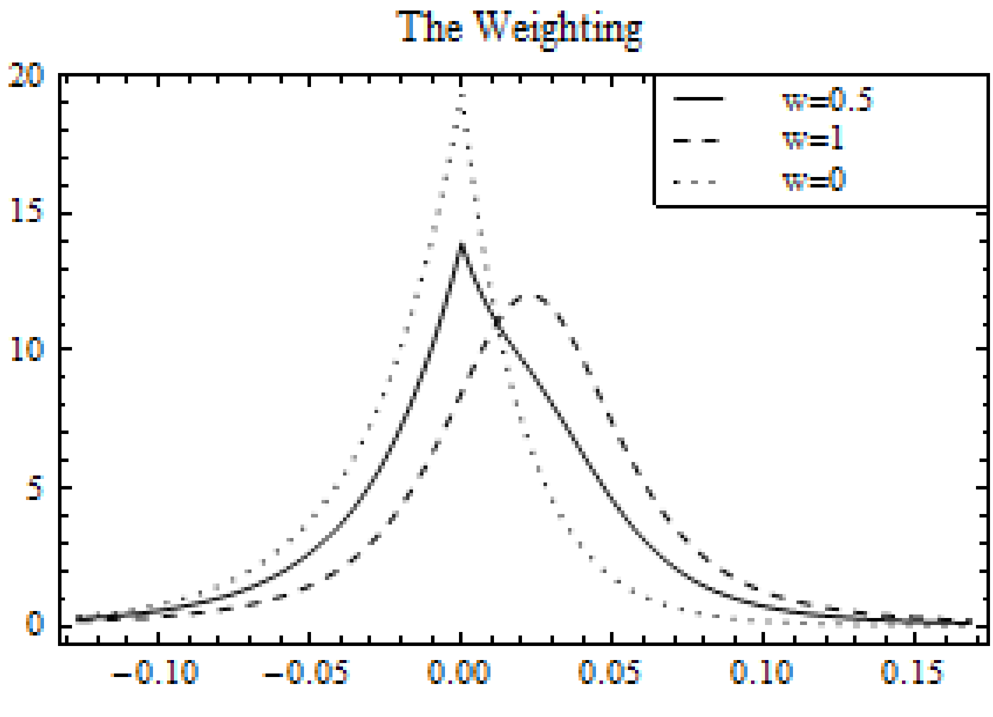
Similarly, we assume that each shock is an independent drawing from either student's t or AL_2 distribution. The probability density distribution of the filtered growth series ($\tilde{z}_t$) can then be described by a weighted sum of student's t and AL_2 as

$$f_{TAL_2(\tilde{z}_t;\theta)} = \frac{w\ \nu^{-\frac{1}{2}+\frac{1+\nu}{2}}\left[\dfrac{1}{\nu+\dfrac{(x-\mu)^2}{\sigma^2}}\right]^{\frac{1+\nu}{2}}}{\sigma Beta\left[\dfrac{\nu}{2},\dfrac{1}{2}\right]} + (1-w)\begin{cases} \dfrac{\alpha\beta}{\alpha+\beta}\exp\{\beta\tilde{z}_t\} & if\ \tilde{z}_t \le 0 \\ \dfrac{\alpha\beta}{\alpha+\beta}\exp\{-\alpha\tilde{z}_t\} & if\ \tilde{z}_t > 0 \end{cases} \quad (4.9)$$

where θ consists of the six parameters (w, μ, ν, σ, α, β).

Equation (4.9) is referred to as the mixed Student's t Asymmetric Laplace-2 (TAL_2) distribution.

Figure 4.5: TAL_2 densities. The Figure shows a pure T(0.022,0.04,2.63), an AL_2(α =48.17, β=32.35) distribution (w=1 and w=0 respectively) and a compound of these two components with w=0.5. Note there is no discontinuity at μ.



5 Estimation and Assessment of Distributional Accuracy

In this chapter, we will use all six distributions in order to find out which one best fits the data. The parameters of all the distributions are estimated by using the method of maximum likelihood (ML). The ML estimates for parameters of the distributions are obtained by numerical maximization of log likelihood of the distribution under a parametric constrain. There are several methods available for numerical maximization of log-likelihood e.g. Nelder and Mead, Simulated annealing, Differential evolution, Random search algorithm, Newton-Raphson method, Method of Scoring, EM (Expectation and Maximization) Algorithm etc. Each method has advantages as well as drawbacks. The EM algorithm is the standard method for maximum likelihood estimation in finite mixture models but it has some drawbacks. For example, the solution depends on choice of initial values and stopping criteria. It is sometimes very slow to converge. The hessian matrix must be calculated manually. Significant implementation effort is required compared to numerical optimization.

For numerical maximization of the log-likelihood, we have used the Nelder and Mead method, proposed by John Nelder and Roger Mead in 1965, which does not require derivative information. This method is simple, intuitive and relatively stable in approaching the optimum, and can be applied to discontinuous problems. It works well when the numbers of estimated parameters are small (up to 10-20). There is, however, no guarantee for the convergence of Nelder and Mead algorithm, even for smooth problems. In practice the performance of the Nelder and Mead algorithm is generally good, see Wright (1995) and Lagariasat at al. (1998).

The Nelder and Mead algorithm is one of the most well known and widely used algorithms for optimization in the fields of statistics, chemical engineering, physical and medical sciences, engineering see Price et al. (2002) and Lewis et al. (2000).“*In late May 2012, Google Scholar displayed more than 2,000 papers published in 2012 that referred to the Nelder–Mead method…*” (Wright, 2012, p.274). Lagarias et al. (1998) stated that “*Two measures of the ubiquity of the Nelder-Mead method are that it appearsin the best-selling handbook Numerical Recipes, where it is called the “amoebaalgorithm," and in Matlab*”(p.112). Olsson (1979) directly searched the maximum of the log-likelihood function of the Mixture Weibull distribution through the Nelder and Mead Procedure. Everitt (1988) introduced a finite mixture density to model the clustering of mixed mode data. He used the Nelder and Mead method to find the ML estimates and showed that its performance is relatively satisfactory by using several small scale numerical examples. Wu (2008) estimated the parameters of a five-parameter generalized Normal Laplace (GNL) and four-parameter Normal Laplace (NL) distributions to grouped income data by maximum likelihood using the Nelder and Mead method. Manoj et al. (2013) proposed a new Binomial mixture distribution called the McDonald Generalized Beta-Binomial distribution (McGBB) and demonstrated that the McGBB mixture distribution fit the data better than the Beta-Binomial and the Kumaraswamy-Binomial distribution. They used the Nelder and Mead method to estimate the ML estimator of the parameters of the McGBB distribution.

It is based on evaluating a function at the vertices of a simplex, then iteratively shrinking the simplex as better points are found until some desired bound is obtained. The Nelder and Mead method is a direct search method. It has four parameters: The reflection parameter α ($\alpha > 0$), the expansion parameter β ($\beta > 1$), contraction parameter ρ ($\rho > 1$) and shrinkage or reduction parameter γ ($\gamma > 1$). The standard values of these parameters are

$\alpha = 1$, $\beta = 2$, $\rho = \frac{1}{2}$ and $\gamma = \frac{1}{2}$.

For a function of *n* variables, the algorithm maintains a set of *n*+ 1 points $\{x_1, x_2, ... x_{n+1}\}$ forming the vertices in *n* dimensional space.

1) At each iteration, order occurs according to the values at the vertices $f(x_1) < f(x_2) < ... < f(x_{n+1})$
2) The centroid of the best n points is denoted by $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
3) Compute the reflection point $x_r = \bar{x} + \alpha(\bar{x} - x_{n+1})$ where $\alpha > 0$ is a reflection parameter. If $f(x_1) < f(x_r) < f(x_n)$, the new point x_r is neither a new worst point nor a new best point, then x_r is replaced by x_{n+1}.

4) If $f(x_r) < f(x_1)$, the new point x_r is better than the current best point, then we will go further in this direction $x_e = \bar{x} + \beta(\bar{x} - x_{n+1})$, where $\beta > 1$ is expansion parameter. If $f(x_e) < f(x_r)$ the expansion is successful, x_e is replaced by x_{n+1} otherwise x_r is replaced by x_{n+1}.

5) If the new point x_r is worse than the 2nd worst point, $f(x_r) \geq f(x_n)$ then the contracted point is defined as

$$x_c = \begin{cases} \bar{x} + \rho(x_{n+1} - \bar{x}), if \ f(x_r) \geq f(x_{n+1}) \\ \bar{x} + \rho(x_t - \bar{x}), \ if \ \ f(x_r) < f(x_{n+1}) \end{cases}$$

where $0 < \rho < 1$ is contracted parameter

If $f(x_c) < Min(f(x_{n+1}), f(x_r))$, the contraction is successful and X_c replaces X_{n+1} otherwise a further contraction is carried out.

6) Replace all but the best point, with point

$$x_i = x_1 + \gamma(x_i - x_1) \ \ i = 2,3,...,n+1$$

If the difference between new best point and old best point or the difference between best functional values for new and old best point are less than the required tolerance level the process is assumed to be converged.

A simulation study is performed to find out the standard errors of the estimated parameters which indicate the performance of the estimation procedure. The process of simulation of a variate from two component mixture distribution was undertaken in two steps

1) First a multivariate **M:** 1, W_1 and W_2 mixture indicator variate is drawn from the multinomial distribution with probabilities equal to the mixture weights.

2) Then given the drawn mixture indicator value, say *k*, the variate ***X*** is drawn from the k^{th} component distribution. The mixture indicator value *k* is used to generate the ***X*** = *x* otherwise discarded.

The process to obtain the standard errors of the estimated parameters using a simulation study is given below:

1) We have drawn 1000 samples of size equal to the length of data from each distribution.

2) For each simulated sample, the ML estimates for the parameter are obtained by using Nelder and Mead optimization method.

3) The standard errors are obtained by taking the standard deviation of these 1000 ML estimates of the parameters.

For independent identically distributed observations, the likelihood is the product of the probability density function evaluated at each of observed value of the data. Consider the sample of *n* independent observation $x_1, x_{2,} ..., x_n$, then the likelihood and log-likelihood functions of *k* components finite mixture model can be written as

$$L(\theta) = \prod_{i=1}^{n} \sum_{j=1}^{k} W_j f(x_i \mid \lambda_j)$$

and

$$l(\theta) = \sum_{i=1}^{n} \log \sum_{j=1}^{k} W_j f(x_i \mid \lambda_j)$$

Here the unknown parameters are the mixing weights W_j and the components parameters λ_j and $\sum_{j=1}^{k} W_j = 1$. Maximization of $l(\theta)$ with respect to θ, for given data x, yields the ML estimate of θ.

The log-likelihood $l(\theta)$ of the NM distribution is:

$$l(\theta)=\sum_{t=1}^{n}\log[W((2\pi\sigma_1^2)^{\frac{1}{2}}\exp(\frac{\tilde{z}_t-\mu_1}{2\sigma_1^2})^2))+(1-W)((2\pi\sigma_2^2)^{\frac{1}{2}}\exp(\frac{\tilde{z}_t-\mu_2}{2\sigma_2^2})^2))]$$

We numerically maximize the above log likelihood function and perform the simulation study to obtain ML estimates and standard errors of the parameters. The ML estimates and standard errors of the parameters for the NM distribution are given below:

Table 5.1: Estimated parameters and standard errors of estimate for the NM distribution

	W	μ_1	σ_1	μ_2	σ_2
US	0.8812 (0.3539)	0.0184 (0.0240)	0.0394 (0.0125)	0.0019 (0.0443)	0.0550 (0.0138)
UK	0.4999 (0.3423)	-0.0137 (0.0108)	0.0438 (0.0158)	0.0518 (0.0295)	0.0420 (0.0552)
CA	0.8143 (0.3292)	0.0328 (0.0220)	0.0385 (0.0073)	-0.0454 (0.0443)	0.0311 (0.0114)

Standard errors for estimates in parentheses

The log-likelihood $l(\theta)$ of the NAL_1 distribution is:

$$l(\theta)=\sum_{t=1}^{n}\log\left[W((2\pi\sigma_1^2)^{\frac{1}{2}}\exp(\frac{\tilde{z}_t-\mu_1}{2\sigma_1^2})^2))+(1-W)\begin{cases}(2\psi)^{-1}\exp(\frac{1}{2\psi}(\tilde{z}_t-\mu))I(\tilde{z}_t\le E(\tilde{z}_t))\\(2\phi)^{-1}\exp(\frac{1}{2\phi}(\mu-\tilde{z}_t))I(\tilde{z}_t>E(\tilde{z}_t))\end{cases}\right]$$

where I is the indicator function.

The ML estimate for μ is the median for the AL_1 distribution. The ML estimates and standard errors of the parameters for the NAL_1 distribution are obtained by numerical maximization of above log likelihood function and simulation study.

Table 5.2: Estimated parameters and standard errors of estimates for the NAL_1 distribution

	W	μ	σ	ϕ	ψ
US	0.87617 (0.28239)	0.01404 (0.00280)	0.04309 (0.01138)	0.02943 (0.021148)	0.01120 (0.01397)
UK	0.98338 (0.43049)	0.01776 (0.00480)	0.05516 (0.02515)	0.04082 (0.02406)	0.00311 (0.02127)
CA	0.94186 (0.30371)	0.02008 (0.00381)	0.04927 (0.01495)	0.02666 (0.01677)	0.03851 (0.01714)

Standard errors for estimates in parentheses

Table 5.2 shows that the Gaussian noise component dominates. In the UK series $\hat{\psi}$ is much smaller than $\hat{\phi}$ which indicates that the growth of shocks that are weaker than trend have a smaller spread than the above trend shocks. Together with a mean growth larger than zero this ensures long-term economic growth.

The log-likelihood $l(\theta)$ of the NAL_2 distribution is:

$$l(\theta) = \sum_{t=1}^{n} \log \left[W((2\pi\sigma_1^2)^{\frac{1}{2}} \exp(\frac{\tilde{z}_t - \mu_1}{2\sigma_1^2})^2)) + (1-W) \begin{cases} \frac{\alpha\beta \exp(-\tilde{z}_t \alpha)}{\alpha+\beta} I(\tilde{z}_t > 0) \\ \frac{\alpha\beta \exp(\tilde{z}_t \alpha)}{\alpha+\beta} I(\tilde{z}_t \leq 0) \end{cases} \right]$$

where I is the indicator function

The ML estimates and standard errors of the parameters for the NAL_2 distribution are given in table 5.3 below. This is done by numerical maximization of above log likelihood function and a simulation study.

Table 5.3: Estimated parameters and standard errors of estimates for the NAL_2 distribution

	W	μ	σ	α	β
US	0.8092	0.0225	0.0402	48.1724	32.3587
	(0.0682)	(0.0026)	(0.0024)	(15.1141)	(39.7381)
UK	0.9697	0.0195	0.0553	58.8707	42.7549
	(0.1670)	(0.0029)	(0.0043)	(8.2575)	(10.2575)
CA	0.3235	0.0514	0.0204	25.40536	27.0923
	(0.1590)	(0.0132)	(0.0119)	(3.7694)	(20.4305)

Standard errors for estimates in parentheses

The above table shows that the Gaussian noise component dominates in US and UK series and for CA series AL_2 noise component dominates.

The log likelihood $l(\theta)$ of the TAL_1 distribution is:

$$l(\theta)=\sum_{t=1}^{n}\log\left[W\frac{\nu^{\frac{-1}{2}+\frac{1+\nu}{2}}\left(\frac{1}{\nu+\frac{(\tilde{z}_t-\mu)^2}{\sigma^2}}\right)^{\frac{1+\nu}{2}}}{\sigma Beta\left[\frac{\nu}{2},\frac{1}{2}\right]}+(1-W)\begin{cases}(2\psi)^{-1}\exp(\frac{1}{2\psi}(\tilde{z}_t-\mu))I(\tilde{z}_t\leq E(\tilde{z}_t))\\(2\phi)^{-1}\exp(\frac{1}{2\phi}(\mu-\tilde{z}_t))I(\tilde{z}_t>E(\tilde{z}_t))\end{cases}\right]$$

where I is the indicator function

The ML estimate for μ is the median for the AL_1 distribution. The numerical maximization of above log likelihood function and simulation study is performed to obtain ML estimates of the parameters and standard errors of the estimated parameters of TAL_1 distribution.

Table 5.4: Estimated parameters and their standard errors for the TAL_1 distribution

	W	μ	σ	ν	ϕ	ψ
US	0.8776	0.0140	0.0429	561.6095	0.0295	0.0111
	(0.4160)	(0.0019)	(0.0183)	(37.9616)	(0.0104)	(0.4160)
UK	0.9627	0.0177	0.0551	226.8639	0.0405	0.0127
	(0.4397)	(0.0045)	(0.0252)	(109.3940)	(0.0174)	(0.0184)
CA	0.9375	0.0202	0.0491	237.7750	0.0279	0.0329
	(0.3781)	(0.0050)	(0.0195)	(90.6533)	(0.0073)	(0.0137)

Standard errors for estimates in parentheses

Above table shows that the Student's t distribution noise component dominates for US, UK and CA GDP series.

The log-likelihood $l(\theta)$ of the TAL_2 distribution is:

$$l(\theta)=\sum_{t=1}^{n}\log\left[W\frac{\nu^{\frac{-1}{2}+\frac{1+\nu}{2}}\left(\dfrac{1}{\nu+\dfrac{(\tilde{z}_t-\mu)^2}{\sigma^2}}\right)^{\frac{1+\nu}{2}}}{\sigma Beta\left[\dfrac{\nu}{2},\dfrac{1}{2}\right]}+(1-W)\begin{cases}\dfrac{\alpha\beta\exp(-\tilde{z}_t\alpha)}{\alpha+\beta}I(\tilde{z}_t>0)\\ \dfrac{\alpha\beta\exp(\tilde{z}_t\alpha)}{\alpha+\beta}I(\tilde{z}_t\leq 0)\end{cases}\right]$$

where I is the indicator function

We have obtained ML estimates and standard errors for the TAL_2 distribution by numerical maximization of above log likelihood function and simulation study.

Table 5.5: Estimated parameters and their standard errors for the TAL_2 distribution

	W	μ	σ	ν	α	β
US	0.8608	0.0188	0.0419	97.6340	39.8234	45.4200
	(0.1520)	(0.0059)	(0.0016)	(7.6632)	(6.9394)	(1.0976)
UK	0.9777	0.0198	0.0550	246.8880	447.2991	67.6721
	(0.0470)	(0.0048)	(0.0032)	(5.5295)	(13.8898)	(22.0135)
CA	0.9745	0.0183	0.0482	8841.4700	30.7224	530.1198
	(0.3927)	(0.0129)	(0.0227)	(101.1644)	(36.1037)	(15.7053)

Standard errors for estimates in parentheses

The above table shows that student's t distribution components are contributing more than the AL_2 part.

5.1 Distributional comparison

In this section, the six fitted distributions are compared by using accuracy measure methods, like the Root Mean Square Error (RMSE), Median Absolute Percentage Error (MdAPE), Symmetric Median Absolute Percentage Error (sMdAPE), Mean Absolute Scaled Error (MASE), Goodness of fit tests including Anderson Darling (AD), Cramer-Von Mises (CVM), Kolmogorov Smirnov (KS), Pearson Chi-square (χ^2), Watson U-square (U^2), Kuiper (V) and graphical method which include Q-Q plots.

The RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{1000}\left[f_K(\tilde{z}_i) - \hat{f}(\tilde{z}_i)\right]^2}{1000}}$$

where $\hat{f}(\tilde{z}_i)$ is the hypothetical distribution, $f_K(\tilde{z}_i)$ is the kernel function of the empirical distribution, and RMSE provides a term by term comparison of the deviation between the hypothetical distribution and the kernel function of the empirical distribution. One thousand equidistant points on the horizontal axis are taken within the range of the data. Hence we have more points where distributions are almost parallel to the x-axis thus providing more weight to these points. The sum in the expression of RMSE is taken over the ordinates of these points. For US data the peak to the left of the median significantly affects the RMSE. A lower value of RMSE indicates a better fit. This scale dependent measure is more sensitive to outliers.

Because of the advantage of being scale independent, percentage error measures are widely used to compare forecasting performance. These measures have some disadvantages. They are undefined at $f_K(\tilde{z}_i)$ = 0, and for values of $f_K(\tilde{z}_i)$ close to zero have an extremely skewed distribution. The MdAPE is defined as:

$$MdAPE = median\left(\frac{100\times|\ f_K(\tilde{z}_i) - \hat{f}(\tilde{z}_i)|}{f_K(\tilde{z}_i)}\right)$$

This measure is better to its close relative Mean Absolute Percentage Error (MAPE) because of the asymmetry, but both MAPE and MdAPE have disadvantage that they give heavier penalty on positive errors than on negative errors. This is the reason Makridakis (1993) advocated so-called "symmetric" measures. One is these MdAPE which can be computed as:

$$sMdAPE = median\left(\frac{200\times|\ f_K(\tilde{z}_i) - \hat{f}(\tilde{z}_i)|}{f_K(\tilde{z}_i) + \hat{f}(\tilde{z}_i)}\right)$$

Another commonly used measure is the MASE defined as:

$$MASE = \frac{1}{1000}\left|\frac{f_K(\tilde{z}_i) - \hat{f}(\tilde{z}_i)}{\frac{1}{999}\sum_{i=2}^{1000}|f_K(\tilde{z}_i) - \hat{f}(\tilde{z}_{i-1})|}\right|$$

Hyndman and Koehler (2006) showed that this measure is less sensitive to outliers and perform better for small samples than other measures. It is widely applicable and easily interpretable. They suggested that MASE was the best available measure of forecast accuracy. All the above five measures are reported in table 5.6.

Table 5.6: Distributional accuracy comparison

	Distributions	RMSE	MdAPE	sMdAPE	MASE
US	N	0.2681	8.0164	8.0650	7.1255
US	NM	0.2908	7.9745	7.8076	7.2800
US	NAL_1	0.4234	10.865	10.768	9.6633
US	NAL_2	0.2433	7.8086	7.9147	6.5789
US	TAL_1	0.4233	10.953	10.8539	9.6838
US	TAL_2	**0.2366**	**7.0385**	**6.9590**	**6.3875**
UK	N	0.1743	6.7088	6.6253	7.9562
UK	NM	0.1991	**5.6624**	**5.6635**	**9.0537**
UK	NAL_1	0.2571	7.3063	7.2687	9.5638
UK	NAL_2	**0.1938**	6.9877	6.8215	9.2546
UK	TAL_1	0.2568	7.8393	7.8045	10.5785
UK	TAL_2	0.2060	6.3067	6.1242	9.2752
CA	N	0.4697	18.470	19.0277	17.4073
CA	NM	**0.1809**	**6.3643**	**6.3206**	**7.1465**
CA	NAL_1	0.4327	16.7095	16.7711	16.7458
CA	NAL_2	0.5433	21.0802	20.8634	19.5698
CA	TAL_1	0.4397	17.0775	16.7821	17.1135
CA	TAL_2	0.4941	19.5556	18.9621	18.3345

For the US series, the TAL_2 distribution using the parameter values in table 5.5 is superior to N, NM, NAL_1, NAL_2 and TAL_1 according to each measure. TAL_2 on average

12.0%, 13.4%, 37.2%, 6.9% and 37.4% better fits comparing to the benchmark N distribution and NM, NAL_1, NAL_2 and TAL_1 respectively. Whereas, by using the estimated parameter in table 5.1, the NM distribution is superior to other distributions for UK GDP series according to all measures except RMSE. The NM shows on average 12.9%, 27.7%, 20.1%, 32.2% and 17.0% better fits comparing with the benchmark N distribution, NAL_1, NAL_2, TAL_1 and $TAL_{2,}$ respectively, for the UK GDP series. For the CA GDP series also NM is superior to all other distributions according to all measures by using the parameter in table 5.1. Finally, for the CA GDP series the NM shows on average 50.6%, 47.9%, 53.9%, 48.4% and 51.7% improvement as compared to the benchmark N distribution, NAL_1, NAL_2, $TAL_1$and $TAL_{2,}$ respectively. According to this numerical comparison, the US GDP series could be looked upon as samples from a TAL_2 whereas UK and CA GDP series from NM distribution with parameter estimates in table 5.5, 5.1 and 5.1 respectively.

Kernel estimation and goodness of fit tests are usually based on subjective choices, both of function and of bandwidth. Tests which are based on "either" of these approaches have lower power which is an established and well known fact. We used the KS, AD, CVM, V, U^2 and χ^2 tests to evaluate how likely it was that the observed sample could have been generated from the distribution in question for the US, UK and CA GDP series.

In the χ^2, the time series data are divided into k class intervals (bins). The χ^2 test statistic is defined as follows

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Here O_i and E_i are the observed and expected number of observations in bin i. The statistic χ^2 is asymptotically distributed as chi-square with k-p degrees of freedom, where p is the number of parameters in the distribution. The χ^2 test is sensitive to the subjective choice of bins and does not have much power.

The KS, AD, CVM, V and U^2 goodness of fit tests are based on the empirical distribution function (EDF) and are often referred to as EDF tests. The EDF tests are more powerful than χ^2 goodness of fit test, see D'Agostino and Stephens (1986), Kotz and Nadarajah

(2000) and Famoye (2000). The AD and CVM are the most powerful tests among the EDF tests; see Kotz and Nadarajah (2000) and Famoye (2000).

The Kolmogorov-Smirnov (KS) test statistic is defined as the maximum value of the absolute difference between two cumulative distribution functions, that is:

$$KS = Max\,|\hat{F}(\tilde{z}_i) - F(\tilde{z}_i)|$$

where $\hat{F}(\tilde{z}_i)$ is the empirical CDF of data and $F(\tilde{z}_i)$ is the theoretical CDF of distribution, the KS test assumes that data comes from a continuous distribution. The KS statistic can be computed as

$$KS = \underset{0 \le i \le N}{Max}\left(\frac{i}{n} - F(\tilde{z}_i), F(\tilde{z}_i) - \frac{i-1}{n}\right)$$

The drawback with the KS test is that it is best suited for finding differences in the middle of the distributions.

The CVM test statistic can be computed as

$$CVM = \frac{1}{12n} + \sum_{i=1}^{n}\left(\frac{2i-1}{2n} - F(\tilde{z}_i)\right)^2$$

where $F(\tilde{z}_i)$ = the distribution function of $\tilde{z}$ and n is the sample size of the time series. The CVM test is more powerful than the KS test and can detect differences between the distributions over their entire range.

The U^2 is a modified version of CVM test. The U^2 test statistic is defined as

$$U^2 = CVM - n\left(\bar{F}(\tilde{z}_i) - \frac{1}{2}\right)^2$$

Where $\bar{F}(\tilde{z}_i) = \frac{\sum_{i=1}^{n} F(\tilde{z}_i)}{n}$ is the distribution function of $\tilde{z}$ and n is the sample size of the time series.

The AD test is a modification of CVM test. This test gives more weight to the tails than the KS test. The AD test statistic is computed as

$$AD = -n - \sum_{i=1}^{n}\left[\frac{2i-1}{n}\left(\log(1-F(\tilde{z}_{n-i+1})) + \log(F(\tilde{z}_i))\right)\right]$$

where $\tilde{z}_1 < ... < \tilde{z}_n$ is the sorted data and $F(\tilde{z}_i)$ is the cumulative distribution function of the specified distribution, the AD test can detect differences between the distributions over their entire width. One drawback of this test is that the distribution of the test statistic depends on the specific distribution being tested, so no general expressions can be given.

The V test is more closely related to KS test. The V test statistic can be computed as

$$V = Max\left[\frac{i}{n} - F(\tilde{z}_i)\right] + Max\left[F(\tilde{z}_i) - \frac{i-1}{n}\right]$$

This test is invariant under cyclic transformations of the independent variable and provides equal sensitivity at the tail as the median.

Table 5.7 reports on the P-values of KS, AD, VCM, V, U^2 and χ^2 tests when testing the null hypotheses $H_{o,1}: y^* \sim N$, $H_{o,2}: y^* \sim NM$, $H_{o,3}: y^* \sim NAL_1$, $H_{o,4}: y^* \sim NAL_2$, $H_{o,2}: y^* \sim TAL_1$ and $H_{o,2}: y^* \sim TAL_2$ for the US, UK and CA series.

Table 5.7: The goodness of fit tests(P-values)

For US						
	AD	KS	χ^2	CVM	V	U^2
N	0.5327	0.4446	0.1635	0.377	0.3970	0.3450
NM	0.9219	0.6340	0.0472	0.7887	0.5871	0.5584
NAL_1	0.8399	0.5398	0.0472	0.7430	0.7245	0.6999
NAL_2	0.9612	0.9084	0.1962	0.9150	0.7820	0.7822
TAL_1	0.8381	0.5368	0.0377	0.7407	0.7222	0.6976
TAL_2	**0.9754**	**0.9182**	**0.3069**	**0.9259**	**0.7967**	**0.8058**
For UK						
	AD	KS	χ^2	CVM	V	U^2
N	0.9550	0.9560	0.7204	0.9688	0.9336	0.9539
NM	**0.9982**	**0.9974**	**0.9516**	0.9954	**0.9921**	0.9798
NAL_1	0.9803	0.9155	0.6919	0.9602	0.9385	0.9611
NAL_2	0.9960	0.9802	0.8144	0.9944	0.9457	0.9816
TAL_1	0.9841	0.9006	0.6363	0.9632	0.9261	0.9510
TAL_2	0.9971	0.9884	0.9088	**0.9960**	0.9682	**0.9838**
For CA						
	AD	KS	χ^2	CVM	V	U^2
N	0.0587	0.1103	0.0485	0.0658	0.0469	0.0752
NM	**0.9725**	**0.9106**	0.4763	**0.9882**	**0.7514**	**0.9551**
NAL_1	0.6048	0.5243	0.0336	0.6908	0.2477	0.4045
NAL_2	0.8547	0.8954	**0.6103**	0.8747	0.70055	0.7077
TAL_1	0.5817	0.5090	0.0385	0.6899	0.2692	0.4118
TAL_2	0.5352	0.5574	0.0677	0.5021	0.2016	0.2462

AD: Anderson Darling, VCM: Cramer-Von Mises, KS: KolmogorovSmirnov, χ^2 : Pearson chi-square, U^2: Watson U-square and V : Kuiper

The result presented in table 5.7 clearly shows that, considering P-values of all the goodness of fit tests for the US GDP series, the TAL_2 fits better compared to other distributions, whereas NAL_2 has second best fit. For the UK GDP series, NM fits the data better compared to other distributions, except for VCM and U^2 test according to which TAL_2 fits the data best. Finally, for CA series NM fits the data better compared to other distributions according to all goodness of fit test except the χ^2 test.

The quantile-quantile (Q-Q) plot is a graphical method for assessing the goodness of fit of the distribution. The Q-Q plot is constructed by plotting the quintiles of the data of the empirical distribution versus the theoretical quantile of a distribution using ML estimates of the parameters.

The distribution fits the data well if in Q-Q plot the quantile of the data and theoretical quantile of the specific distribution roughly lie along the central diagonal, i.e. the plotted points fall on or close to the line with slope value one. The empirical quantiles are just the sorted observations of the data. The theoretical quantile Q_i corresponding to the i^{th} ordered observation is obtained by solving

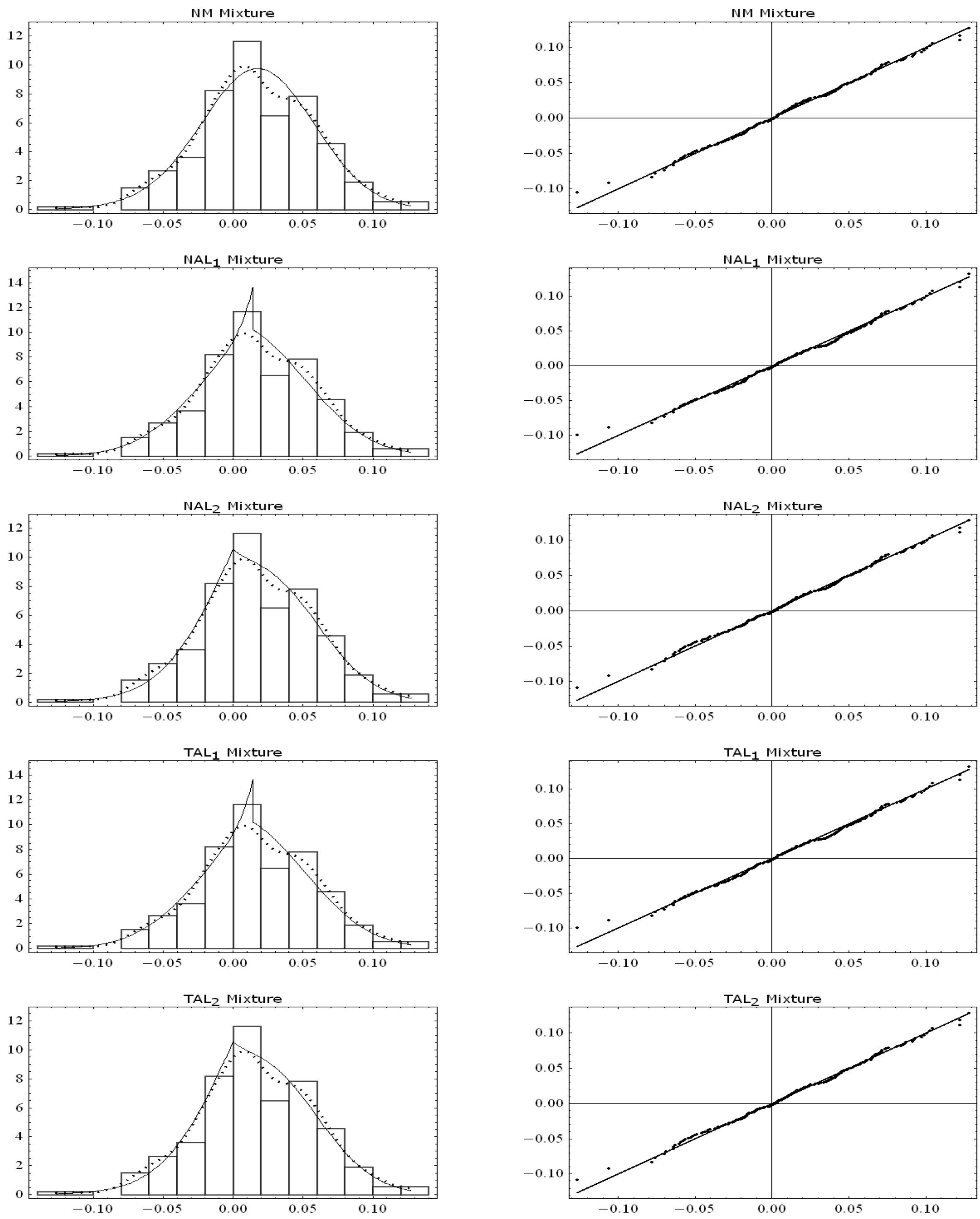
$$F(Q_i) = p_i$$

where $p_i = \frac{i}{n+1}$ and n is the number of observations therefore

$$Q_i = F^{-1}(p_i) \qquad (3.18)$$

Unfortunately, in many situations no closed-form exists for the inverse of the cdf of the distribution. So, equation (3.18) has to be solved by numerically using i.e. Secant method, the Newton's method, etc. We used Secant method to solve equation (3.18)

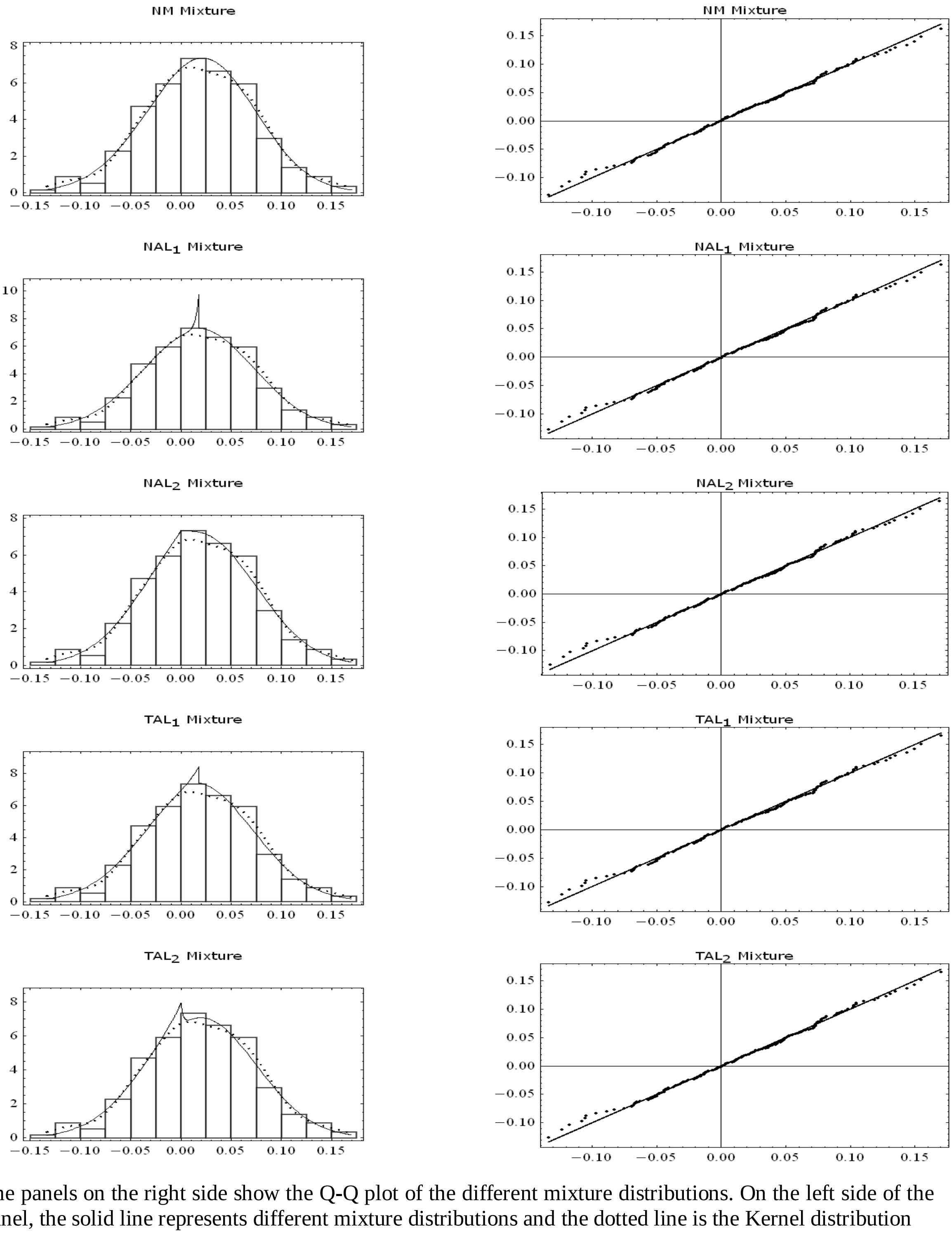
Figure 5.1: Goodness of fit plot for the US GDP series



The panels on the right side show the Q-Q plot of the different mixture distributions. On the left side of the panel, the solid line represents different mixture distributions and the dotted line is the Kernel distribution.

We can clearly see from above graph for the US GDP series TAL_2 density is closer to the Kernel density as compared to other distributions, and in the Q-Q plot points are closer to the line. This confirms table 5.6 and 5.7 that the TAL_2 distribution fits better to US GDP data compared to other distributions.

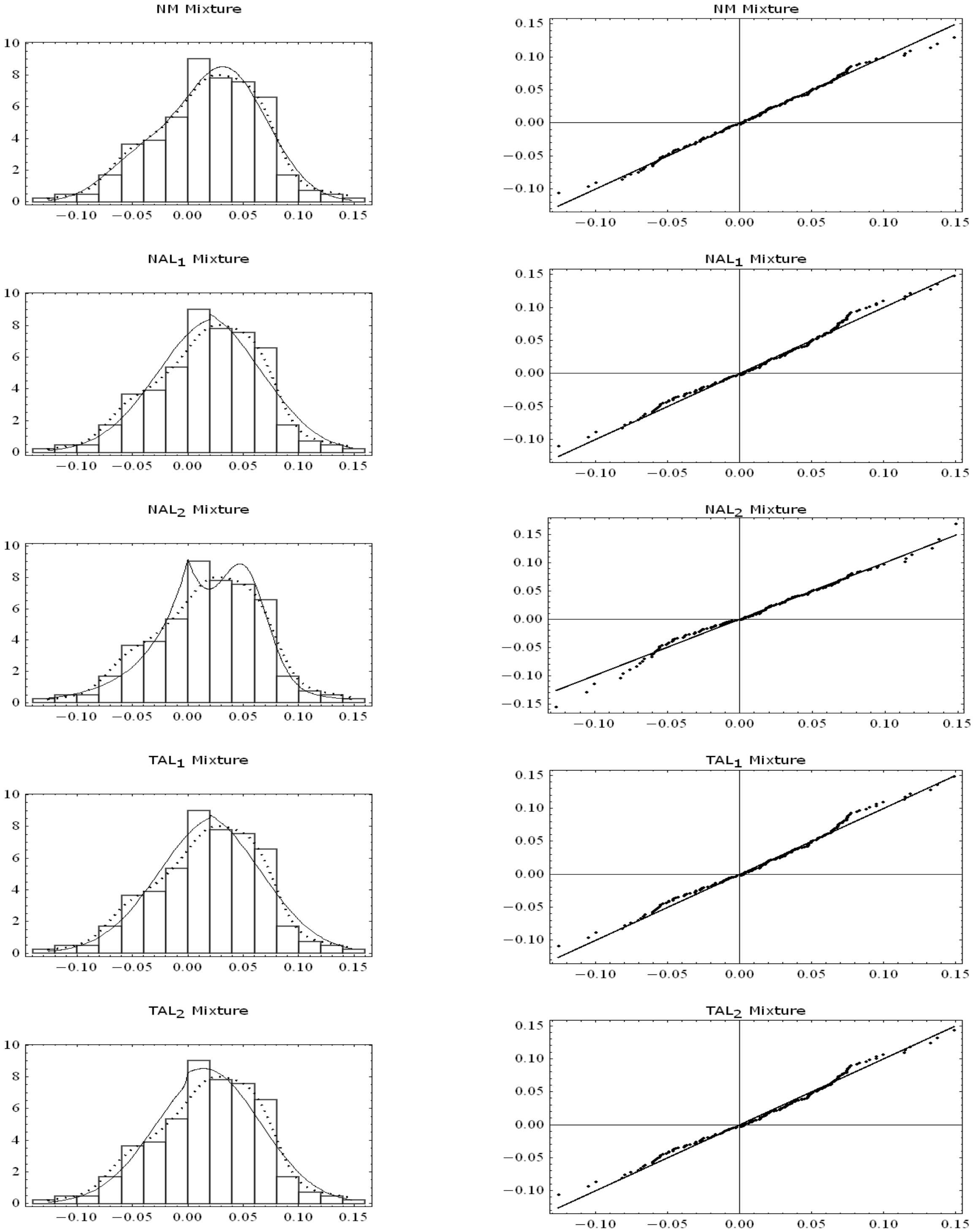
Figure 5.2: Goodness of fit plots for the UK GDP series.



The panels on the right side show the Q-Q plot of the different mixture distributions. On the left side of the panel, the solid line represents different mixture distributions and the dotted line is the Kernel distribution

From the above graph we can see that the NM density is closer to the Kernel density compared to other distributions. In the Q-Q plot the theoretical quantiles from NM, using the estimated parameter in table 5.1, are close to the line y=x, indicating that NM fits the data better compared to other distributions. Table 5.6 and 5.7 also confirm this.

Figure 5.3: Goodness of fit plots for the CA GDP series



The panel on the right side shows the Q-Q plot of the different mixture distributions. On the left side of the panel, the solid line represents different mixture distributions and the dotted line is the Kernel distribution.

The above graph clearly shows that the NM density is closer to the Kernel density compared to other distributions. The theoretical quantiles of the NM, using the ML estimates from table 5.1, are closer to 45-degree reference line in the Q-Q plot and indicates that the NM fits the data better. This can also be confirmed by table 5.6 and 5.7.

6 Conclusions

The growth rate of GDP has been found to exhibit heteroscedasticity, leptokurtosis (Fat tails) and skewness (asymmetry around the mean). Heteroscedasticity was removed by using the filter proposed by Stockhammar and Öller (2011).

The Laplace distribution and the asymmetric Laplace distribution are unable to explain the asymmetries and a slight leptokurtic shape. A mixed Student t Asymmetric Laplace-2 (TAL_2) distribution is introduced. For the US GDP, which is more skewed and leptokurtic, TAL_2 is shown to better describe the density distribution of growth than the N, NM, NAL_1, NAL_2, TAL_1 and L distributions. In the TAL_2 distribution, student's t distribution component was dominant. For UK and CA GDP series where data was skewed but slightly leptokurtic, the NM distribution showed better fit.

The TAL_2 implies a breakdown of the shocks into AL_2 and student's t components, and NM implies a breakdown into two normally distributed components. The six parameters of TAL_2 and the five parameters of NM are able to describe the mean, variance, skewness and kurtosis of the data. The ML estimates of the parameters of the distributions were estimated by the maximization of log likelihood. This was done by using the Nelder and Mead method.

Because of the close distributional fit, the TAL_2 and NM distributions are better choices for density forecasting. The GDP series has been studied in this thesis. These distributions could also prove useful in density forecasting of any heteroscedastic, asymmetric and leptokurtic time series.

References

Barro, R.J. (1991) Macroeconomic growth in a cross section of countries. *The Quarterly Jurnal of Macroeconomics*, 106, 407-444.

Blattberg, R.C. and Gonedes, N.J. (1974) A comparison of the stable and student distributions as statistical models for stock prices. *The Journal of Business*, 47, 244-280.

Bodman, P.M. (1998) Asymmetry and duration dependence in Australian GDP and unemployment. *Economic Record*, 74, 399-411.

Bottazzi, G., and Duenas, M. (2012) The Evolution of the Business Cycles and Growth Rates Distributions. Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Canning, D., Amaral, L.A.N., Lee, Y., Meyer, M. and Stanley, H.E. (1998) Scaling the volatility of GDP growth rates. *Economics Letters*, 60, 335–341.

Castaldi, C. and Dosi, G. (2009) The patterns of output growth of firms and countries: Scale invariances and scale specificities. *Empirical Economics*, 37, 475–495.

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

D'Agostino, R.B. and Stephens, M.A. (1986) *Goodness-of-Fit Techniques*. New York: Marcel Dekker.

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.

Dyer, A.R. (1974) Comparisons of tests for normality with a cautionary note. *Biometrika*, 61, 185-189.

Everitt, B. S. (1988) A finite mixture model for the clustering of mixed-mode data. *Statistics & probability letters*, *6*, 305-309.

Everitt, B.S. and Hand, D.J. (1981) *Finite Mixture Distributions*. Chapman and Hall.

Fagiolo, G., Napoletano, M. and Roventini, A. (2008) Are output Growth-Rate distributions fat- tailed? Some evidence from OECD. *Journal of Applied Econometrics*, 23, 639–669.

Famoye, F. (2000) Goodness-of-fit tests for generalized logarithmic series distribution. *Computational Statistics and Data Analysis*, 33, 59–67.

Frühwirth-Schnatter, S. (2006) *Finite Mixture and Markov Switching Models*. NY: Springer.

Gridgeman, N.T. (1970) A Comparison of Two Methods of Analysis of Mixtures of Normal Distributions. *Technometrics*, 12, 823-833.

Harandi, S.S. and Alamatsaz, M.H. (2013) Alpha-Skew-Laplace distribution. *Statistics and Probability Letters*, 83, 774-782.

Hass, M., Mittnik, S. and Paolella, M.S. (2006) Modelling and predicting market risk with Laplace- Gaussian mixture distributions. *Applied Financial Economics*, 16, 1145-1162.

Hess, G.D. and Iwata, S. (1997) Asymmetric persistence in GDP? A deeper look at depth. *Journal of Monetary Economics*, 40, 535-554.

Hodrick, R.J. and Prescott, E.C. (1997) Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29, 1-16.

Hyndman, R.J. and Koehler, A.B. (2006) Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22, 679-688.

Jayakumar, K. and Kuttykrishnan, A.P. (2007) A time-series model using asymmetric Laplace distribution. *Statistics and Probability Letters*, 77, 1636–1640.

Jones, P.N. and McLachlan, G.J. (1990) Laplace-normal mixtures fitted to wind shear data. *Journal of Applied Statistics*, 17, 271-276.

Julia, O. and Vives-Rego, J. (2008) A microbiology application of the skew-Laplace distribution. *Sort: Statistics and Operations Research Transactions*, 32, 141-150.

Kamaruzzaman, Z.A., Isa, Z. and Ismail M.T. (2012) Mixtures of normal distributions: Application to Bursa Malaysia Stock Market Indices. *World Applied Sciences Journal*, 16, 781-790.

Kanji, G.K. (1985) A mixture model for wind shear data. *Journal of Applied Statistics*, 12, 49-58.

Klar, B. and Meintanis, S.G. (2005) Tests for normal mixtures based on the empirical characteristic function. *Computational Statistics and Data Analysis*, 49, 227-242.

Kon, S.J. (1984) Models of stock returns a comparison. *The Journal of Finance*, 39, 147-165.

Kotz, S. and Nadarajah, S. (2000) *Extreme Value Distributions: Theory and Applications*. London: Imperial College Press.

Kotz, S., Kozubowski, T.J. and Podgorski, K. (2001) *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Exonomics, Engineering, and Finance*. Birkhäuser, Boston.

Kozubowski, T.J. and Nadarajah, S. (2010) Multitude of Laplace distribution. *Statistical Papers*, 51,127–148.

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

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

Lagarias, J.C. Reeds, J. A. Wright, M. H. and Wright, P. E. (1998) Convergence Properties of the Nelder-Mead Simplex Algorithm in Low Dimensions. *SIAM Journal on Optimization*, 9, 112–147.

Lee, Y., Amaral, L.A.N., Canning, D., Meyer, M. and Stanley, H.E. (1998) Universal features in the growth dynamics of complex organizations. *Physical Review Letters*, 81, 3275–3278.

Lewis, R. M., Torczon, V. and Trosset, M. W. (2000) Direct search methods: then and now. *Journal of Computational and Applied Mathematics*, *124*, 191-207.

Linden, M. (2001) A model for stock return distribution. *International Journal of Finance and Economics*, 6, 159-169.

Lindsay, B. G. (1995) Mixture models: theory, geometry, and applications, Volume 5 of NSF-CBMS Regional Conference Series in Probability and Statistics. *Institute for Mathematical Statistics: Hayward, CA*.

Makridakis, S. (1993) Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, 9, 527-529.

McGill, W.J. (1962) Random fluctuations of response rate. *Psychometrika*, 27, 3-17.

McLachlan, G. and Peel, D. (2000) *Finite Mixture Models*. NY: John Wiley & Sons.

McLachlan, G.J. and Basford, K.E. (1988) *Mixture Models: Inference and Applications to Clustering*. NY: Marcel Dekker.

Mengersen, K., Robert, C. and Titterington, M. (2011) *Mixture: Estimation and Application*. NY: John Wiley & Sons

Narayan, P.K. (2009) Are health expenditures and GDP characterized by asymmetric behaviour? Evidence from 11 OECD countries. *Applied Economics*, 41, 531–536.

Nelder, J.A. and Mead, R. (1965) A simplex method for function minimization. *Computer Journal*, 7, 308-313.

Newcomb, S. (1963) A generalized theory of combination of observations so as to obtain the best result. *American Journal of Mathematics*, 8, 343-366.

Pearson, K. (1895) Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London. A*, 186, 343-414.

Praetz, P.D. (1972) The distribution of share price changes. *Journal of Business*, 45, 49-55.

Press, S.J. (1967) A compound events model for security prices. *Journal of Business*, 40, 317-335.

Price, C. J., Coope, I. D., and Byatt, D. (2002) A convergent variant of the Nelder–Mead algorithm. *Journal of Optimization Theory and Applications*.113, 5-19.

Razali, N.M. and Wah Y.B. (2011) Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling test. *Journal of Statistical Modelling and Analytics*, 2, 21-33.

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, 1733-1753.

Schlattmann, P. (2009) *Medical Applications of Finite Mixture Models*. NY: Springer.

Silverman, B.W. (1986) *Density estimation for statistics and data analysis*. Chapman and Hall, London.

Stockhammar, P. and Öller, L.-E. (2011) On the probability distribution of economic growth. *Journal of Applied Statistics,* 38, 2023-2041.

Stockhammar, P. and Öller, L.-E. (2011) A heteroscedasticity removing filter. *Communications in Statistics -Theory and Methods,* 41, 281-299.

Tay, A. S. and Wallis, K. F. (2002) Density forecasting: a survey. *Companion to Economic Forecasting*, 45-68.

Thadewald, T. and Buning, H. (2007) Jarque–Bera Test and its competitors for testing normality -A Power Comparison. *Journal of Applied Statistics*, 34, 87–105.

Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985) *Statistical Analysis of finite mixture distributions*. New York : John Wiley & Sons.

Wirjanto, T. S., & Xu, D. (2009) The Applications of Mixtures of Normal Distributions in Empirical Finance: A Selected Survey. Working paper, University of Waterloo.

Wright, M. H. (2012) Nelder, Mead, and the other simplex method. *Documenta Math. Extra Volume ISMP*.

Wright, M.H. (1995) Direct search methods: Once scorned, now respectable, in Numerical Analysis, Proceedings. 1995 Dundee Bienneal Conference in Numerical Analysis, Griffiths, D.F. and Watson, G.A. eds., 1996, *Addison–Wesley Longman, Harlow,* U.K., 191–208.

Wu, F. (2008) *Applications of the normal laplace and generalized normal laplace distributions* (Doctoral dissertation, University of Victoria).

Yu, K. and Zhang, J. (2005) A three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics-Theory and Methods,* 34, 1867-1879.

Appendix

R and Mathematica code for US GDP Series

#R code for table 2.1

```
library(Rcmdr);library(stats);library(class);library(e1071);library(urca);library(FinTS);

y<- read.table(file.choose(), header=TRUE, sep=",",na.strings="NA", dec=".", strip.white=TRUE)

data<-diff(log(y$US),lag=1)

mean(data)

sd(data)

skewness(data,type=2)

kurtosis(data,type=2)

ArchTest (data, lags=2, demean = FALSE)

summary(ur.df(data, type = c("drift"), lags = 0))
```

 # Mathematica code for figure 2.1

```
uuss=Import["CDG.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;

G=SmoothKernelDistribution[data];g1=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},

ImagePadding->20,ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Diff in US GDPq 2007-2012","Label"

,Small]]],PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.0160,\[Sigma]->0.0113}

,{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G,y],{y,Min[data],Max[data]}

,PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g11=ListLinePlot[data,DataRange->{1947,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},

PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Diff in US GDPq 2007-2012","Label",Small]]];

ukuk=Import["CDG.xlsx",{"xlsx","Data",2}];TableView[ukuk];data1=ukuk[[2;;230,{1}]]//Flatten;G1=SmoothKernelDistribution[data1];

g2=Show[Histogram[data1,15,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},

ImagePadding->20,ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Diff in UK GDPq 1955-2012","Label",

Small]]],PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.0192,\[Sigma]->0.0188},

{x,Min[data1],Max[data1]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G1,y],{y,Min[data1],

Max[data1]},PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g22=ListLinePlot[data1,DataRange->{1955,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},
```

```
PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Diff in UK GDPq 1955-2012","Label",Small]]];

caca=Import["CDG.xlsx",{"xlsx","Data",3}];TableView[caca];data3=caca[[2;;206,{1}]]//Flatten;

G3=SmoothKernelDistribution[data3];

g3=Show[Histogram[data3,14,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,

ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Diff in Canada GDPq 1961-2012","Label",Small]]],

PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.019,\[Sigma]->0.013},{x,Min[data3],

Max[data3]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G3,y],{y,Min[data3],Max[data3]},

PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g33=ListLinePlot[data3,DataRange->{1961,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},

PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Diff in Canada GDPq 1961-2012","Label",Small]]];

Show[GraphicsGrid[{{g1,g11},{g2,g22},{g3,g33}}]]
```

 # R code for filter proposed by Stockhammar and Oller

```
library(Rcmdr);library(stats);library(class);library(e1071);library(mFilter)

library(bitops);library(caTools)

data<- read.table(file.choose(), header=TRUE, sep=",",na.strings="NA", dec=".", strip.white=TRUE)

y<-diff(log(data$US),lag=1);ma<-runmean(y,15);z<-y-ma

ma2<- runmean((z^2),15);sqma2<-sqrt(ma2/14);hp1<-hpfilter(sqma2,freq=1600)

hp<-as.vector(hp1$trend);Req1<-z/hp

sr<-(Req1*sd(y))+mean(y);aa<-as.data.frame(sr)

write.table(aa, "C:/Users/MAHMOOD/Desktop/USZ.csv", sep=",", col.names=TRUE, row.names=TRUE, quote=TRUE, na="NA")
```

#R code for table 3.1

```
mean(aa$sr);sd(aa$sr)

kurtosis(aa$sr,type=2);skewness(aa$sr,type=2)

library(FinTS)

ArchTest (aa$sr, lags=2, demean = FALSE)
```

#R code for table 3.2

```
mean(aa$sr[1:131]);mean(aa$sr[132:262])

sd(aa$sr[1:131]);sd(aa$sr[132:262])

skewness(aa$sr[1:131],type=2);skewness(aa$sr[132:262],type=2)
```

kurtosis(aa$sr[1:131],type=2);kurtosis(aa$sr[132:262],type=2)

#R code for table 3.3

library(zoo);library(quadprog);library(tseries);library(nortest);

ad.test(aa$sr);shapiro.test(aa$sr);lillie.test(aa$sr);jarque.bera.test(aa$sr)

pearson.test(aa$sr);cvm.test(aa$sr);sf.test(aa$sr)

#Mathematica Code for figure 3.1

uuss=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;

G=SmoothKernelDistribution[data];g1=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Filtered Diff in US GDPq 2007-2012","Label",Small]]],PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.01629,\[Sigma]->0.04158},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g11=ListLinePlot[data,DataRange->{1947,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Filtered Diff in US GDPq 2007-2012","Label",Small]]];

ukuk=Import["GDPZ.xlsx",{"xlsx","Data",2}];TableView[ukuk];data1=ukuk[[2;;230,{1}]]//Flatten;

G1=SmoothKernelDistribution[data1];g2=Show[Histogram[data1,20,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Filtered Diff in UK GDPq 1955-2012","Label",Small]]],PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.01905,\[Sigma]->0.05501},{x,Min[data1],Max[data1]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G1,y],{y,Min[data1],Max[data1]},PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g22=ListLinePlot[data1,DataRange->{1955,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Filtered Diff in UK GDPq 1955-2012","Label",Small]]];

caca=Import["GDPZ.xlsx",{"xlsx","Data",3}];TableView[caca];

data3=caca[[2;;206,{1}]]//Flatten;G3=SmoothKernelDistribution[data3];

g3=Show[Histogram[data3,14,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame-> True,PlotLabel->Text[Style["Filtered Diff in Canada GDPq 1961-2012","Label",Small]]],PDFplot=Plot[PDF[NormalDistribution[\[Mu],\[Sigma]],x]/.{\[Mu]->0.0183065,\[Sigma]->0.0482706},{x,Min[data3],Max[data3]},PlotStyle->{Thin,Black},PlotRange->All],Plot[PDF[G3,y],{y,Min[data3],Max[data3]},PlotStyle->{Thick,Dotted,Black},PlotRange->All,Frame->True,Axes->False]];

g33=ListLinePlot[data3,DataRange->{1961,2012},Frame->True,Axes->False,PlotStyle->{Thin,Black},PlotRange->All,ImageSize->{290,175},ImagePadding->20,PlotLabel->Text[Style["Filtered Diff in Canada GDPq 1961-2012","Label",Small]]];

Show[GraphicsGrid[{{g1,g11},{g2,g22},{g3,g33}}]]

 # Mathematica Code for NM distribution table 5.1 and table 5.6

nn=MixtureDistribution[{p1,p2},{NormalDistribution[Subscript[\[Mu], 1],Subscript[\[Sigma], 1]],NormalDistribution[Subscript[\[Mu], 2],Subscript[\[Sigma], 2]]}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;

```
G=SmoothKernelDistribution[data];div=FindDivisions[{Min[data],Max[data]},1000];
```

 # Function for maximization of log likelihood with accuracy measure

```
result[data_,w_]:=Module[{n=Length[data],sdata=Sort[data],m=Mean[data],s2=Variance[data],optimalMixture,quantilesMixture,RMSE1,RMSE,rr,MDAPE,sMDAPE,MASE},

optimalMixture=Quiet[Check[NMaximize[{Total[Log[PDF[nn,data]]],Join[{Subscript[\[Sigma], 1]>0},{Subscript[\[Sigma], 2]>0},{0<p1<1},{0<p2<1},{p1+p2==1}]},Join[{{Subscript[\[Sigma], 1],Sqrt[s2]-.01,Sqrt[s2]+.01},{Subscript[\[Sigma], 2],Sqrt[s2]-.01,Sqrt[s2]+.01},{Subscript[\[Mu], 1],Mean[data]-.1,Mean[data]+.1},{Subscript[\[Mu], 2],Mean[data]-.1,Mean[data]+.1}},{{p1,w-.1,w+.1}},{{p2,w-.1,w+.1}}]],None]];

quantilesMixture:=Quiet[Table[x/.FindRoot[(CDF[nn,x]/.optimalMixture[[2]])==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]];

RMSE1:=Row[{" RMSE1 = ",Chop[RootMeanSquare[quantilesMixture-sdata],10^-6]}];

a=PDF[G,div];b=PDF[nn,div]/.optimalMixture[[2]];RMSE=Sqrt[Total[(a-b)^2]/1000];MDAPE=Median[(100*Abs[(a-b)])/(a)];sMDAPE=Median[(200*Abs[(a-b)])/(a+b)];MASE=((1/1000)*Total[Abs[a-b]])/((1/999)*Total[Abs[Differences[a]]]);rr:={optimalMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE};rr]
```

 # Run the above function for different value of weight (P) to find out optimal solution and accuracy measures

```
uu={result[data,0.10],result[data,0.20],result[data,0.30],result[data,0.40],result[data,0.50],result[data,0.60],result[data,0.70],result[data,0.80],result[data,0.90]}
```

#Mathematica Code to find simmulated Standard Error for parameter estimate of NM table 5.1

```
bootstrap=Table[prob=RandomVariate[BinomialDistribution[1,0.8812438502042144],Length[data]];

data1=prob*(RandomVariate[NormalDistribution[0.018406036847296987,0.0394500521564434],Length[data]])+(1-prob)*(RandomVariate[NormalDistribution[0.0019546485083310966,0.05501044895827489],Length[data]]);

{Subscript[\[Sigma], 1],Subscript[\[Sigma], 2],Subscript[\[Mu], 1],Subscript[\[Mu], 2],p}/.Last[Quiet[Check[NMaximize[{Total[Log[PDF[nn,data1]]],Join[{Subscript[\[Sigma], 1]>0},{Subscript[\[Sigma], 2]>0},{0<p<1}]},Join[{{Subscript[\[Sigma], 1],StandardDeviation[data1]-.01,StandardDeviation[data1]+.01},{Subscript[\[Sigma], 2],StandardDeviation[data1]-.01,StandardDeviation[data1]+.01},{Subscript[\[Mu], 1],Mean[data1]-.1,Mean[data1]+.1},{Subscript[\[Mu], 2],Mean[data1]-.1,Mean[data1]+.1}},{{p,0.35-.1,0.35+.1}}]],{{Subscript[\[Sigma], 1]-> 0,Subscript[\[Sigma], 2]-> 0,Subscript[\[Mu], 1]-> 0,Subscript[\[Mu], 2]-> 0,p-> 0}}]]],{1000}];

StandardDeviation[DeleteCases[bootstrap,{0,0,0,0,0}]]
```

 # Mathematica code to find the goodness of fit test for NM distribution table 5.7

```
bb=nn/.{Subscript[\[Sigma], 1]->0.0394500521564434`,Subscript[\[Sigma], 2]->0.05501044895827489`,Subscript[\[Mu], 1]->0.018406036847296987`,Subscript[\[Mu], 2]->0.0019546485083310966`,p->0.8812438502042144`};

\[ScriptCapitalH]=DistributionFitTest[data,bb,"HypothesisTestData"];

\[ScriptCapitalH]["TestDataTable",All]
```

 # Mathematica Code for NAL$_1$ distribution table 5.2 and table 5.6

```
ALL=ProbabilityDistribution[f11,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{NormalDistribution[\[Mu],\[Sigma]],ALL}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];

TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;G=SmoothKernelDistribution[data];

div=FindDivisions[{Min[data],Max[data]},1000];FindDistributionParameters[data,ALL1,ParameterEstimator -> "MethodOfMoments"]

f1=Piecewise[{{Exp[(x-0.017034650008291793)/\[Psi]]/(2\[Psi]), x<=0.017034650008291793`}}, Exp[(0.017034650008291793`-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f1,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{NormalDistribution[0.017034650008291793,\[Sigma]],ALL}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];

data=uuss[[2;;263,{1}]]//Flatten;
```

 # Function for maximization of log likelihood with accuracy measure

```
result[data_,w_]:=

Module[{n=Length[data],sdata=Sort[data],m=Mean[data],s2=Variance[data],optimalMixture,quantilesMixture,RMSE1,RMSE, MDAPE, sMDAPE, MASE,rr},

optimalMixture=Quiet[Check[NMaximize[{Total[Log[PDF[nn,data]]],Join[{\[Sigma]>0},{ \[Phi]>0,\[Psi]>0},{0<p<1}}],Join[{{\[Sigma],Sqrt[s2]-.01,Sqrt[s2]+.01},{\[Phi],0.0286-.01,0.0286+.01},{\[Psi],0.03-0.01,0.03+.01}},{{p,w-.1,w+.1}}]],None]];

quantilesMixture:=Quiet[Table[x/.FindRoot[(CDF[nn,x]/.optimalMixture[[2]])==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]];

RMSE1:=Row[{" RMSE1 = ",Chop[RootMeanSquare[quantilesMixture-sdata],10^-6]}];

a=PDF[G,div];b=PDF[nn,div]/ .optimalMixture[[2]];RMSE=Sqrt[Total[(a-b)^2]/1000];

MDAPE=Median[(100*Abs[(a-b)])/(a)];sMDAPE=Median[(200*Abs[(a-b)])/(a+b)];

MASE=((1/1000)*Total[Abs[a-b]])/((1/999)*Total[Abs[Differences[a]]]);

rr:={optimalMixture,RMSE1, RMSE,MDAPE,sMDAPE,MASE };rr]
```

 # Run the above function for different value of weight (P) to find out optimal solution and accuracy measures

```
uu={result[data,0.10],result[data,0.20],result[data,0.30],result[data,0.40],result[data,0.50],result[data,0.60],result[data,0.70],result[data,0.80],result[data,0.90]}
```

#Mathematica Code to find simulated Standard Error for parameter estimate of NAL1 table 5.2

```
AL1rn[n_,\[Mu]_,\[Psi]_,\[Phi]_]:=Module[{vec},

vec=Table[0,{n}];

For[i=0,i<n,{x=RandomVariate[UniformDistribution[{0,1}],{1}],If[ x[[1]]<= 0.5,
```

```
vec[[i]]=\[Mu]+\[Psi] Log[2 x[[1]]]],vec[[i]]=\[Mu]-\[Phi] Log[2 (1-x[[1]])]},i++];vec]

f33=Piecewise[{{Exp[(x-Median[data1])/\[Psi]]/(2\[Psi]), x<=Median[data1]}}, Exp[(Median[data1]-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f33,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{NormalDistribution[Median[data1],\[Sigma]],ALL}];

bootstrap=Table[prob=RandomVariate[BinomialDistribution[1, 0.8761693288029527],Length[data]];

data1=prob*(RandomVariate[NormalDistribution[0.0140432,0.04308871498517691],262])+

(1-prob)*(AL1rn[262,0.0140432,0.011195661894971605,0.029432059190111725]);

{\[Mu]=Median[data1],\[Sigma],\[Phi],\[Psi],p}/.Last[Quiet[Check[NMaximize[{Total[Log[PDF[nn,data1]]],

Join[{\[Sigma]>0},{ \[Phi] >0,\[Psi]>0},{0<p<1}]},Join[{{\[Sigma],StandardDeviation[data1]-.01,

StandardDeviation[data1]+.01},{\[Phi],0.029432059190111725-.01,0.029432059190111725+.01},

{\[Psi],0.011195661894971605-0.01,0.011195661894971605+.01}},{{p,0.8761693288029527-.1,0.8761693288029527+.1}}]],

{ {\[Sigma]-> 0,\[Phi]-> 0,\[Psi]-> 0,p-> 0}}]]],{1000}];

StandardDeviation/@Transpose[DeleteCases[bootstrap,{0,0,0,0,0}]]
```

 # Mathematica code to find the goodness of fit test for NAL1 table 5.7

```
bb=nn/.{\[Mu]-> 0.0140432,\[Sigma]->0.04308871498517691`,\[Phi]->0.029432059190111725`,\[Psi]->0.011195661894971605`,p->0.8761693288029527`};

\[ScriptCapitalH]=DistributionFitTest[data,bb,"HypothesisTestData"];

\[ScriptCapitalH]["TestDataTable",All]
```

 # Mathematica Code for NAL_2 distribution table 5.3 and table 5.6

```
f1=Piecewise[{{\[Alpha] \[Beta] Exp[\[Beta] x]/(\[Alpha]+\[Beta]), x<=0}},\[Alpha] \[Beta] Exp[-\[Alpha] x]/(\[Alpha]+\[Beta])];

f2=ProbabilityDistribution[f1,{x,-Infinity,Infinity},Assumptions->{ \[Alpha]>0,\[Beta]>0}];

nn=MixtureDistribution[{p,1-p},{NormalDistribution[\[Mu],\[Sigma]],f2}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;

G=SmoothKernelDistribution[data];div=FindDivisions[{Min[data],Max[data]},1000];
```

 # Function for maximization of log likelihood with accuracy measure

```
result[data_,w_]:=Module[{n=Length[data],sdata=Sort[data],m=Mean[data],s2=Variance[data],optimalMixture,quantilesMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE,a,b,rr},

optimalMixture:=Quiet[Check[NMaximize[{Total[Log[PDF[nn,data]]],Join[{ \[Alpha]>0,\[Beta]>0},{\[Sigma]>0},{0<p<1}]},Join[{{\[Sigma],StandardDeviation[data1]-.1,StandardDeviation[data1]+.1},{\[Mu],Mean[data1]-.1,Mean[data1]+.1},{\[Alpha],(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])-0.1,(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+0.1},{\[Beta],1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2
```

```
Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))-0.1,1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))+0.1}},{{p,0.20-.1,0.20+.1}}]],None]];

quantilesMixture:=Quiet[Table[x/.FindRoot[(CDF[nn,x]/.optimalMixture[[2]])==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]];

RMSE1:=Row[{" RMSE1 = ",Chop[RootMeanSquare[quantilesMixture-sdata],10^-6]}];

a=PDF[G,div];b=PDF[nn,div]/ .optimalMixture[[2]];RMSE=Sqrt[Total[(a-b)^2]/1000];

MDAPE=Median[(100*Abs[(a-b)])/(a)];sMDAPE=Median[(200*Abs[(a-b)])/(a+b)];

MASE=((1/1000)*Total[Abs[a-b]])/((1/999)*Total[Abs[Differences[a]]]);

rr:={optimalMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE};rr]
```

 # Run the above function for different value of weight (P) to find out optimal solution and accuracy measures

```
uu={result[data,0.10],result[data,0.20],result[data,0.30],result[data,0.40],result[data,0.50],result[data,0.60],result[data,0.70],result[data,0.80],result[data,0.90]}
```

#Mathematica Code to find simmulated Standard Error for parameter estimate of NAL$_2$ table 5.3

```
par={\[Sigma]->0.040239444476853875`,\[Mu]->0.02251649574543454`,\[Alpha]->48.172434462297986`,\[Beta]->32.35877311627851`,p->0.8092099118475501`};

bootstrap=Table[prob=RandomVariate[BinomialDistribution[1, 0.8092099118475501],Length[data]];

data1=(1-prob)(RandomVariate[ExponentialDistribution[48.172434462297986],Length[data]]-RandomReal[ExponentialDistribution[32.35877311627851],Length[data]])+prob*(RandomVariate[NormalDistribution[0.02251649574543454,0.040239444476853875],Length[data]]);

{\[Sigma],\[Mu],\[Alpha],\[Beta],p}/.Last[Quiet[Check[NMaximize[{Total[Log[PDF[nn,data1]]],Join[{ \[Alpha] >0,\[Beta]>0},{\[Sigma]>0},{0<p<1}]},Join[{{\[Sigma],StandardDeviation[data1]-.1,StandardDeviation[data1]+.1},{\[Mu],Mean[data1]-.1,Mean[data1]+.1},{\[Alpha],(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])-0.1,(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+0.1},{\[Beta],1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))-0.1,1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))+0.1}},{{p,0.20-.1,0.20+.1}}]],{ {\[Sigma]-> 0,\[Mu]-> 0,\[Alpha]-> 0,\[Beta]-> 0,p-> 0}}]]],{1000}];

StandardDeviation[DeleteCases[bootstrap,{0,0,0,0,0}]]
```

 # Mathematica code to find the goodness of fit test for NAL2 distribution table 5.7

```
bb=nn/.{\[Sigma]->0.040239444476853875`,\[Mu]->0.02251649574543454`,\[Alpha]->48.172434462297986`,\[Beta]->32.35877311627851`,p->0.8092099118475501`};

\[ScriptCapitalH]=DistributionFitTest[data,bb,"HypothesisTestData"];

\[ScriptCapitalH]["TestDataTable",All]
```

 # Mathematica Code for TAL1 distribution table 5.4 and table 5.6

```
f1=Piecewise[{{Exp[(x-\[Mu])/\[Psi]]/(2\[Psi]), x<=\[Mu]}}, Exp[(\[Mu]-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f1,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ALL}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;

G=SmoothKernelDistribution[data];div=FindDivisions[{Min[data],Max[data]},1000];

FindDistributionParameters[data,ALL,ParameterEstimator -> "MethodOfMoments"]

f1=Piecewise[{{Exp[(x-0.017034650008291793)/\[Psi]]/(2\[Psi]), x<=0.017034650008291793`}}, Exp[(0.017034650008291793`-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f1,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{StudentTDistribution[0.017034650008291793,\[Sigma],\[Nu]],ALL}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;
```

 # Function for maximization of log likelihood with accuracy measure

```
result[data_,w_]:=Module[{n=Length[data],sdata=Sort[data],m=Mean[data],s2=Variance[data],optimalMixture,quantilesMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE,rr},

optimalMixture=Quiet[Check[NMaximize[{Total[Log[PDF[nn,data]]],Join[{\[Sigma]>0,\[Nu]>1},{ \[Phi]>0,\[Psi]>0},{0<p<1}]},Join[{{\[Sigma],0.0411451-.01,0.0411451+.01},{\[Nu],96.8119-1,96.8119+1},{\[Phi],0.0286449-.01,0.0286449+.01},{\[Psi],0.030144-0.01,0.030144+.01}},{{p,w-.1,w+.1}}]],None]];

quantilesMixture:=Quiet[Table[x/.FindRoot[(CDF[nn,x]/.optimalMixture[[2]])==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]];

RMSE1:=Row[{" RMSE1 = ",Chop[RootMeanSquare[quantilesMixture-sdata],10^-6]}];

a=PDF[G,div];b=PDF[nn,div]/ .optimalMixture[[2]];RMSE=Sqrt[Total[(a-b)^2]/1000];

MDAPE=Median[(100*Abs[(a-b)])/(a)];sMDAPE=Median[(200*Abs[(a-b)])/(a+b)];

MASE=((1/1000)*Total[Abs[a-b]])/((1/999)*Total[Abs[Differences[a]]]);

rr:={optimalMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE};rr]
```

 # Run the above function for different value of weight (P) to find out optimal solution and accuracy measures

```
uu={result[data,0.10],result[data,0.20],result[data,0.30],result[data,0.40],result[data,0.50],result[data,0.60],result[data,0.70],result[data,0.80],result[data,0.90]}
```

#Mathematica Code to find simmulated Standard Error for parameter estimate of TAL1 table 5.4

```
uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];

TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;AL1rn[n_,\[Mu]_,\[Psi]_,\[Phi]_]:=Module[{vec},

vec=Table[0,{n}];For[i=0,i<n,{x=RandomVariate[UniformDistribution[{0,1}],{1}],If[ x[[1]]<= 0.5,

vec[[i]]=\[Mu]+\[Psi] Log[2 x[[1]]]],vec[[i]]=\[Mu]-\[Phi] Log[2 (1-x[[1]])]},i++];vec]

f33=Piecewise[{{Exp[(x-Median[data1])/\[Psi]]/(2\[Psi]), x<=Median[data1]}}, Exp[(Median[data1]-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f33,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

nn=MixtureDistribution[{p,1-p},{StudentTDistribution[Median[data1],\[Sigma],\[Nu]],ALL}];

bootstrap=Table[prob=RandomVariate[BinomialDistribution[1, 0.8776600352724089],Length[data]];

data1= prob *(RandomVariate[StudentTDistribution[ 0.0140432,0.042989398728593996,561.6095472598373],Length[data]])+(1-prob)*(AL1rn[Length[data],0.0140432,0.011141022213614358,0.029497763511069332]);

{\[Mu]=Median[data1],\[Sigma],\[Nu],\[Phi],\[Psi],p}/.Last[Quiet[Check[NMaximize[{Total[Log[PDF[nn, data1]]],Join[{\[Sigma]>0,\[Nu]>1},{ \[Phi] >0,\[Psi]>0},{0<p<1}]},Join[{{\[Sigma],0.0411451-.01,0.0411451+.01},{\[Nu],96.8119-1,96.8119+1},{\[Phi],0.0286449-.01,0.0286449+.01},{\[Psi],0.030144-0.01,0.030144+.01}},{{p,0.35-.1,0.35+.1}}]],{ {\[Sigma]-> 0,\[Nu]-> 0,\[Phi]-> 0,\[Psi]-> 0,p-> 0}}]]]],{1000}];

StandardDeviation/@Transpose[DeleteCases[bootstrap,{0,0,0,0,0,0}]]
```

 # Mathematica code to find the goodness of fit test for TAL1 distribution table 5.7

```
bb=nn/.{\[Mu]->0.0140432,\[Sigma]->0.042989398728593996`,\[Nu]->561.6095472598373`,\[Phi]->0.029497763511069332`,\[Psi]->0.011141022213614358`,p->0.8776600352724089`};

\[ScriptCapitalH]=DistributionFitTest[data,bb,"HypothesisTestData"];

\[ScriptCapitalH]["TestDataTable",All]
```

 # Mathematica Code for TAL2 distribution table 5.5 and table 5.6

```
f1=Piecewise[{{\[Alpha] \[Beta] Exp[\[Beta] x]/(\[Alpha]+\[Beta]), x<=0}},\[Alpha] \[Beta] Exp[-\[Alpha] x]/(\[Alpha]+\[Beta])];

f2=ProbabilityDistribution[f1,{x,-Infinity,Infinity},Assumptions->{ \[Alpha]>0,\[Beta]>0}];

nn=MixtureDistribution[{p,1-p},{StudentTDistribution[\[Mu],\[Sigma],\[Nu]],f2}];

uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;
```

 # Function for maximization of log likelihood with accuracy measure

```
result[data_,w_]:=

Module[{n=Length[data],sdata=Sort[data],m=Mean[data],s2=Variance[data],optimalMixture,quantilesMixture,RMSE1,RMSE,rr,MDAPE,sMDAPE,MASE},

optimalMixture=Quiet[Check[NMaximize[{Total[Log[PDF[nn,data]]],Join[{\[Sigma]>0,\[Nu]>1},{ \[Alpha] >0,\[Beta]>0},{0<p<1}]},Join[{{\[Mu],(\[Mu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-.1,(\[Mu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -
```

>"MethodOfMoments"])+.1},{\[Sigma],(\[Sigma]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-.1,(\[Sigma]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])+.1},{\[Nu],(\[Nu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-1,(\[Nu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])+1},{\[Alpha],(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])-1,(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+1},{\[Beta],1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))-1,1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))+1}},{{p,0.83-.1,0.83+.1}}]],None]];

quantilesMixture:=Quiet[Table[x/.FindRoot[(CDF[nn,x]/.optimalMixture[[2]])==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]];

RMSE1:=Row[{" RMSE1 = ",Chop[RootMeanSquare[quantilesMixture-sdata],10^-6]}];

a=PDF[G,div];b=PDF[nn,div]/.optimalMixture[[2]];RMSE=Sqrt[Total[(a-b)^2]/1000];

MDAPE=Median[(100*Abs[(a-b)])/(a)];sMDAPE=Median[(200*Abs[(a-b)])/(a+b)];

MASE=((1/1000)*Total[Abs[a-b]])/((1/999)*Total[Abs[Differences[a]]]);

rr:={optimalMixture,RMSE1,RMSE,MDAPE,sMDAPE,MASE};rr]

 # Run the above function for different value of weight (P) to find out optimal solution and accuracy measures

uu={result[data,0.10],result[data,0.20],result[data,0.30],result[data,0.40],result[data,0.50],result[data,0.60],result[data,0.70],result[data,0.80],result[data,0.90]}

#Mathematica Code to find simulated Standard Error for parameter estimate of TAL_2 table 5.5

par={\[Mu]->0.01877842826558326`,\[Sigma]->0.041974660396799576`,\[Nu]->97.63400030538081`,\[Alpha]->39.82340185945522`,\[Beta]->45.42001387694496`,p->0.8607655885843857};

bootstrap=Table[prob=RandomVariate[BinomialDistribution[1, 0.8607655885843857],Length[data]]

data1=(1-prob)(RandomVariate[ExponentialDistribution[25.405363621223035],Length[data]]-RandomVariate[ExponentialDistribution[27.092359793165315],Length[data]])+prob*(RandomVariate[StudentTDistribution[0.01877842826558326,0.041974660396799576,97.63400030538081],Length[data]]);

{\[Mu],\[Sigma],\[Nu],\[Alpha],\[Beta],p}/.Last[Quiet[Check[NMaximize[{Total[Log[PDF[nn,data1]]],Join[{\[Sigma]>0,\[Nu]>1},{\[Alpha]>0,\[Beta]>0},{0<p<1}]},Join[{{\[Mu],(\[Mu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-.1,(\[Mu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])+.1},{\[Sigma],(\[Sigma]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-.1,(\[Sigma]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEsti

```
mator ->
"MethodOfMoments"])+.1},{\[Nu],(\[Nu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])-1,(\[Nu]/.FindDistributionParameters[data1,StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ParameterEstimator -> "MethodOfMoments"])+1},{\[Alpha],(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])-1,(Moment[data1,1]-Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+1},{\[Beta],1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))-1,1/(2 Moment[data1,1]^2-Moment[data1,2]) (-2 Moment[data1,1]+(2 Moment[data1,1]^3)/(2 Moment[data1,1]^2-Moment[data1,2])-(Moment[data1,1] Moment[data1,2])/(2 Moment[data1,1]^2-Moment[data1,2])-(2Moment[data1,1]^2 Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2])+(Moment[data1,2] Sqrt[-3 Moment[data1,1]^2+2 Moment[data1,2]])/(2 Moment[data1,1]^2-Moment[data1,2]))+1}},{{p,0.83-.1,0.83+.1}}]],{ {\[Mu]-> 0,\[Sigma]-> 0,\[Nu]->0,\[Alpha]-> 0,\[Beta]-> 0,p-> 0}}]]],{2000}];
```

```
StandardDeviation[DeleteCases[bootstrap,{0,0,0,0,0,0}]]
```

 # Mathematica code to find the goodness of fit test for TAL_2 distribution table 5.6

```
bb=nn/.{\[Mu]->0.01877842826558326`,\[Sigma]->0.041974660396799576`,\[Nu]->97.63400030538081`,\[Alpha]->39.82340185945522`,\[Beta]->45.42001387694496`,p->0.8607655885843857};
```

```
\[ScriptCapitalH]=DistributionFitTest[data,bb,"HypothesisTestData"];\[ScriptCapitalH]["TestDataTable",All]
```

 # Mathematica code for figure 5.1

```
nn=MixtureDistribution[{p,(1-p)},{NormalDistribution[μ1,σ1],NormalDistribution[μ2,σ2]}];
f11=Piecewise[{{Exp[(x-μ)/ψ]/(2ψ), x≤μ}}, Exp[(μ-x)/ϕ]/(2ϕ)];
ALL=ProbabilityDistribution[f11,{x,-∞,∞},Assumptions→{ϕ>0,ψ>0,μ∈Reals}];
NAL1=MixtureDistribution[{p,1-p},{NormalDistribution[μ,σ],ALL}];
f1=Piecewise[{{αβ Exp[β x]/(α+β), x≤0}},αβ Exp[-α x]/(α+β)];
f2=ProbabilityDistribution[f1,{x,-Infinity,Infinity},Assumptions→{ α>0,β>0}];
NAL2=MixtureDistribution[{p,1-p},{NormalDistribution[μ,σ],f2}];
f3=Piecewise[{{Exp[(x-μ)/ψ]/(2ψ), x≤μ}}, Exp[(μ-x)/ϕ]/(2ϕ)];
ALL1=ProbabilityDistribution[f3,{x,-∞,∞},Assumptions→{ϕ>0,ψ>0,μ∈Reals}];
TAL1=MixtureDistribution[{p,1-p},{StudentTDistribution[μ,σ,ν],ALL1}];
f4=Piecewise[{{αβ Exp[β x]/(α+β), x≤0}},αβ Exp[-α x]/(α+β)];
f5=ProbabilityDistribution[f4,{x,-Infinity,Infinity},Assumptions→{ α>0,β>0}];
TAL2=MixtureDistribution[{p,1-p},{StudentTDistribution[μ,σ,ν],f5}];
uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];
TableView[uuss];data=uuss[[2;;263,{1}]]//Flatten;G=SmoothKernelDistribution[data];
div=FindDivisions[{Min[data],Max[data]},1000];{n=Length[data],sdata=Sort[data]};
g2=Show[Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle→{Thickness[0.007],Dotted,Black},PlotRange→All,Frame→True,Axes→False],PDFplot=Plot[PDF[nn,x]/.{σ1→0.0394500521564434`,σ2→0.05501044895827489`,μ1→0.018406036847296987`,μ2→0.0019546485083310966`,p→0.8812438502042144`},{x,Min[data],Max[data]},PlotStyle→{Thin,Black},PlotRange→All],PlotLabel→Text[Style["NN Mixture","Label",Small]]];
g22=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[nn,x]/.{σ1→0.0394500521564434`,σ2→0.05501044895827489`,μ1→0.018406036847296987`,μ2→0.0019546485083310966`,p→0.8812438502042144`})□i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}←,PlotRange→All,ImageSize→{290,175},PlotStyle→{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle→Black],Frame→True,PlotLabel→Text[Style["NN Mixture","Label",Small]]];
```

```
g3=Show[Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle→{Thickness[0.007],Dotted,Black},PlotRange
→All,Frame→True,Axes→False],PDFplot=Plot[PDF[NAL1,x]/.{μ→
0.0140432,σ→0.04308871498517691`,ϕ→0.029432059190111725`,ψ→0.011195661894971605`,p→0.87
61693288029527`},{x,Min[data],Max[data]},PlotStyle→{Thin,Black},PlotRange→All],PlotLabel→Text[
Style["NAL1 Mixture","Label",Small]]];

g33=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[NAL1,x]/.{μ→
0.0140432,σ→0.04308871498517691`,ϕ→0.029432059190111725`,ψ→0.011195661894971605`,p→0.87
61693288029527`})□i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}←,PlotRange→All,ImageSize→{290,175},PlotStyle→{Black,PointSi
ze[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle→Black],Frame→True,PlotLabel→Text[Style["NA
L1 Mixture","Label",Small]]];
g4=Show[Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle→{Thickness[0.007],Dotted,Black},PlotRange
→All,Frame→True,Axes→False],PDFplot=Plot[PDF[NAL2,x]/.{σ→0.040239444476853875`,μ→0.0225
1649574543454`,α→48.172434462297986`,β→32.35877311627851`,p→0.8092099118475501`},{x,Min[
data],Max[data]},PlotStyle→{Thin,Black},PlotRange→All],PlotLabel→Text[Style["NAL2
Mixture","Label",Small]]];
g44=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[NAL2,x]/.{σ→0.040239444476853875`,μ→0.0
2251649574543454`,α→48.172434462297986`,β→32.35877311627851`,p→0.8092099118475501`})□i/(n
+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}←,PlotRange→All,ImageSize→{290,175},PlotStyle→{Black,PointSi
ze[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle→Black],Frame→True,PlotLabel→Text[Style["NA
L2 Mixture","Label",Small]]];

g5=Show[Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle→{Thickness[0.007],Dotted,Black},PlotRange
→All,Frame→True,Axes→False],PDFplot=Plot[PDF[TAL1,x]/.{μ→0.0140432,σ→0.0429893987285939
96`,ν→561.6095472598373`,ϕ→0.029497763511069332`,ψ→0.011141022213614358`,p→0.8776600352
724089`},{x,Min[data],Max[data]},PlotStyle→{Thin,Black},PlotRange→All],PlotLabel→Text[Style["TA
L1 Mixture","Label",Small]]];
g55=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[TAL1,x]/.{μ→0.0140432,σ→0.0429893987285
93996`,ν→561.6095472598373`,ϕ→0.029497763511069332`,ψ→0.011141022213614358`,p→0.8776600
352724089`})□i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}←,PlotRange→All,ImageSize→{290,175},PlotStyle→{Black,PointSi
ze[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle→Black],Frame→True,PlotLabel→Text[Style["TAL
1 Mixture","Label",Small]]];

g6=Show[Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle→{Thickness[0.007],Dotted,Black},PlotRange
→All,Frame→True,Axes→False],PDFplot=Plot[PDF[TAL2,x]/.{μ→0.022350521324623212`,σ→0.03994
1828989823784`,ν→94.825641579638`,α→49.3910794403953`,β→32.5634458698108`,p→0.815974038
830743`},{x,Min[data],Max[data]},PlotStyle→{Thin,Black},PlotRange→All,Frame→True,Axes→False],
PlotLabel→Text[Style["TAL2 Mixture","Label",Small]]];
g66=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[TAL2,x]/.{μ→0.022350521324623212`,σ→0.0
39941828989823784`,ν→94.825641579638`,α→49.3910794403953`,β→32.5634458698108`,p→0.81597
4038830743`})□i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}←,PlotRange→All,ImageSize→{290,175},PlotStyle→{Black,PointSi
ze[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle→Black],Frame→True,PlotLabel→Text[Style["TAL
2 Mixture","Label",Small]]];

Show[GraphicsGrid[{{g2,g22},{g3,g33},{g4,g44},{g5,g55},{g6,g66}}]]
```

#Mathematica Code for figure 5.2

```
nn=MixtureDistribution[{p,(1-p)},{NormalDistribution[μ1,σ1],NormalDistribution[μ2,σ2]}];
f11=Piecewise[{{Exp[(x-μ)/ψ]/(2ψ), x≤μ}}, Exp[(μ-x)/ϕ]/(2ϕ)];
```

```
ALL=ProbabilityDistribution[f11,{x,-∞,∞},Assumptions→{ϕ>0,ψ>0,μ∈Reals}];
NAL1=MixtureDistribution[{p,1-p},{NormalDistribution[μ,σ],ALL}];
f1=Piecewise[{{αβ Exp[β x]/(α+β), x≤0}},αβ Exp[-α x]/(α+β)];
f2=ProbabilityDistribution[f1,{x,-Infinity,Infinity},Assumptions→{ α>0,β>0}];
NAL2=MixtureDistribution[{p,1-p},{NormalDistribution[μ,σ],f2}];
f3=Piecewise[{{Exp[(x-μ)/ψ]/(2ψ), x≤μ}}, Exp[(μ-x)/ϕ]/(2ϕ)];
ALL1=ProbabilityDistribution[f3,{x,-∞,∞},Assumptions→{ϕ>0,ψ>0,μ∈Reals}];
TAL1=MixtureDistribution[{p,1-p},{StudentTDistribution[μ,σ,ν],ALL1}];
f4=Piecewise[{{αβ Exp[β x]/(α+β), x≤0}},αβ Exp[-α x]/(α+β)];
f5=ProbabilityDistribution[f4,{x,-Infinity,Infinity},Assumptions→{ α>0,β>0}];
TAL2=MixtureDistribution[{p,1-p},{StudentTDistribution[μ,σ,ν],f5}];
uuss:=Import["GDPZ.xlsx",{"xlsx","Data",1}];
TableView[uuss];
data=uuss[[2;;263,{1}]]//Flatten;

g1=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["Normal distribution","Label",Small]]],PDFplot=Plot[PDF[NormalDistribution[μ,σ],x]/.{μ→0.01629,σ→0.04158},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
g2=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["NN Mixture","Label",Small]]],PDFplot=Plot[PDF[nn,x]/.{σ1→0.0394500521564434`,σ2→0.05501044895827489`,μ1→0.018406036847296987`,μ2→0.0019546485083310966`,p→0.8812438502042144`},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
g3=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["NAL-1 Mixture","Label",Small]]],PDFplot=Plot[PDF[NAL1,x]/.{μ→ 0.0140432,σ→0.04308871498517691`,ϕ→0.029432059190111725`,ψ→0.011195661894971605`,p→0.8761693288029527`},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
g4=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["NAL-2 Mixture","Label",Small]]],PDFplot=Plot[PDF[NAL2,x]/.{σ→0.040239444476853875`,μ→0.022516495745434 54`,α→48.172434462297986`,β→32.35877311627851`,p→0.8092099118475501`},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
g5=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["TAL-1 Mixture","Label",Small]]],PDFplot=Plot[PDF[TAL1,x]/.{μ→0.0140432,σ→0.042989398728593996`,ν→ 561.6095472598373`,ϕ→0.029497763511069332`,ψ→0.011141022213614358`,p→0.8776600352724089`},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
g6=Show[Histogram[data,13,"ProbabilityDensity",PlotRange→All,ImageSize→{290,175},ImagePadding→20,ChartStyle→White,PlotLabel→Text[Style["TAL-2 Mixture","Label",Small]]],PDFplot=Plot[PDF[TAL2,x]/.{μ→0.022350521324623212`,σ→0.039941828989823784`,ν→94.825641579638`,α→49.3910794403953`,β→32.5634458698108`,p→0.815974038830743`},{x,Min[data],Max[data]},PlotStyle→{Thick,Black},PlotRange→All]];
Show[GraphicsGrid[{{g1,g2},{g3,g4},{g5,g6}}]]
```

#Mathematica Code for figure 5.3

```
nn=MixtureDistribution[{p,(1-p)},{NormalDistribution[Subscript[\[Mu], 1],Subscript[\[Sigma], 1]],NormalDistribution[Subscript[\[Mu], 2],Subscript[\[Sigma], 2]]}];

f11=Piecewise[{{Exp[(x-\[Mu])/\[Psi]]/(2\[Psi]), x<=\[Mu]}}, Exp[(\[Mu]-x)/\[Phi]]/(2\[Phi])];

ALL=ProbabilityDistribution[f11,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

NAL1=MixtureDistribution[{p,1-p},{NormalDistribution[\[Mu],\[Sigma]],ALL}];
```

```
f1=Piecewise[{{\[Alpha] \[Beta] Exp[\[Beta] x]/(\[Alpha]+\[Beta]), x<=0}},\[Alpha] \[Beta] Exp[-\[Alpha] x]/(\[Alpha]+\[Beta])];

f2=ProbabilityDistribution[f1,{x,-Infinity,Infinity},Assumptions->{ \[Alpha]>0,\[Beta]>0}];

NAL2=MixtureDistribution[{p,1-p},{NormalDistribution[\[Mu],\[Sigma]],f2}];

f3=Piecewise[{{Exp[(x-\[Mu])/\[Psi]]/(2\[Psi]), x<=\[Mu]}}, Exp[(\[Mu]-x)/\[Phi]]/(2\[Phi])];

ALL1=ProbabilityDistribution[f3,{x,-\[Infinity],\[Infinity]},Assumptions->{\[Phi]>0,\[Psi]>0,\[Mu]\[Element]Reals}];

TAL1=MixtureDistribution[{p,1-p},{StudentTDistribution[\[Mu],\[Sigma],\[Nu]],ALL1}];

f4=Piecewise[{{\[Alpha] \[Beta] Exp[\[Beta] x]/(\[Alpha]+\[Beta]), x<=0}},\[Alpha] \[Beta] Exp[-\[Alpha] x]/(\[Alpha]+\[Beta])];

f5=ProbabilityDistribution[f4,{x,-Infinity,Infinity},Assumptions->{ \[Alpha]>0,\[Beta]>0}];

TAL2=MixtureDistribution[{p,1-p},{StudentTDistribution[\[Mu],\[Sigma],\[Nu]],f5}];

caca=Import["GDPZ.xlsx",{"xlsx","Data",3}];TableView[caca];data=caca[[2;;206,{1}]]//Flatten;

G=SmoothKernelDistribution[data];div=FindDivisions[{Min[data],Max[data]},1000];

{n=Length[data],sdata=Sort[data]};

g2=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame->True],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle->{Thickness[0.007],Dotted,Black},PlotRange->All,Frame->True,Axes->False],PDFplot=Plot[PDF[nn,x]/.{Subscript[\[Mu], 1]->0.032845323825797454,Subscript[\[Sigma], 1]->0.0385767978218417,Subscript[\[Mu], 2]->-0.04546372285360516,Subscript[\[Sigma], 2]->0.031108177166300228`,p->0.8143399989936488`},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],PlotLabel->Text[Style["NN Mixture","Label",Small]]];

g22=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[nn,x]/.{Subscript[\[Mu], 1]->0.032845323825797454,Subscript[\[Sigma], 1]->0.0385767978218417,Subscript[\[Mu], 2]->-0.04546372285360516,Subscript[\[Sigma], 2]->0.031108177166300228`,p->0.8143399989936488`})==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}\[Transpose],PlotRange->All,ImageSize->{290,175},PlotStyle->{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle->Black],Frame->True,PlotLabel->Text[Style["NN Mixture","Label",Small]]];

g3=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame->True],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle->{Thickness[0.007],Dotted,Black},PlotRange->All,Frame->True,Axes->False],PDFplot=Plot[PDF[NAL1,x]/.{\[Mu]->0.02007556539610321,\[Sigma]->0.049265688449921664`,\[Phi]->0.026658123656732025`,\[Psi]->0.03851228130854295`,p->0.9418599382767202`},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],PlotLabel->Text[Style["NAL1 Mixture","Label",Small]]];

g33=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[NAL1,x]/.{\[Mu]->0.02007556539610321,\[Sigma]->0.049265688449921664`,\[Phi]->0.026658123656732025`,\[Psi]->0.03851228130854295`,p->0.9418599382767202`})==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}\[Transpose],PlotRange->All,ImageSize->{290,175},PlotStyle->{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle->Black],Frame->True,PlotLabel->Text[Style["NAL1 Mixture","Label",Small]]];

g4=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding->20,ChartStyle->White,Frame->True],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle->{Thickness[0.007],Dotted,Black},PlotRange->All,Frame->True,Axes-
```

```
>False],PDFplot=Plot[PDF[NAL2,x]/.{\[Sigma]->0.02039432051457115`,\[Mu]-
>0.05146400839968819`,\[Alpha]->25.405363621223035`,\[Beta]->27.092359793165315`,p-
>0.32349517456284693`},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],PlotLabel-
>Text[Style["NAL2 Mixture","Label",Small]]];

g44=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[NAL2,x]/.{\[Sigma]-
>0.02039432051457115`,\[Mu]->0.05146400839968819`,\[Alpha]->25.405363621223035`,\[Beta]-
>27.092359793165315`,p->0.32349517456284693`})==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}\[Transpose],PlotRange->All,ImageSize->{290,175},PlotStyle-
>{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle->Black],Frame->True,PlotLabel-
>Text[Style["NAL2 Mixture","Label",Small]]];

g5=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding-
>20,ChartStyle->White,Frame->True],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle-
>{Thickness[0.007],Dotted,Black},PlotRange->All,Frame->True,Axes-
>False],PDFplot=Plot[PDF[TAL1,x]/.{\[Mu]->0.0202188,\[Sigma]->0.049139994085544474`,\[Nu]-
>237.7750287042269`,\[Phi]->0.02798486569279492`,\[Psi]->0.0329275845911434`,p-
>0.9375217977681525`},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange->All],PlotLabel-
>Text[Style["TAL1 Mixture","Label",Small]]];

g55=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[TAL1,x]/.{\[Mu]->0.0202188,\[Sigma]-
>0.049139994085544474`,\[Nu]->237.7750287042269`,\[Phi]->0.02798486569279492`,\[Psi]-
>0.0329275845911434`,p->0.9375217977681525`})==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}\[Transpose],PlotRange->All,ImageSize->{290,175},PlotStyle-
>{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle->Black],Frame->True,PlotLabel-
>Text[Style["TAL1 Mixture","Label",Small]]];

g6=Show[Histogram[data,13,"ProbabilityDensity",PlotRange->All,ImageSize->{290,175},ImagePadding-
>20,ChartStyle->White,Frame->True],Plot[PDF[G,y],{y,Min[data],Max[data]},PlotStyle-
>{Thickness[0.007],Dotted,Black},PlotRange->All,Frame->True,Axes-
>False],PDFplot=Plot[PDF[TAL2,x]/.{\[Sigma]->0.04814739300894285`,\[Nu]-
>8841.479758018255`,\[Mu]->0.018307665488183626`,\[Alpha]->30.722429536875456`,\[Beta]-
>530.1198669661604`,p->0.974594556},{x,Min[data],Max[data]},PlotStyle->{Thin,Black},PlotRange-
>All,Frame->True,Axes->False],PlotLabel->Text[Style["TAL2 Mixture","Label",Small]]];

g66=Show[ListPlot[{sdata,Quiet[Table[x/.FindRoot[(CDF[TAL2,x]/.{\[Sigma]-
>0.04814739300894285`,\[Nu]->8841.479758018255`,\[Mu]->0.018307665488183626`,\[Alpha]-
>30.722429536875456`,\[Beta]->530.1198669661604`,p->0.974594556})==i/(n+1.),

{x,sdata[[i]],sdata[[i]]+.05}],{i,n}]]}\[Transpose],PlotRange->All,ImageSize->{290,175},PlotStyle-
>{Black,PointSize[0.01]}],Plot[x,{x,First[sdata],Last[sdata]},PlotStyle->Black],Frame->True,PlotLabel-
>Text[Style["TAL2 Mixture","Label",Small]]];

Show[GraphicsGrid[{{g2,g22},{g3,g33},{g4,g44},{g5,g55},{g6,g66}}]]
```