MONASH UNIVERSITY

AUSTRALIA

NON-LINEAR TIME SERIES MODELLING AND DISTRIBUTIONAL FLEXIBILITY

Jenny N. Lye and Vance L. Martin

Working Paper No. 14/91

December 1991

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 982 X

NON-LINEAR TIME SERIES MODELLING AND DISTRIBUTIONAL FLEXIBILITY

Jenny N. Lye and Vance L. Martin

Working Paper No. 14/91

December 1991

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

NON-LINEAR TIME SERIES MODELLING AND DISTRIBUTIONAL FLEXIBILITY

by

Jenny N. Lye and Vance L. Martin

University of Melbourne, Parkville, Victoria, 3052

Abstract

Most of the existing work in non-linear time series analysis has concentrated on generating flexible functional models by specifying non-linear specifications for the mean of a particular process without much, if any, attention given to the distributional properties of the model. However, as Martin (1991) has shown, greater flexibility in perhaps a more natural way, can be achieved by consideration of distributions from the generalized exponential class. This paper represents an extension of the earlier work of Martin by introducing a flexible class of non-linear time series models which can capture a wide range of empirical behaviour such as skewed, fat-tailed and even multimodal distributions. This class of models is referred to as GENTS: Generalized Exponential Non-linear Time Series. A maximum likelihood algorithm is given for estimating the parameters of the model, and the framework is applied to estimating the distribution of the movements of the exchange rate.

KEY WORDS: Non-linear time series, generalized exponential distributions, skewness, fat-tails, multimodality, maximum likelihood

1 INTRODUCTION

The identification of limit cycles, chaos, self-excitation, asymmetric distributions, leptokurtosis, sudden jumping behaviour, time deformation,and time series exhibiting irreversibility characteristics, has led to the development of nonlinear time series models. One class of nonlinear time series models consists of letting the pertinent variable at time t be either a function of polynomials in the autoregressive terms (Granger and Newbold, 1977), or a function of polynomials in the moving average terms (Keenan, 1985), or as a product of autoregressive and moving average terms (Granger and Anderson, 1978; and Rao and Gabr, 1984). Another class consists of allowing the dependent variable to switch between linear autoregressive models (Tong, 1983).

One important implication of nonlinear time series analysis is that in modelling data it is necessary to consider not only the first two moments of the distribution, but higher order moments such as the third (skewness) and the fourth (kurtosis) moments. However, in most nonlinear time series models higher order moments are not modelled explicitly. For example, in the class of models suggested by Engle (1982) called ARCH, kurtosis can be modelled by allowing the (conditional) variance to have an autoregressive representation. This class of models can also capture the bunching characteristic of price data that was first observed by Mandelbrot (1963), and has been extended by Nelson (1991) to explain skewness. An exception is the work of Rao and Gabr (1984) and Ashley and Patterson (1989) which is based on bispectral analysis and which consists of a double Fourier transformation of the third moment.

The reason for the implicit treatment of higher order moments can be attributed to a preoccupation with constructing nonlinear models of the mean and the variance, with little, if any consideration given to the specification of the stochastics of the model. The approach adopted in this paper, and in contrast with most of the existing nonlinear time series models, is to construct a nonlinear time series model which explicitly takes into account higher order moments. The approach consists of using a generalization of the Student t distribution to model the error distribution. This distribution which has been studied recently by Martin (1990) and Lye and Martin (1990), provides a flexible framework by which to build nonlinear time series models since most of the observed nonnormal characteristics often identified in data can be captured directly. A special feature of the generalized Student t distribution is that it can model not only symmetric fat-tailed distributions, but also distributions that are skewed, and possibly even multimodal. This last characteristic is particularly important when attempting to model the switching regime class of Tong (1983) and Martin (1991).

The rest of the paper proceeds as follows. The theoretical apparatus of the GENTS model is given in Section 2. An iterative maximum likelihood estimation procedure is given in Section 3, while the GENTS model is contrasted with some existing time series models in Section 4. In Section 5 the GENTS class of nonlinear time series models is applied to modelling the rate of growth of the nominal exchange rate. An important feature of the estimated model is that the exchange rate distribution is bimodal over the period 1977 to 1989. This not only suggests that movements in the exchange rate are the result of the exchange rate jumping between equilibria in a zone of multiple equilibria, but it also highlights the difficulties of modelling the rate of growth of the exchange rate with a linear model. Some concluding comments and suggestions for future work are given in Section 6.

3 GENERALIZED EXPONENTIAL NON-LINEAR TIME SERIES (GENTS)

3.1 The Model

Consider the following non-linear autoregressive time series model

$$Y_t = \theta_0(Y_{t-1}, Y_{t-2}, \ldots) + U_t, \tag{1}$$

where $\theta_0(.)$ represents a non-linear function, and U_t is an error term with zero mean and variance equal to σ_u^2. The distribution of U_t and hence Y_t, is often assumed to be normal. However, as discussed in the previous section, there are a number of reasons as to why the distribution of Y_t is not normal. One appropriate class of distributions to be used as a model of Y_t stems from the work of Cobb (1978), Cobb, Koppstein and Chen (1983), Martin (1991), and Lye and Martin (1990). This class of distributions is based on a generalization of the Pearson system. In particular, one subordinate in this class which is able to model the leptokurtosis observed in a number of data sets, represents a generalization of the Student t distribution. The generalized Student t distribution is derived from the following differential equation

$$\frac{df}{du} = \frac{-g(u)f(u)}{h(u)}, \tag{2}$$

where

$$g(u) = \sum_{i=0}^{M-1} \alpha_i u^i, \tag{3}$$

$$h(u) = \gamma^2 + u^2, \qquad -\infty < u < \infty, \tag{4}$$

where f is the density function of U and γ^2 is referred to as the "degrees of freedom" parameter. In the standard Pearson system, g(u) is a polynomial in U of degree less than or equal to one, whereas h(u) is a polynomial in U of degree less than or equal to two. The general solution of (2) for M=6 is

$$f(u) = \exp\left[\theta_1 \tan^{-1}(u/\gamma) + \theta_2 \log(\gamma^2 + u^2) + \sum_{i=3}^{6} \theta_i u^{i-2} - \eta\right], \quad -\infty<u<\infty, \tag{5}$$

where the normalizing constant is given by

$$\eta = \log\int \exp\left[\theta_1 \tan^{-1}(u/\gamma) + \theta_2 \log(\gamma^2 + u^2) + \sum_{i=3}^{6} \theta_i u^{i-2}\right] du, \tag{6}$$

and

$$\begin{aligned}
\theta_1 &= -\alpha_0 + \gamma\alpha_2 - \gamma^2\alpha_4,\\
\theta_2 &= -\alpha_1/2 + \gamma^2\alpha_3/2 - \gamma^4\alpha_5/2,\\
\theta_3 &= -\alpha_2 + \gamma^2\alpha_4,\\
\theta_4 &= -\alpha_3/2 + \gamma^2\alpha_5/2,\\
\theta_5 &= -\alpha_4/3,\\
\theta_6 &= -\alpha_5/4.
\end{aligned}$$

The conditional distribution of Y_t can be simply obtained from using (1) to transform (5) into a density in terms of Y_t. This yields a generalized conditional Student t distribution which encompasses a number of well-known distributions. For example, the normal distribution is given by setting $\theta_1=\theta_2=0$ and $\theta_i=0$, i>4. The Student t distribution is given by $\theta_1=0$, $\theta_i=0$, i>2, and $\theta_2=-(1+\gamma^2)/2$. The Pearson Type 1V distribution is given by $\theta_i=0$, i>4. The generalized normal distribution is given by $\theta_1=\theta_2=0$.

The distribution given by (5) is flexible enough to model distributions that are fat-tailed, skewed, or even multimodal. In general, the functions θ_i, i=1,6, in (5) can be time varying by assuming that these parameters have autoregressive representations which need not be linear

$$\theta_{it} = \theta_{it}(Y_{t-1}, Y_{t-2}, \ldots) \qquad i=1,2,\ldots,6. \tag{7}$$

An implication of (7) is that there is both a direct and an indirect relationship between Y_t and lagged Y_t. The direct relationship is given by (1), whereas the indirect relationship is via the distributional parameters θ_{it}, i=1,2,..., which affect U_t and hence Y_t. An allowance for both direct and indirect linkages makes the relationship between Y_t and lagged Y_t non-linear. Of course the model is even more non-linear if θ_{it}, i=0,1,2,..., are also non-linear.

3.2 Choice of Predictor

In the standard case, the distribution of U_t is unimodal and hence g(.) in (3) has a single (real) root. The choice of the best predictor of the random variable, or some function of it, depends upon the adopted loss function. Whatever the choice of the loss function is, the set of predictors is small in general. If the distribution is both unimodal and symmetric, this set is reduced even further. A typical choice is the mean of the distribution.

Difficulties arise when the distribution is bimodal, or even multimodal. In the multimodal case, the usual conventions such as the mean for example, would seem to be inappropriate. Some potential new conventions are (see, for example, Cobb, Koppstein and Chen, 1983):

Global: Choose the global maximum at each point in time.

Delay: Do not leave a mode until that mode disappears.

Nearest: Choose the mode or antimode closest to the data.

The "global" convention, is based on the assumption that the process always moves (jumps) to a region of higher probability. With the "delay" convention, it is assumed that a process never leaves a stable region (ie delays moving) until that region becomes unstable. The "nearest" convention consists of choosing the mode, stable of unstable, which is closest to the data. This convention allows for the possibility that the process can settle at unstable equilibria (known as self-organized criticality) as well as settle at stable equilibria. This convention can also be adjusted to allow for delays by choosing the mode at time t which is closest to the data at time $t-\tau$.

3.3 Interpreting Parameters

Consider the following generalized normal distribution which is a special case of the generalized Student t distribution

$$f(u) = \exp\left[\theta_3 u + \theta_4 u^2 - u^4/4 - \eta\right], \quad -\infty<u<\infty. \tag{8}$$

The bimodal region can be determined from the sign of Cardan's discriminant

$$\delta = (\theta_1/2)^2 - (\theta_2/3)^3, \tag{9}$$

Unimodality: $\delta > 0$,

Bimodality: $\delta < 0$.

Also note that a necessary condition for bimodality is that $\theta_2 > 0$.

It can be shown that the parameters θ_3 and θ_4 in (8) have the following interpretations (see Cobb, Koppstein and Chen, 1983)

Unimodality: θ_1 is a measure of skewness,

θ_2 is a measure of kurtosis.

Bimodality: θ_1 is a measure of the relative heights of the two modes,

θ_2 is a measure of the separateness of the two modes.

In the case of the generalized Student t distribution given by (5), some of the properties of this distribution are highlighted in Figure 1.

4 MAXIMUM LIKELIHOOD ESTIMATION

The principle of maximum likelihood provides a means of choosing a best asymptotically normal, BAN, estimator which is identified as the global maximum of the log - likelihood function. Since the model is non-linear, it is necessary to compute the parameter estimates by an iterative optimization procedure. Letting θ be a vector of K parameters, then the generalized Student t density can be written as

$$f(u) = \exp\left[\sum_i \theta_i R_i(u;\theta) - \eta\right], \qquad (10)$$

where $R(u;\theta)$ is a general function of u which is dependent on the set of parameters θ, and the normalizing constant is defined as

$$\eta = \log\int \exp\left[\sum_i \theta_i R_i(u;\theta)\right] du. \qquad (11)$$

For a sample of size t=1,2,...,T, the log of the likelihood is

$$\mathcal{L}og\ \mathcal{L} = \sum_t \log f(u_t) = \sum_t \sum_i \theta_i R(u_t;\theta) - \sum_t \eta_t. \qquad (12)$$

Numerical iterative optimization procedures are required to maximize the log - likelihood function. In this paper both the Newton-Raphson (NR) and Berndt, Hall, Hall, Hausman (BHHH) algorithms are used. The BHHH algorithm is a Quasi - Newton method which approximates the Hessian at each iteration by the cross - product of the first derivatives of the log - likelihood. For computational convenience all derivatives are computed numerically.

A property of the log - likelihood in this case is that it may be multimodal, (see, for example, Gabrielson, 1982). To guard against this problem, it is necessary to choose several sets of starting values to ensure that the global maximum is achieved.

At each iteration numerical integrations are required to compute the normalizing constant in (10) since there does not exist closed - form solutions for these integrations. The authors have found that Gaussian-Legendre quadrature based on the GAUSS program INTQUAD1, yields more than satisfactory results.

5 RELATIONSHIP WITH EXISTING MODELS

5.1 SETAR

Tong (1983) has introduced a class of models known as self-exciting threshold autoregressive (SETAR). This class of models represents a set of piece-wise linear time series models where at each point in time, the model used to predict the dependent variable is based on

$$Y_t = \begin{cases} \sum_i \alpha_i Y_{t-i} , & Y_{t-\tau} \le \delta \\ \sum_i \beta_i Y_{t-i} , & Y_{t-\tau} > \delta \end{cases} \tag{13}$$

where τ is the delay parameter and δ is the threshold parameter. The GENTS model captures the essential characteristics of the SETAR model in the case where the distribution in (5) is bimodal. In this case, there are two predictors and the choice of the predictor is based on the adopted criterion as given in Section 3.2 of the paper.

5.2 ARCH

The original ARCH model of Engle (1982) is based on assuming that both the mean and in particular the variance have linear autoregressive representations. For example, the ARCH(M) model is given by

$$Y_t = \sum_i \alpha_i Y_{t-i} + U_t \tag{14}$$

$$U_t = \sigma_t Z_t, \; Z_t \sim \text{iid}, \; \mu_z = 0, \; \sigma_z^2 = 1,$$

$$\sigma_t^2 = \phi_0 + \sum_{i=1}^{M} \phi_i U_{t-i}^2.$$

In the ARCH(M) model, the conditional variance has a simple nonlinear structure: σ_t^2 is a nonlinear function of U_{t-i}^2. However, some other non-linear conditional variance specifications have been suggested recently by Nelson (1991). The model in this case is called EGARCH and the conditional variance is given by for the EGARCH(1,1) model

$$\sigma_t^2 = \phi_0 + \phi_1(|U_{t-1}/\sigma_{t-1}| - (2/\pi)^{0.5}) + \phi_2 \ln \sigma_{t-1}^2 + \phi_3 U_{t-1}/\sigma_{t-1}$$

In the GENTS model, estimates of the conditional variance can be obtained by computing

$$\sigma_t^2 = \int (Y_t - \mu_t)^2 f(y_t | Y_{t-1}, \ldots), \tag{15}$$

where $f(y_t|Y_{t-1},\ldots)$ is given by the conditional generalized Student t distribution in (5). However, it is natural to consider higher order conditional moments which can be derived from

$$M_t^j = \int (Y_t - \mu_t)^j f(y_t|Y_{t-1},\ldots), \quad j=1,2,3,4,\ldots \tag{16}$$

This expression provides a generalization of the ARCH class of models in two ways. First, the relationship between σ_t^2 and U_{t-1} tends to be more nonlinear than it is in the existing ARCH specifications. The standard ARCH model is realized when $f(.|.)$ is chosen as the conditional normal distribution. Second, the allowance for higher order conditional moments generalizes the existing approach that is adopted in ARCH modelling where only the second conditional moments is concentrated on. This extension is likely to be useful in applied work where a number of researchers have found that the standard ARCH model has not been able to capture all of the observed kurtosis in the data.

5.3 MATS

The multipredictor autoregressive time series model (MATS) suggested by Martin (1991) is a special case of the GENTS model. The MATS model provides a framework for embedding piece wise linear autoregressive models in a multimodal model with the constraint that the predictors from the competing models correspond to the modes of the distribution. That is, the MATS model is obtained by constraining the GENTS model to be multimodal over the entire sample. These constraints amount to a set of nonlinear restrictions which can be tested by using a Wald test (see Martin, 1990). This test provides a way of examining the validity of the nesting framework adopted in the MATS model.

5.4 Testing For Nonnormality

Consider the generalized Student t distribution in (5) with the restriction that $\theta_1=\theta_2=0$.

$$f(u) = \exp\left[\theta_3 u + \theta_4 u^2 + \theta_5 u^3 + \theta_6 u^4 - \eta\right], \quad -\infty<u<\infty. \tag{17}$$

It can be shown that a Lagrange multiplier test of the hypothesis

$$H_0\colon \theta_5 = \theta_6 = 0$$

yields the test statistic

$$T[\mu_3^2/6\mu_2^3 + (\mu_4/\mu_2^2 - 3)^2/24], \tag{18}$$

which is asympotically distributed under H_0 as χ_2^2. Alternatively, consider the following generalized Student t distribution

$$f(u) = \exp\left[\theta_1 \tan^{-1}(u/\gamma) + \theta_2 \log(\gamma^2 + u^2) + \sum_{i=3}^{4} \theta_i u^{i-2} - \eta\right], \quad -\infty<u<\infty. \tag{19}$$

It can be easily shown that a Lagrange multiplier test of the hypothesis

$$\theta_1 = \theta_2 = 0,$$

yields the test statistic (18), by simply noting that the second terms in the Taylor series expansions of $\tan^{-1}(u/\gamma)$ and $\log(\gamma^2 + u^2)$, correspond to u^3 and u^4 respectively.

The test statistic given by (18) is simply the Bera-Jarque (1982) test for normality which was derived by defining (3) and (4) as

$$g(u) = \alpha_0 + \alpha_1 u,$$

$$h(u) = \beta_0 + \beta_1 u + \beta_2 u^2,$$

and testing that the $\beta_1=\beta_2=0$ by use of a Lagrange multiplier procedure. To a certain extent it is not too surprising that the two testing frameworks yield identical test statistics since both procedures are based on a Lagrange multiplier test where the distribution under the null is the same; namely, the unimodal (Pearson) normal distribution. From a practical point view however, an implication of a statistically significant test statistic which results in the null hypothesis being rejected, means that it is not clear what is the alternative distribution. If the null hypothesis of a unimodal normal distribution is rejected, this may simply reflect the presence of a generalized Student t distribution.

5.5 ARMA

The linear ARMA model is a special case of the GENTS model, arising when $\theta_3 \neq 0$, $\theta_4 = -1$, with the remaining parameters equal to zero. In this case, the error distribution has a unimodal normal distribution.

6 APPLICATION TO EXCHANGE RATES

The framework of the generalized exponential nonlinear time series model is now applied to modelling the growth rate of the nominal exchange rate. The distribution of the growth rate of the exchange rate has already been extensively studied. The main conclusion from this work is that the empirical distribution is leptokurtic which has been modelled in a variety of ways. Typical distributions that have been used are: Paretian stable distribution (Mandelbrot, 1963; and Westerfield, 1977); mixture of normals

(Boothe and Glassman, 1987); Student t (Rogalski and Vinso, 1978); normal with time varying parameters (Friedman and Vandersteel, 1982); mixed jump processes (Akgiray and Booth, 1988; and Tucker and Pond, 1988); and ARCH with nonnormal errors (Hsieh, 1989).

A feature of most of the empirical work on the distribution of the rate of growth of the exchange rate is that the distribution is assumed to be time invariant. This assumption can be investigated by using a nonlinear time series model such as the GENTS framework, which generates temporal exchange rate distributions. In particular, it is hypothesised that for those points in time where the movements in the exchange rate are very large, the distribution is not a unimodal-leptokurtic distribution, but is bimodal with the large movements reflecting the exchange rate jumping between equilibria in a zone of multiple equilibria.

6.1 Data

The rate of growth of the exchange rate is defined as

$$e_t = 100 \log(E_t/E_{t-1}),$$

where E_t is the \$US/\$Australian nominal exchange rate. The data is monthly and the sample period begins in 1977:1 and ends in 1989:10, a sample size of 154 observations. The data comes from the dx database.

6.2 The Linear Specification

The results for estimating a linear AR(1) time series model are given by

$$\hat{e}_t = \underset{(-0.848)}{-0.207} + \underset{(0.404)}{0.032}\, e_{t-1} \qquad (20)$$

R^2(LINEAR) = 0.001, AR(1) = 2.082, ARCH(1) = 0.151, HETERO = 0.440,

Sk = -6.626, Kt = 11.693, KEENAN = 0.058, RESET = 0.159.

where t-statistics are given in brackets, and the tests for first order autocorrelation with a lagged dependent variable [AR(1)], first order ARCH [ARCH(1)] and Breusch-Pagan heteroskedasticity [HETERO] are distributed as χ_1^2, whereas the test for skewness [Sk] and for kurtosis [Kt] are distributed asymptotically as N(0,1). Two general tests of nonlinearity are also reported; namely, the Keenan test [KEENAN] (see Keenan, 1985) and the Ramsey test [RESET] (see Ramsey, 1969), which are distributed as $F_{1,T-6}$, and N(0,1) respectively.

Clearly, the AR(1) model performs very badly with R^2(LINEAR) = 0.001. The reason(s) as to why the model is performing badly is (are) not detected by the usual diagnostics based on tests of autocorrelation, ARCH and heteroskedasticity. There is also no evidence of any nonlinearities as based on the Keenan and Ramsey tests. There is, however, significant negative skewness and kurtosis, suggesting that the assumption that the residuals are normal should be rejected. From the discussion in Section 5.4, the rejection of the normality assumption could reflect that the pertinent distribution is not normal, but say Generalized Student t. In particular, evidence against normality may also reflect that the distribution is multimodal.

6.3 The GENTS Specification

Given that there is strong evidence of both skewness and kurtosis in the residuals, it is appropriate to fit a GENTS model based on the Generalized Student t distribution. A range of models were estimated with the best model given by

$$\text{GEM-T} = \exp\left[\theta_{1t}\tan^{-1}(e_t/\gamma) + \theta_{2t}\log(\gamma^2 + e_t^2) + \sum_{i=3}^{4}\theta_{it}e_t^{i-2} - \eta_t\right], \qquad (21)$$

where $\gamma = 2.577$,

$$\theta_{1t} = 24.386 - 0.754e_{t-1},$$

$$\theta_{2t} = -(1 + \gamma^2)/2,$$

$$\theta_{3t} = -8.350 + 0.162e_{t-1},$$

$$\theta_{4t} = -1/2.$$

The distributional properties of the estimated model are highlighted in Figure 2. The main observation to note is that the exchange rate distribution is bimodal over the entire sample period. Figure 2.1 gives a time series plot of the data together with the two modes, which occur at around 0 and -6, and the antimode which occurs at around -2, of the distribution. The bimodality property of the distribution is further highlighted in Figure 2.3, where the sequence of exchange rate distributions over time is given. Notice that the global mode occurs at around 0, whereas the local mode is at around -6.

To highlight the potential advantages of the GENTS approach to nonlinear time series analysis, Figure 2.2 compares the predictions based on the nearest neighbour with the forecasts from a standard AR(1) model. The AR(1) model has a goodness of fit of R^2(LINEAR)=0.001, suggesting that the growth rate in the exchange rate cannot be predicted. However, the results of the GENTS model show that there are potential gains in forecasting with the goodness of fit of R^2(GENTS)=0.423. An alternative interpretation of these results is that they show why it is so difficult to predict movements in the exchange rate from a linear model; namely, the exchange rate is jumping between equilibria as it is operating in a zone of multiple equilibria and this characteristic cannot be captured by a linear model since it is based on the assumption that there is only a single equilibrium.

Although the distribution of the exchange rate is bimodal for the entire sample period, Figure 2.3 shows that the height of the smaller mode has become relatively larger since the floating of the Australian dollar in 1983:4. This property is further highlighted in Figure 2.4 which gives the area around the global mode which occurs at zero. For most of the time period, the area around the local mode is only about 0.05. However, there are four points in time where the area is between 0.2 and 0.4; namely, in 1983:4, 1985:3, 1986:7 and 1989:3. These four points in time correspond to the four largest falls in the $US/$Australian exchange rate.

The first four conditional moments of the generalized Student t distribution are given in Figure 3. Both the conditional mean, Figure 3.1, and the conditional variance, Figure 3.2, identify the same points in time which correspond to the the four largest falls in the $US/$Australian exchange rate. Moreover, the conditional variance represents a nonlinear ARCH process and thus identifies the periods of (excess) volatility. Figure 3.3 shows that for all points in time, the distribution is negatively skewed, which of course reflects the occurrence of another mode to the left of the global mode. Finally, Figure 3.4 shows that there is conditional excess kurtosis with a value at around 8, but which falls during the periods when there are large falls in the $US/$Australian exchange rate.

The importance of identifying bimodality in the exchange rate distribution is highlighted in Figure 2.1 which shows that exchange rate volatility is the result of the exchange rate switching between modes. This result provides one explanation of the fat-tails observed in exchange rate distributions: fatness is the result of the exchange rate switching between equilibria (stable or unstable). This result also highlights the limitations of ignoring the time varying nature of the exchange rate distribution since behind an atemporal leptokurtic empirical distribution is a process exhibiting multiple equilibria. Finally, the empirical results also suggest that the adoption of a flexible exchange rate has resulted in relatively greater volatility, and thus provides evidence against the hypothesis that exchange rates are less volatile under a flexible exchange rate system than under a fixed exchange rate system (Dornbusch, 1976).

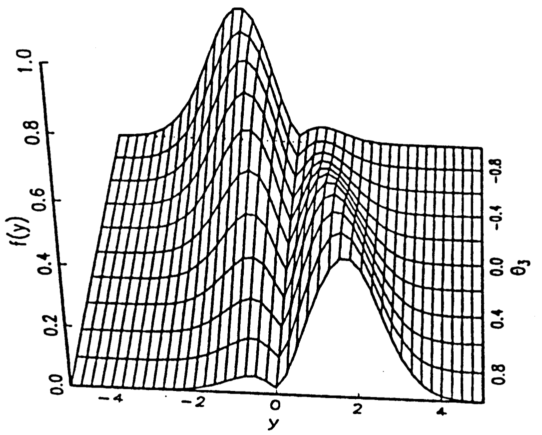
7 CONCLUSIONS

This paper has introduced a general class of nonlinear time series models based upon a generalized Student t distribution. This class of models is referred to as GENTS: Generalized Exponential Non-linear Time Series, and offers greater flexibility than a number of existing nonlinear time series approaches as it can capture a wide range of empirical behaviour such as skewed, fat-tailed and even multimodal distributions. The model can be estimated by an iterative maximum likelihood procedure and was shown to be related to other nonlinear time series models, as well as ARCH models and tests of normality.

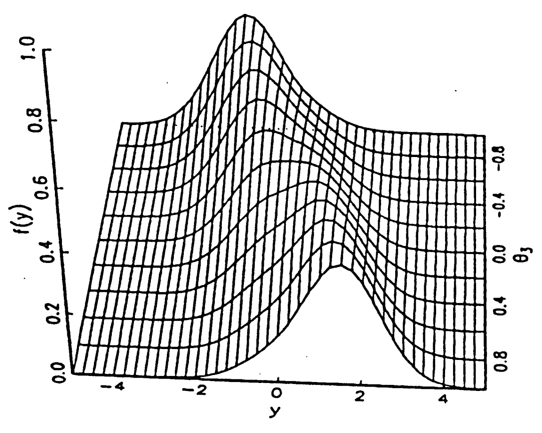
The GENTS model was applied to analysing the monthly movements in the US/Australian exchange rate for the period 1977:1 to 1989:10. The main empirical finding is that the exchange rate distribution is bimodal over the entire sample period, with the degree of bimodality increasing since the floating of the Australian currency. This suggests that the exchange rate has been operating in a zone of multiple equilibria and that the adoption of a flexible exchange rate has led to greater instability. Furthermore, the occurrence of large movements in the exchange rate is not the result of drawings from a unimodal-leptokurtic time invariant distribution, but reflects the exchange rate is jumping between equilibria in a zone of multiple equilibria. Conditional moments were also computed and used to identify the occurrence of volatility in the foreign exchange market. One important implication of the empirical results is that it provides a reason as to why it is so difficult to predict movements in the exchange rate; namely, the observed volatility in the exchange rate reflects the exchange rate switching between equilibria.

Figure 1: The Generalized Student t Distribution

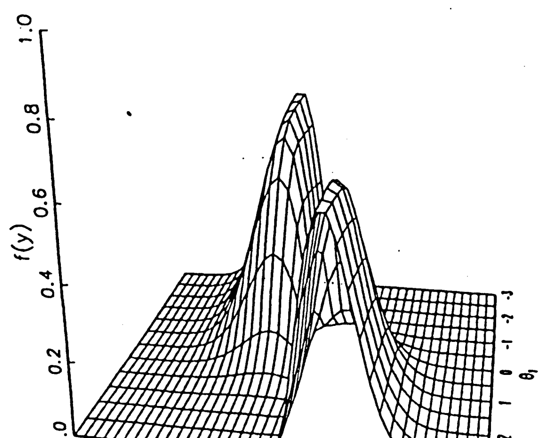
$f(y) = \exp(0.5\log(0.1^2+y^2)+\Theta_3 y-0.5y^2 - \eta)$



$f(y) = \exp(0.5\log(1.0^2+y^2)+\Theta_3 y-0.5y^2 - \eta)$



$f(y) = \exp(\Theta_1 \tan^{-1}(y/0.1)+0.5\log(0.1^2+y^2)$
$-0.5y^2 - \eta)$



$f(y) = \exp(\Theta_1 \tan^{-1}(y/1.0)+0.5\log(1.0^2+y^2)$
$-0.5y^2 - \eta)$

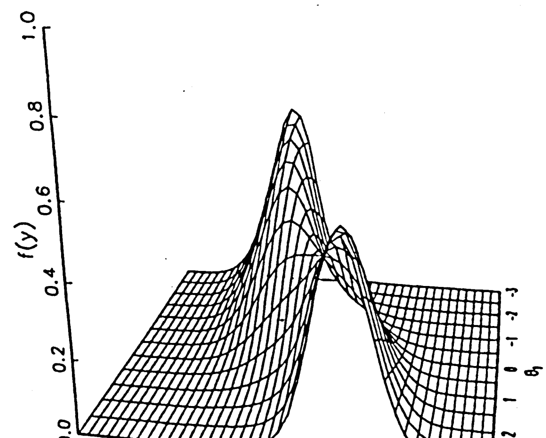


Figure 2: Temporal distributions of the rate of growth of the exchange rate

Figure 2.1: Actual and Modes

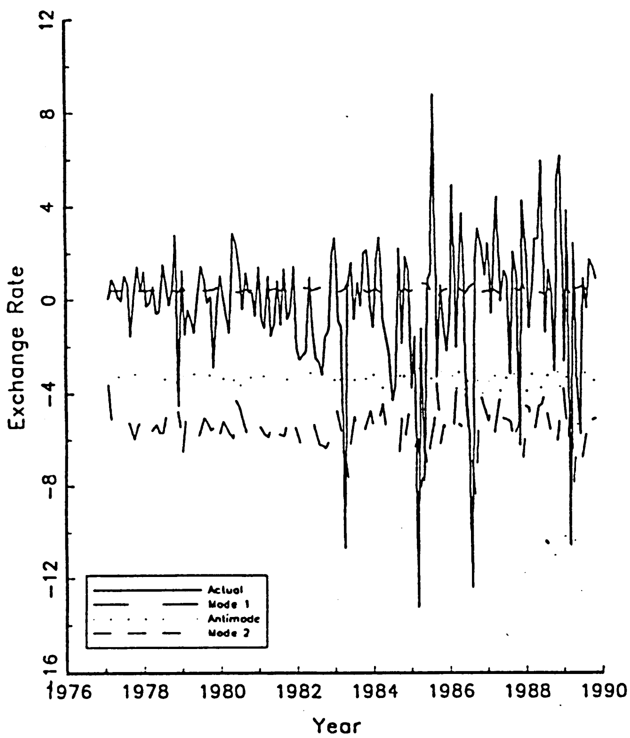


Figure 2.2: Actual and Predicted

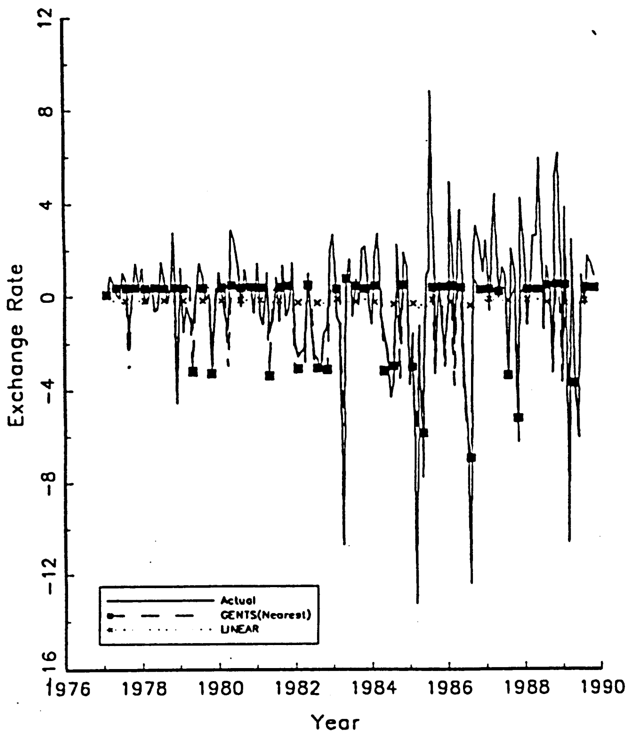


Figure 2.3: The temporal exchange rate distribution,

1977:1 to 1989:10

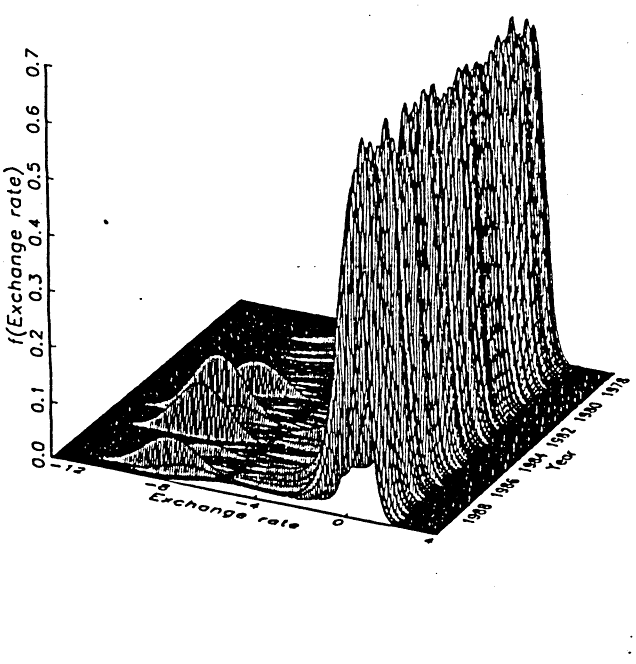


Figure 2.4: The area around the global mode

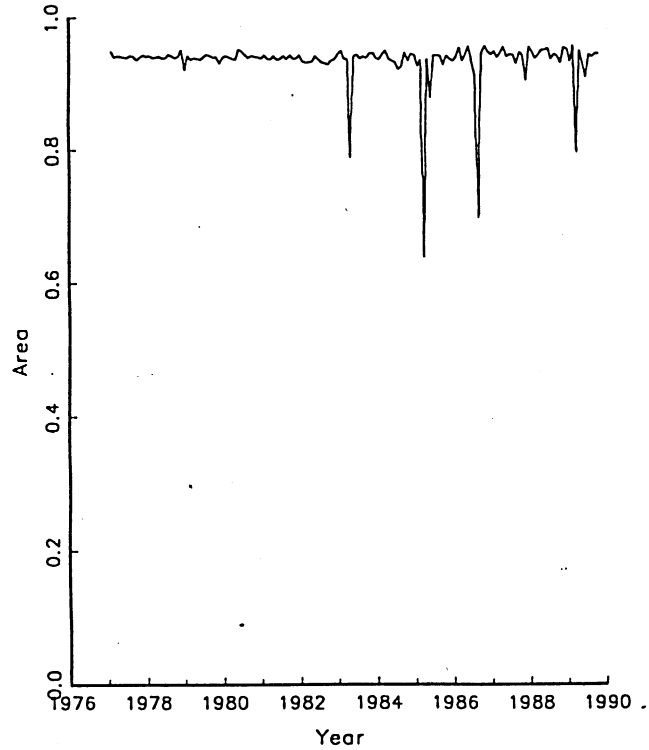


Figure 3: Conditional moments of the GENTS model

Figure 3.1: Conditional Mean

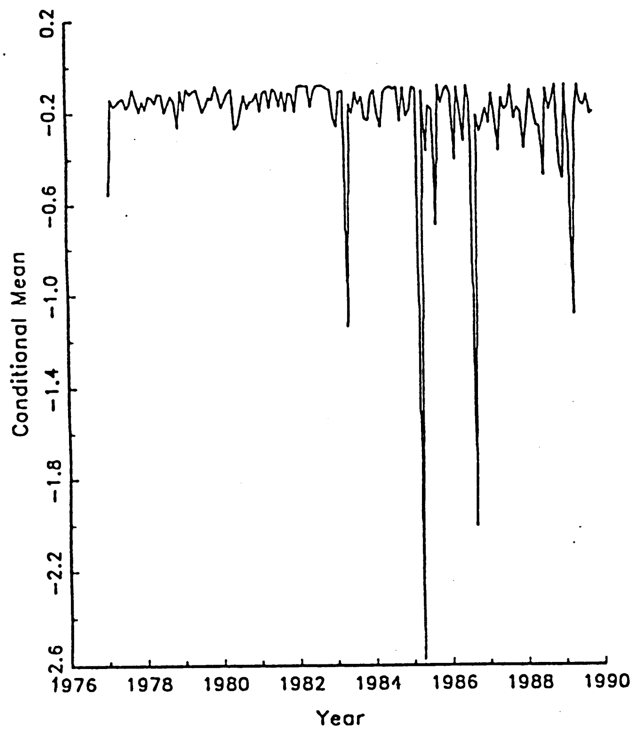


Figure 3.2: Conditional Variance

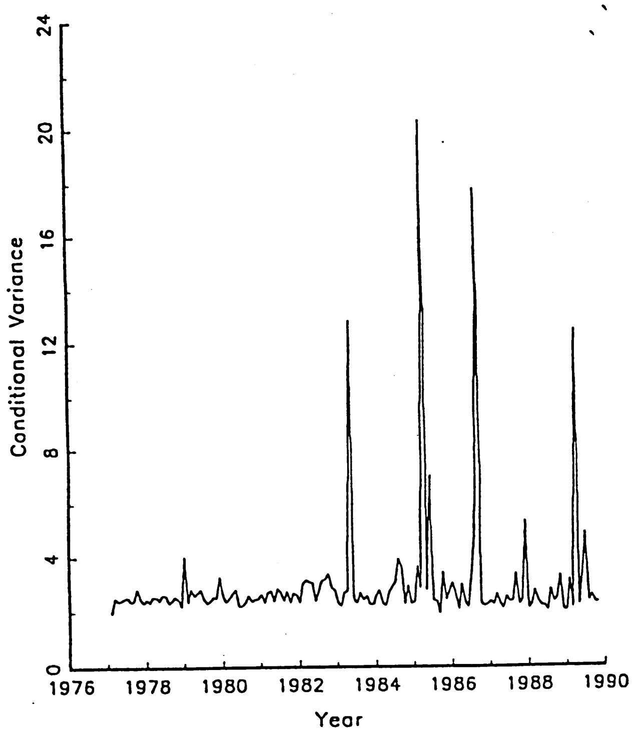


Figure 3.3: Conditional Skewness

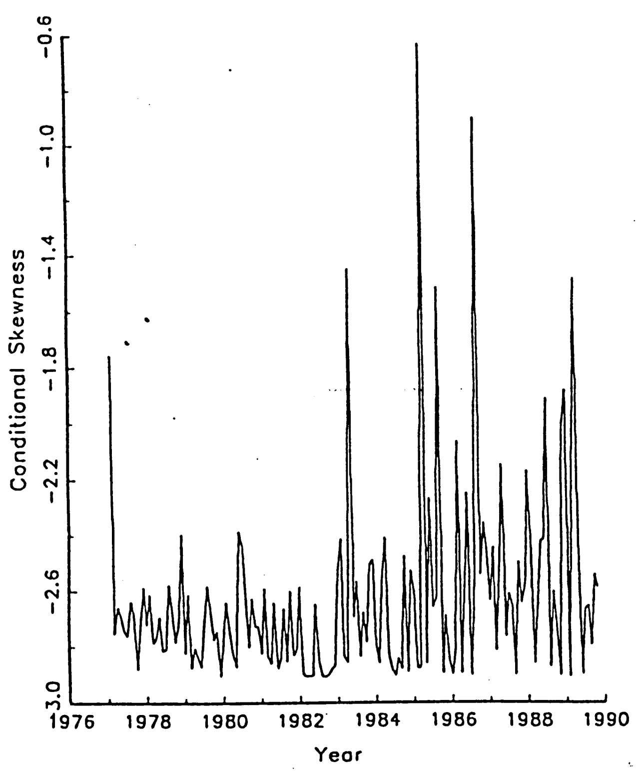
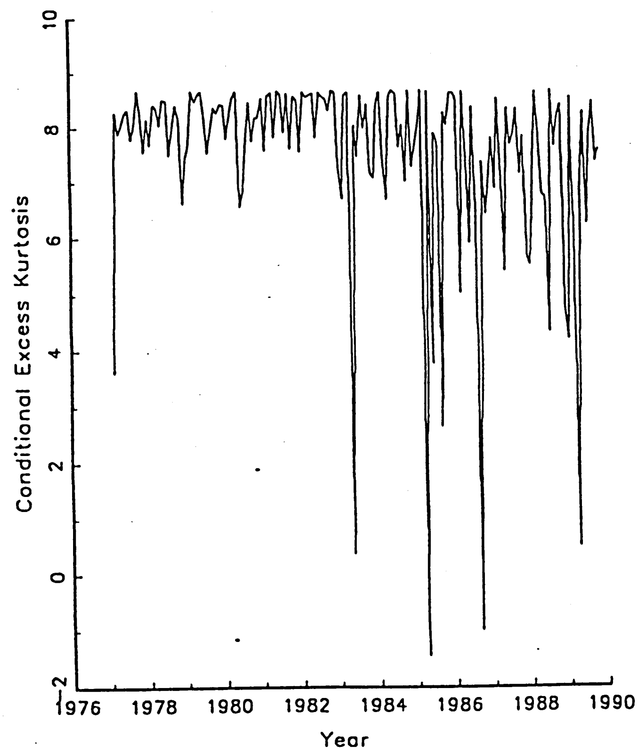


Figure 3.4: Conditional Excess Kurtosis



REFERENCES

Akgiray, V. and Booth, G.G. (1988), "Mixed Diffusion - Jump Process Modelling of Foreign Exchange Rate Movements", *The Review of Economics and Statistics*, Vol. 770, 631-637.

Ashley, R.A. and Patterson, D.M. (1989), "Linear Versus Nonlinear Macroeconomics: A Statistical Test", *International Economic Review*, 30, 685-704.

Bera, A.K. and Jarque, C.M. (1982), "Model Specification Tests: A Simultaoneous Approach", *Journal of Econometrics*, 20, 59-82.

Boothe, P. and Glassman, D. (1987), "The Statistical Distribution of Exchange Rates: Empirical Evidence and Economic Implications", *Journal of International Economics*, 22, 297-319.

Cobb, L. (1978), "Stochastic Catastrophe Models and Multimodal Distributions", *Behavioural Science*, 23, 360-374.

Cobb, L, Koppstein, P. and Chen, N.H. (1983), "Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family", *Journal of the American Statistical Association*, 78, 124-130.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.

Dornbusch, R. (1976), "Expectations and Exchange Rate Dynamics", *Journal of Political Economy*, 84, 1161-1176.

Friedman, M. and Vandersteel, S. (1982), "Short - Run Fluctuations in Foreign Exchange Rates: Evidence from the Data 1973 - 79", *Journal of International Economics*, 13, 171-186.

Gabrielson, G. (1982), "On the Unimodality of the Likelihood for the Cauchy Distribution : Some Comments", *Biometrika*, 69, 677-678.

Granger, C.W.J. and Anderson, A.P. (1978), *An Introduction to Bilinear Time Series Models*, Gottingin: Vandenhoeck and Ruprecht.

Granger, C.W.J. and Newbold, P. (1977), *Forecasting Economic Time Series*, Academic Press, New York.

Hsieh, D.A. (1989), "Modeling Heteroskedasticity in Daily Foreign-Exchange Rates", *Journal of Business and Economic Statistics*, 7, 307-317.

Keenan, D.M. (1985), "A Tukey Nonadditivity-Type Test for Time Series Nonlinearity", *Biometrica*, 72, 39-44.

Lye, J. and Martin, V.L. (1990), "Robust Estimation, Non-normalities and General Exponential Distributions", Research Paper No.286, Department of Economics, University of Melbourne.

Mandelbrot, B.B. (1963), "The Variation of Certain Speculative Prices", *Journal of Business*, 36, 394-419.

Martin, V.L. (1990), *Properties and Applications of Distributions from the Generalized Exponential Family*, unpublished Ph.D. Thesis, Monash University.

Martin, V.L. (1991), "Threshold Time Series Models as Multimodal Distribution Jump Processes", *Journal of Time Series Analysis*, forthcoming.

Nelson, D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, 59, 347-370.

Rao, S.T. and Gabr, M.M (1984), *An Introduction to Bispectral Analysis and Bilinear Time Series Models*, Lecture Notes in Statistics, Springer-Verlag.

Ramsey, J.B. (1969), "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis", *Journal of Royal Statistical Society*, Series B, 350-371.

Rogalski, R.J. and Vinso, J.D. (1978), "Empirical Properties of Foreign Exchange Rates", *Journal of International Business Studies*, 9, 69-79.

Tong, H. (1983), *Threshold Models in Non-linear Time Series Analysis*, Lecture Notes in Statistics, Springer-Verlag.

Tucker, A.L. and Pond, L. (1988), "The Probability Distribution of Foreign Exchange Price Changes: Tests of Candidate Processes", *The Review of Economics and Statistics*, 70, 638-647.

Westerfield, J.M. (1977), "An Examination of Foreign Exchange Risk under Fixed and Floating Rate Regimes", *Journal of International Economics*, 7, 181-200.