Rochester Center for

Economic Research

Autoregressive Conditional Density Estimation

Hansen, Bruce E.

Working Paper No. 322
June 1992

University of

Rochester

Autoregressive Conditional Density Estimation

Bruce E. Hansen

Rochester Center for Economic Research
Working Paper No. 322

Autoregressive Conditional Density Estimation

Bruce E. Hansen*
Department of Economics
University of Rochester

May, 1992

Abstract

Engle's ARCH model is extended to permit parametric specifications for conditional dependence beyond the mean and variance. The suggestion is to model the conditional density with a small number of "parameters", and then model these parameters as functions of the conditioning information, in the same manner as the conditional variance is modeled in standard ARCH models. Models of this form will be important for predictive density estimation, and option pricing. This method is applied to two data sets. The first application is to the monthly excess holding yield on U.S. Treasury securities, where the conditional density used is a student's t distribution. The shape parameter (the "degrees of freedom") is found to be highly sensitive to the conditioning information, implying that the conditional density varies between an extremely fat-tailed density and the standard normal. The second application is to the U.S. dollar/Swiss Franc exchange rate, using a new "skewed student t" conditional distribution. Again, the shape parameters are found to be significantly sensitive to the conditioning information.

*My thanks go to Adrian Pagan for numerous discussions which stimulated my interest in this area, and to the National Science Foundation for financial support.

1 Introduction

A typical econometric problem is to obtain an approximation to the distribution of a variable y_t, conditional on another (vector-valued) variable x_t. This includes the dynamic context where x_t contains lagged values of y_t.

Most applications include estimates of the conditional mean:

$$\mu_t = E(y_t \mid x_t). \tag{1}$$

The conditional mean may be thought of as the leading term in the conditional distribution. Many econometric applications are concerned with nothing further than the mean. The remaining error

$$e_t = y_t - \mu_t$$

in these contexts is implicitly modeled as independent of x_t.

Many applications include as well estimates of the conditional variance

$$\sigma_t^2 = \sigma^2(x_t) = E\left((y_t - \mu_t)^2 \mid x_t\right) \tag{2}$$

which may be thought of as the second term in the conditional distribution.

The conditional variance can be used to define the normalized error

$$z_t = \frac{e_t}{\sigma_t} = \frac{y_t - \mu_t}{\sigma_t}. \tag{3}$$

The normalized error z_t is a random variable whose conditional distribution is derived from the conditional distribution of y_t by the transformations (1) and (2). In most regression models, however, the conditional distribution of z_t is simply *assumed* to be independent of the conditioning variable x_t. This is typical, for example, in the "ARCH" literature which has sprung from the pioneering work of Engle [2]. While a useful simplifying assumption, there is absolutely no reason to expect the conditional distribution of the derived variable z_t to be independent of the conditioning information. Another way of saying this is that there is no reason to assume, in general, that the only features of the conditional distribution which depend upon the conditioning information are the mean and variance. Indeed, it seems quite reasonable that other features of the distribution (such as skewness and kurtosis) will depend on the conditioning information. Gallant, Hsieh and Tauchen [7] have made a clever argument of this form. They show that if the innovations e_t are generated by the mixture model $e_t = I_t^{1/2}\xi_t$ where ξ_t is iid and independent of I_t, then the variance of e_t, conditional on the past history of e_t alone, will not (in general) equal I_t, and thus the normalized error z_t will generally have a non-constant conditional distribution.

The reason why most applications have ignored higher-order features of the conditional distribution may be because only the conditional mean and variance generate significant excitement. But this lack of excitement does not imply that higher-order features should be completely ignored. First, efficient estimation of the equations for the conditional mean and variance require a complete description of the conditional distribution. Second, the aim of conditional models is often prediction, and the accuracy of predictive distributions is critically dependent upon knowledge of the correct conditional distribution for the normalized error. This point has been recently made in Baillie and Bollerslev [1]. Third, empirical models of asset pricing are incomplete unless the full conditional model is specified. Full specification may be especially important in the context of options pricing, where the price is determined by not just the conditional mean and variance, but more complicated functions of the conditional distribution.

While it might be agreed that it is desirable to allow the conditional density of z_t to depend on x_t, it is probably not clear at all how to achieve this goal. One approach, offered by Gallant, Hsieh and Tauchen [7], is to model the joint density of y_t and x_t using a series expansion about the Gaussian density. This is an innovative approach, and has the potential to reveal a lot of information concerning the underlying distribution without having to impose a great deal of *a priori* information or structure. Their approach has several drawbacks, however, First, their parameterization is not parsimonious, and therefore requires very large data sets in order to achieve a reasonable degree of precision. Second, the methods are computationally expensive, and may lay outside the reach of many routine applications. Third, the techniques may be sensitive to choices of the number of expansion terms. Theorists haven't yet completely solved many questions concerning implementation and the selection of the order of the expansion. As a result, these techniques will probably remain primarily in the hands of specialists.

This paper suggests an alternative parametric approach to modeling the conditional density of the normalized error. The approach may be regarded as a direct extension of Engle's idea to model the conditional variance as a function of lagged errors. My suggestion is to select a distribution which depends upon a low-dimensional parameter vector, and then let this "parameter vector" vary as a function of the conditional variables. In the applications presented in this paper, the student's t density and a generalization which allows for skewness are used.

This method is applied to two financial data set. The first is the excess holding yield on U.S. Treasury securities. The second is the Dollar/Franc exchange rate. In both applications strong evidence is found for variation in the conditional distribution beyond the mean and variance.

2 ARCD Model

2.1 Probability Model

The observed sample is $(y_t, x_t : t = 1, ..., n)$ which is assumed to be a realization of some jointly stationary process. We do not need to restrict the variables x_t to lie in a finite-dimensional space, so we can allow, for example, the variable x_t to include all of the (observed) past values of y_t.

We will restrict attention to distribution functions which have densities which can be written in the form

$$f(y \mid \alpha(x_t, \theta)) = \frac{d}{dy} P(y_t \leq y \mid x_t) \tag{4}$$

where θ is a finite-dimensional parameter vector and

$$\alpha_t = \alpha(x_t, \theta)$$

is a low-dimensional "time varying parameter" which fully describes the influence of x_t upon the conditional distribution. When the dimension of x_t is constant and finite, there is of course no loss in generality in writing the density function in this form, but when x_t is infinite dimensional or has a dimension which depends on t, then this class represents a meaningful restriction of the class of potential models. For reasons which will become apparent, we will denote this class of models by the name "autoregressive conditional density models" (ARCD).

For the applied model builder the conditional density function $f(y|\alpha_t)$ should be chosen so that it can capture the possible variations in the conditional distribution, subject to the limitations of the data set. In applied time series, little attention has been given to the shape of conditional densities. The density function which is almost universally used is the Gaussian (normal), where α_t is merely two-dimensional (representing the mean and variance). In a smaller number of applications, the density is either the student's t distribution or the generalized exponential (each with three parameters). On occasion, non-parametric density functions are used (for an interesting recent application, see [4]), which in practice means that the parameter α_t is high dimensional. It is interesting to observe that there are few intermediate cases in regular use. It is hard to believe that density functions with only two or three parameters can be sufficiently flexible to capture the wealth of likely distributional behaviors. On the other hand, the typical nonparametric methods go to the other extreme, employing far more "parameters" than can be adequately modeled using time-series methods in even large sample sizes. Flexible parametric density functions are sorely lacking in applied econometrics. I will return to this issue in the fourth section of the paper, where I introduce a generalization of the student's t distribution which permits skewed densities.

2.2 Normalized Parameterizations

It is particularly convenient for the reporting of applied research to rewrite the density function in terms of location and scale parameters. I will restrict attention in this exposition to cases where the location parameter is the conditional mean, and the scale parameter is the conditional variance, but the generalizations to cases where the mean or variance does not exist is straightforward and merely involves changes in notation. The idea is to parameterize the function $f(y|\alpha)$ so that we have the partition

$$\alpha_t = (\mu_t, \sigma_t^2, \eta_t)$$

where

$$\mu_t = \mu(\theta, x_t) = E(y_t \mid x_t) \tag{5}$$

is the conditional mean,

$$\sigma_t^2 = \sigma^2(\theta, x_t) = E((y_t - \mu_t)^2 \mid x_t). \tag{6}$$

is the conditional variance, and

$$\eta_t = \eta(\theta, x_t)$$

contain the remaining parameters of the conditional distribution, which we will sometimes refer to as "shape" parameters.

The conditional mean and variance allow us to define the normalized variable

$$z_t(\theta) = \frac{y_t - \mu(\theta, x_t)}{\sigma(\theta, x_t)}. \tag{7}$$

We will denote the conditional density function for z_t by

$$g(z|\eta_t) = \frac{d}{dz}P(z_t < z|\eta_t) \tag{8}$$

say. Then densities (4) and (8) are of course related as

$$f(y_t|\mu_t, \sigma_t^2, \eta_t) = \frac{1}{\sigma_t}g(z_t|\eta_t).$$

Most ARCH-type applications use probability models of the form (5)-(8), but with η_t assumed to be time invariant. The ARCD modeling strategy simply builds on this foundation by allowing the shape parameters of the density function to be time varying as well.

This formalization is convenient since there is a large literature which concerns the specification of the mean equation (5) and the variance equation (6). Parametric models include ARCH, GARCH, E-GARCH, N-ARCH, A-ARCH, plus ARCH-M versions of each (see Hentshcel [9] for a recent summary). Non-parametric models for the mean and variance equations have also been suggested, as in Pagan and Hong [17] and Gourerioux and Monfort [6].

2.3 Specification of Laws of Motion for Shape Parameters

It is necessary to specify laws of motion for the "parameters" α_t. Many strategies are possible, but the one suggested here is to follow the lead of Engle [2]. Engle's ARCH model and its generalizations have all made σ_t^2 a function of the lagged errors

$$e_t = y_t - \mu_t.$$

Since this approach has been empirically successful for the conditional variance, then it seems reasonable to believe that this strategy could also work well for other time-varying parameters in η_t. That is, the proposed modeling strategy will be to specify laws of motion of the form

$$\eta_t = \eta(e_{t-1}, e_{t-2}, ..., e_1).$$

As in the ARCH literature, we have to pay attention to boundary constraints. The conditional variance, for example, is constrained to be positive. Thus specifications of the form $\sigma_t^2 = a + be_{t-1}$ are avoided since they cannot guarantee positivity of the estimated variance sequence. One common solution (in this context) is to use specifications of the form $\sigma_t^2 = a + be_{t-1}^2$. Another solution is to use an appropriate transformation of the variance, such as $\ln \sigma_t^2 = a + be_{t-1} + ce_{t-1}^2$. Both methods have been used in the ARCH literature.

This constraint problem will certainly arise in the general ARCD context. Shape parameters arising from typical density functions often need to lie in restricted regions of the real line. Without the guidance of *a priori* theory, there is no uniformly correct approach, but a practical method which will "work" is to use a logistic transformation. Suppose that η_t is real valued and is related to a variable λ_t as

$$\eta_t = L + \frac{(U - L)}{1 + \exp(-\lambda_t)}.$$

Even if λ_t is allowed to vary over the entire real line, η_t will be constrained to lie in the region $[L, U]$. L and U should be chosen to reflect the region of interest for η_t. Combined with a law of motion for η_t such as

$$\lambda_t = a + be_{t-1} + ce_{t-1}^2$$

we obtain a relationship $\eta_t = \eta(e_{t-1})$ which is flexible yet constrained to the region $[L, U]$.

2.4 Estimation and Inference

We can write the conditional log-likelihood function as

$$\ln L(\theta \mid x_1, x_2, ..., x_n) = \sum_{t=1}^{n} l_t(\theta) \tag{9}$$

where

$$l_t(\theta) = \ln\, g(z_t(\theta)|\eta_t(\theta)) - \ln\, \sigma(\theta, x_t).$$

The maximum likelihood estimate (MLE) of the model is the value $\hat{\theta}$ which maximizes the conditional log-likelihood (9). The optimum may be found using an appropriate optimization technique.

Under the assumption of correct specification, the likelihood scores

$$\frac{\partial}{\partial\theta} l_t(\theta) = \frac{\partial}{\partial\theta} \ln g(z_t(\theta)|\eta_t(\theta)) - \frac{\partial}{\partial\theta} \ln \sigma(\theta, x_t)$$

are martingale differences and have variance

$$V = V(\theta_0), \quad V(\theta) = E\frac{\partial}{\partial\theta} l_t(\theta) \frac{\partial}{\partial\theta} l_t(\theta)' = -E\frac{\partial}{\partial\theta\partial\theta'} l_t(\theta),$$

where θ_0 denotes the true parameter value. If $El_t(\theta) < \infty$ and $E\frac{\partial}{\partial\theta}l_t(\theta) < \infty$ uniformly in θ then the MLE will be consistent. If as well $V < \infty$ and the likelihood is sufficiently well behaved in the neighborhood of θ_0 then the MLE will be asymptotically normal as well. While these are not unreasonable expectations, it is my expectation that a rigorous proof will be quite difficult to accomplish in this general setting. Lumsdaine [12] established consistency and asymptotic normality for the Gaussian GARCH(1,1) pseudo-MLE under the assumption that z_t is iid with 32 finite moments. Lee and Hansen [11] achieved a similar result under the weaker condition that z_t has a bounded conditional fourth moment. Lee [10] extended these results to incorporate the Gaussian GARCH-M model. All of these papers have confined attention to the case in which the conditional density used for estimation is the standard normal. Extension of this theory to cover the general context considered here would be desirable, but beyond the scope of the present study. We will simply assume that such theorems hold, and proceed conventionally.

Since any particular probability model is unlikely to be the "correct" model, but should more accurately be viewed as an approximation to the underlying probability structure, it is reasonable to report "robust" standard errors, as suggested by White [19], in addition to the more conventional standard errors. These give asymptotically valid confidence intervals for the "pseudo-true" parameter values which minimize the information distance between the true probability measure and the modeler's likelihood. The robust standard errors are the square roots of the diagonal elements of the matrix

$$\hat{\Omega} = \hat{M}^{-1}\hat{V}\hat{M}^{-1}$$

where

$$\hat{M} = -\sum_{t=1}^{n} \frac{\partial}{\partial\theta\partial\theta'} l_t(\hat{\theta})$$

and

$$\hat{V} = \sum_{t=1}^{n} \frac{\partial}{\partial\theta} l_t(\hat{\theta}) \frac{\partial}{\partial\theta} l_t(\hat{\theta})'.$$

2.5 Parameter Constancy

A parameter constancy test has been introduced by Lee and Hansen [11] which is particularly easy to implement. The test statistic is a member of the family of tests introduced by Nyblom [16] and modified by Hansen [8]. The statistic is an approximate LM test of the null that the parameters θ are constant against the alternative that the parameters θ follow a martingale process. The statistic is based on the cummulative moments

$$S_t = \sum_{i=1}^{t} \frac{\partial}{\partial\theta} l_i(\hat{\theta})$$

and takes the form

$$L = \frac{1}{n} \sum_{t=1}^{n} S_t' \hat{V}^{-1} S_t.$$

Under the same regularity conditions which guarantee asymptotic normality of the pseudo-MLE, the statistic L has an asymptotic distribution which depends only on the number of parameters in θ. This distribution is tabulated in [16] and [8]. The statistic L tests the null that the entire vector θ is stable against the alternative that the entire vector may be unstable. A statistic which tests the stability of an individual parameter is given by

$$L_k = \frac{1}{n} \sum_{t=1}^{n} S_{kt}^2 / \hat{V}_{kk}$$

where S_{kt} is the kth element of S_t and $\hat{V}_{kk}$ is the kth diagonal element of $\hat{V}$. The asymptotic 1% critical value for the individual statistics is 0.75, and the asymptotic 5% critical value is 0.47.

2.6 Non-Parametric Density Comparisons

A interesting yet informal diagnostic can be obtained by comparing the density function for the errors implied by the model with that calculated using a non-parametric kernel technique. Discrepancies can suggest useful modifications to the model specification. Such a procedure is outlined in this subsection.

The parametric assumption is that the density function for $z_t = z_t(\theta)$ is given by $g(z_t|\eta_t)$ where $\eta_t = \eta(\theta_0, x_t)$. Although θ_0 is unknown, the function is estimated by $g(z_t|\hat{\eta}_t)$ where $\hat{\eta}_t = \eta(\hat{\theta}, x_t)$. When $\eta(\theta_0, x_t)$ doesn't depend on x_t, this gives a fixed density function which we can plot. In the general case in which η_t varies with x_t, this is not possible, since z_t has a density whose shape varies across different values of x_t. Insight is gained, however, by noting that although the *conditional* density of z_t is a function of x_t, the *unconditional* density of z_t is a simple function which can be plotted. It is given by

$$g(z) = E_\eta g(z|\eta)$$

where E_η takes expectations over η. $g(z)$ is a *mixtures* distribution, where the mixing is over the shape parameters η.

The fact that $g(z)$ can be represented as an expectation suggests that it is naturally estimated by the empirical expectation

$$\hat{g}(z) = \hat{E}_\eta g(z|\eta) = \frac{1}{n}\sum_{t=1}^{n} g(z|\hat{\eta}_t)$$

where $\hat{E}_\eta$ is the probability measure which puts a mass of $1/n$ at each value of $\hat{\eta}_t$, (that is, the empirical distribution of the estimates values of $\hat{\eta}_t$). Since the $\hat{\eta}_t$ are calculable from $\hat{\theta}$ and $g(z|\eta)$ is a known function, $\hat{g}(z)$ is easily calculable. This gives the estimate of the density of z_t which is implied by the model.

The density $g(z)$ may be a good approximation to the actual unconditional density of the standardized errors z_t, but it need not be. The model will restrict the class of permissible density shapes. For example, a conditional student's t distribution will restrict $g(z)$ to be unimodal and symmetric, but these basic features need not be valid descriptions of the underlying errors.

An estimator of the unconditional density which is not dependent on the model structure is given by a nonparametric kernel estimator applied to the standardized residuals z_t. The Rosenblatt-Parzen estimator, for example, is given by

$$\tilde{g}(z) = \frac{1}{nh} \sum_{t=1}^{n} K \left(\frac{\hat{z}_t - z}{h} \right)$$

where h is the bandwidth which controls the amount of local smoothing, and $K(\cdot)$ is a kernel function.

The adequacy of the parametric model can be informally assessed by plotting the parametric estimate $\hat{g}(z)$ with the nonparametric estimate $\tilde{g}(z)$. Discrepancies can help lead to reformulated models with better fit. Unfortunately, formal comparisons are quite difficult. In general, the asymptotic distribution of the empirical probability measures involved depend on the fact that there are estimated parameters. Thus while analogs to the Kolmogorov-Smirnov class of statistics can be calculated, there is no easily accessible large sample theory to provide guidance concerning critical values.

3 A Conditional Student Model for the Term Structure

3.1 Basic Structure

This section describes a study concerning the short-run term structure of interest rates. The data, monthly observations on returns to U.S. Treasury securities for the period December 1946 to February 1987, come from Appendix II, Table A-1 of McCulloch [13]. His returns series were calculated from the prices of whole securities, and were adjusted for changes in tax legislation. Figure 1 plots the one-month yield rate R_t and the instantaneous yield rate r_t.

From his tables, the excess holding yield, y_t was calculated as

$$y_t = \frac{(1+R_t)^2}{1+r_{t+1}} - (1+r_t),$$

and the interest differential, i_t, was calculated as

$$i_t = R_t - r_t.$$

These two series are displayed in Figure 2.

In our earlier notation, $x_t = (y_{t-1}, y_{t-2}, ...; i_t, i_{t-1}, ...)$, since we are interested in obtained the distribution of the excess holding yield, conditional on the current interest differential and lagged values of these two series. As discussed in [5] and [17], the interest differential plays an important role in empirical models of the excess holding yield, even though the expectations hypothesis implies otherwise.

3.2 Specification of the Conditional Mean

The main thrust of this exercise is not on the conditional mean or variance, but is to demonstrate that allowing for higher-order dependence yields significant gains. Yet the specification of the mean and variance equations cannot be taken lightly, for it is clear that errors in their specification may result in spurious higher-order findings. At the same time, it is important (from both computational and precision viewpoints) not to heavily over-parameterize the model. The approach adapted in this application is to model the equations sequentially, using the vehicle of the Gaussian likelihood to select the equations for the mean and variance. This will enable us to feasibly estimate and compare a large number of models.

The use of a misspecified Gaussian likelihood has been justified by the asymptotic theory of Lee and Hansen [11] (for the GARCH(1,1) model) and Lee [10] (for the GARCH-M model). These papers showed that so long as the mean and variance equations properly describe the conditional mean and variance, the Gaussian pseudo-likelihood parameter estimates will be consistent and asymptotically normally distributed. Their work, unlike the earlier theoretical literature, did not require the standardized error z_t to be an independent sequence, thus allowing for a general ARCD model to be generating the data.

Table 1 reports the Gaussian maximum likelihood estimates of a fairly general specification of the conditional mean, with a fairly simple specification of the conditional variance. In all of the tables, the maximum likelihood estimates, the conventional standard errors, and the White robust standard errors are reported. The Nyblom L_k statistics for each parameter are reported. In the variance equation, the variance is reported as a linear function of σ_{t-1}^2 and $e_{t-1}^2 - \sigma_{t-1}^2$. This was done so that the coefficient on the former can be interpreted as a measure of persistence in the variance, as it is unity in the "IGARCH" model.

In Table 1, a large number of the individual coefficients appear insignificantly different from zero. A more parsimonious model was selected by successfully eliminating the variable with the smallest t-statistic, until the model reported in Table 2 was obtained. The only exceptions to the smallest t-statistic rule were that the intercept was always maintained, and the conditional standard deviation was retained until the final step. The latter was done since the possibility of a significant "GARCH-M" effect has long been believed to be important for the excess holding yield on Treasury securities. The model of Table 2 has eight fewer parameters than the model of Table 1, with an increase in the log-likelihood of only 3.04, which is far from a statistically significant difference.

It is interesting to compare these results with an alternative, simpler specification reported in Table 3. The major difference is that only the

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Mean Equation				
intercept	0.02	0.03	0.03	0.05
σ_t	0.09	0.11	0.11	0.07
i_t	1.17	0.15	0.17	0.17
i_t^2	-0.78	0.23	0.23	0.04
i_{t-1}	0.25	0.14	0.16	0.03
i_{t-1}^2	0.14	0.19	0.18	0.03
i_{t-2}	0.24	0.14	0.16	0.10
i_{t-2}^2	-0.12	0.15	0.12	0.36
i_{t-3}	0.12	0.14	0.16	0.05
i_{t-3}^2	0.26	0.11	0.12	0.03
y_{t-1}	0.03	0.09	0.10	0.12
y_{t-1}^2	0.10	0.05	0.04	0.08
$i_t y_{t-1}$	0.25	0.16	0.17	0.04
$i_{t-1} y_{t-1}$	-0.35	0.18	0.16	0.04
$i_{t-2} y_{t-1}$	0.02	0.17	0.12	0.07
$i_{t-3} y_{t-1}$	-0.19	0.17	0.15	0.05
Variance Equation				
intercept	-0.00001	0.0008	0.0010	0.09
$e_{t-1}^2 - \sigma_{t-1}^2$	0.21	0.06	0.10	0.12
i_t^2	0.07	0.03	0.03	0.26
σ_{t-1}^2	1.01	0.02	0.04	0.04
Log L	323.7			
Nyblom L	4.12			

Table 1: Excess Holding Yield: Unrestricted Gaussian Model

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Mean Equation				
intercept	0.04	0.23	0.25	0.04
i_t	1.24	0.12	0.13	0.13
i_t^2	-0.51	0.14	0.13	0.10
i_{t-1}	0.35	0.12	0.12	0.08
i_{t-2}	0.19	0.09	0.10	0.11
i_{t-3}^2	0.30	0.09	0.09	0.03
y_{t-1}^2	0.12	0.09	0.03	0.12
$i_{t-1}y_{t-1}$	-0.29	0.03	0.10	0.07
Variance Equation				
intercept	-0.0001	0.0007	0.0009	0.10
$e_{t-1}^2 - \sigma_{t-1}^2$	0.22	0.06	0.10	0.14
i_t^2	0.07	0.03	0.03	0.30
σ_{t-1}^2	1.01	0.03	0.04	0.04
Log L	326.7			
Nyblom L	2.46			

Table 2: Excess Holding Yield: Restricted Gaussian Model

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Mean Equation				
intercept	0.02	0.04	0.06	0.10
σ_t	0.32	0.11	0.16	0.06
i_t	0.99	0.13	0.17	0.70
y_{t-1}	0.10	0.06	0.07	0.42
Variance Equation				
intercept	0.0004	0.0013	0.0017	0.20
$e_{t-1}^2 - \sigma_t^2$	0.21	0.04	0.07	0.09
i_t^2	0.16	0.07	0.10	0.34
σ_{t-1}^2	0.97	0.03	0.05	0.08
Log L	342.9			
Nyblom L	3.57			

Table 3: Excess Holding Yield: Naive Gaussian Model

current value of the interest differential is included in the conditional mean equation. In this specification, the conditional standard deviation appears to be statistically significant in the mean equation, as is commonly found in this literature. Note that the likelihood ratio statistic for this restricted model is 38.4, which is statistically significant at the 1% level. This restricted model also fails the Nyblom-Hansen parameter stability test. The L statistic of 3.6 exceeds the 1% null critical value of 2.6. The individual stability tests suggest that the coefficient on i_t is not stable. Note that these problems do not arise for the general models of Tables 1 and 2, where extra lags of the interest differential are included. An important lesson here is that the stability tests are useful diagnostics. If the model of Table 3 were estimated first, the large stability test statistics would alert a careful researcher that further study of the dynamic specification is needed.

Another interesting contrast between the models of Table 3 and Tables 1 and 2 is the difference between the conventional standard errors and the robust standard errors. In Tables 1 and 2 the two estimates are nearly the same, but in Table 3 the estimates are quite different. This is also informal evidence against the specification (this informal comparison could be made rigorous using a White information matrix test).

For the rest of the analysis, we will use the specification for the conditional mean and variance as given in Table 2. The specification of the conditional variance was also examined. Additional lags of the e_{t-1}^2 and i_t^2 were also included, but were not statistically significant and so the model was not augmented. It appears that the model reflected in Table 2 provides a good specification for the conditional mean and variance. We now turn to modeling other features of the conditional distribution.

3.3 Student T Likelihood

The fit of the Gaussian model can be informally assessed by comparing a nonparametric estimate of the density of the standardized residuals z_t with the standard normal density. These are displayed in figure 3. The kernel estimate reveals a more peaked and fat-tailed density than the standard normal. As a first approximation, it appears that a student's t distribution might make a better fit.

A student's t density function normalized to have unit variance is given by

$$f(z \mid \eta) = \frac{\Gamma(\frac{\eta+1}{2})}{\sqrt{\pi(\eta-2)}\Gamma(\frac{\eta}{2})}\left(1+\frac{z^2}{(\eta-2)}\right)^{-(\eta+1)/2}, \; -\infty < z < \infty \qquad (10)$$

where $2 < \eta < \infty$. As $\eta \to \infty$, this density approaches the standard normal; in fact, the match is quite good for η above 30. The "degrees of freedom" η is constrained to exceed two, as we have normalized z to have a finite variance. As $\eta \to 2$, the density becomes increasing peaked, and is ill-behaved in the neighborhood of 2.

We start with a conventional student's t model with a constant degrees of freedom parameter. The MLE for this model are given in Table 4. The parameter estimates and standard errors for the conditional mean and variance are not dramatically different than those from the Gaussian MLE. The degrees of freedom parameter is estimated to be 5.7, which implies a fairly fat tail. The fit of the model is a dramatic improvement over the Gaussian, with the log-likelihood changing by 11.1.

To assess the fit of the model, we display in Figure 4 the non-parametric and parametric estimates of the density normalized residuals. Here the parametric estimate is simply a student's t density with 5.7 degrees of freedom. The fit appears to be much better than for the Gaussian pseudo-likelihood.

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Mean Equation				
intercept	0.04	0.02	0.02	0.05
i_t	1.17	0.13	0.16	0.32
i_t^2	-0.45	0.15	0.16	0.03
i_{t-1}	0.35	0.10	0.10	0.05
i_{t-2}	0.19	0.09	0.10	0.24
i_{t-3}^2	0.28	0.09	0.11	0.02
y_{t-1}^2	0.11	0.04	0.03	0.05
$i_{t-1}y_{t-1}$	-0.29	0.12	0.11	0.04
Variance Equation				
intercept	-0.00008	0.00090	0.0010	0.13
$e_{t-1}^2 - \sigma_t^2$	0.20	0.07	0.10	0.10
i_t^2	0.11	0.05	0.07	0.31
σ_{t-1}^2	0.99	0.03	0.04	0.06
Degrees of Freedom	5.7	1.56	1.60	.16
Log L	315.6			
Nyblom L	3.16			

Table 4: Excess Holding Yield: Student's t Model

It is hard to know if the remaining differences are due to random error or not.

3.4 Conditional Student Likelihood

As discussed in the introduction, there is no reason to believe that the only time-varying features of the conditional distribution are the mean and variance. We now allow for the shape of the conditional density to be time-varying through the degrees of freedom parameter, using a specification of the form presented in section 2.3.

A logistic function was used to bound the time-varying conditional degrees of freedom parameter to lie between a lower and an upper bound, which were chosen to be 2.1 and 30, respectively. The upper bound was selected simply because the student's t distribution is virtually indistinguishable from the standard normal for any value of η above 30. The lower bound is perhaps more critical. The normalized student's t density is not defined for $\eta = 2$, so needs to be bounded away from 2. Some visual experimentation suggested that setting $L = 2.1$ wasn't too extreme a choice, and the numerical operations didn't appear to find this choice offensive. The function was completed by making the logistically transformed η_t a quadratic function of the information set. The complete specification is

$$\frac{\eta_t - 2.1}{27.9} = \frac{1}{1 + \exp(-\lambda_t)}$$

$$\lambda_t = \lambda_0 + \lambda_1 e_{t-1} + \lambda_2 e_{t-1}^2 + \lambda_3 i_t + \lambda_4 i_t^2 + \lambda_5 e_{t-1} i_t. \qquad (11)$$

This function is quite flexible and will allow for a wide range of relationships.

To optimize the global likelihood, I found that it was easiest to first use the normalized residuals from the previously estimated model, and fit equation (11) alone. This provided a good set of starting values for the complete likelihood.

The estimates are reported in Table 5. Most of the coefficient estimates of the mean and variance equations are quite similar to those of Table 4, and most of the standard errors are smaller. The change in the log-likelihood (from the student t model) is 6.4, yielding a likelihood ratio statistic of 12.8 which has a p-value of 2.5% using a chi-square distribution with five degrees of freedom. While we cannot be certain of the validity of the asymptotic

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Mean Equation				
intercept	0.02	0.02	0.02	0.05
i_t	1.14	0.11	0.14	0.30
i_t^2	-0.34	0.08	0.07	0.03
i_{t-1}	0.42	0.10	0.09	0.07
i_{t-2}	0.16	0.09	0.10	0.27
i_{t-3}^2	0.28	0.09	0.11	0.02
y_{t-1}^2	0.12	0.03	0.02	0.06
$i_{t-1}y_{t-1}$	-0.35	0.10	0.10	0.03
Variance Equation				
intercept	0.00003	0.00112	0.001142	0.14
e_{t-1}^2	0.23	0.08	0.12	0.13
i_t^2	0.09	0.05	0.05	0.25
σ_{t-1}^2	1.03	0.04	0.05	0.05
Degrees of Freedom				
intercept	-2.44	0.55	0.60	0.08
e_{t-1}	-0.23	0.66	0.48	0.23
e_{t-1}^2	-0.05	0.37	0.23	0.07
i_t	3.33	1.97	1.94	0.14
i_t^2	3.27	2.59	2.64	0.04
$e_{t-1}i_t$	-4.14	2.44	2.39	0.03
Log L	309.2			
Nyblom L	3.81			

Table 5: Excess Holding Yield: Conditional Student's t Model

approximation, it seems reasonable to believe that this provides evidence against the assumption that the conditional distribution of the normalized errors is independent of the conditioning information. This particular model (the conditional student t) may not be the "truth", but it does appear to give a statistically significant increase in fit, and therefore a better description of the time series process for excess holding yields.

To assess the fit of the model we can examine the densities of the residuals and their normalized counterparts. The non-parametric estimates were obtain by kernel estimation as before. To obtain parametric estimates, we now have to average over the realized values of $\hat{\eta}_t = \eta(x_t, \hat{\theta})$ as discussed in section 2.6. The estimates are displayed in Figure 5. They appear neither better nor worse than those obtained from the student t pseudo-likelihood estimates.

Parameter estimates from tables often do not give a good feel between conditioning variables and the objects of interest, and this is certainly true concerning the estimated relationship for the degrees of freedom, so I have displayed the non-linear relationship in a 3-D graph in Figure 6. The vertical axis gives the estimated degrees of freedom, and the other axes the interest differential and lagged residual. It is easy to see a strong quadratic effect in the interest differential (so the degrees of freedom is small for i_t near zero, and a more mild quadratic effect in e_{t-1}.

Figure 7 displays the estimated degrees of freedom parameter over the sample period. Note that most of the estimates are close to 5, with some visits down to the lower boundary of 2.1 (implying a very fat tailed distribution) and some up towards, and even hitting, the upper boundary of 30 (implying a near-Gaussian distribution). Unfortunately, the "degrees of freedom" parameterization disguises some information, since the shape of the density is much more sensitive to changes in η when η is small than when it is large. The plot of Figure 7 emphasizes the large movements between 10 and 30, which are probably less significant than the movements between 2 and 3. To alleviate this deficiency, we plot in Figure 8 the inverse of the degrees of freedom, $1/\eta_t$. In this picture, the lower boundary, 0, represents a Gaussian density, and the upper boundary, 1/2, represents the limit of the fattailed densities. Another method to assess the behavior of the estimated process for the degrees of freedom parameter is to estimate its unconditional density. This is shown in Figure 9. This shows clearly that η_t is typically close to the modal value, 5.

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Variance Equation				
intercept	0.033	0.0025	0.030	0.34
$e_{t-1}^2 - \sigma_t^2$	0.15	0.04	0.05	0.23
σ_{t-1}^2	1.01	0.02	0.02	0.40
Degrees of Freedom	8.2	2.8	2.5	1.79
Log L	1142.6			
Nyblom L	2.3			

Table 6: Exchange Rate: Student T Model

4 A Skewed Student's T Model for the Exchange Rate

One commonly analyzed series in the ARCH literature is the monthly dollar/Swiss Franc exchange rate. Engle and Bollerslev [3] studied this series, and suggested a GARCH(1,1) specification with a student's t density. Maximum likelihood estimates for this specification are given in Table 6. Figure 10 displays the nonparametric and parametric estimates of the density of the standardized residuals. While this model survives a number of standard specification tests (such as tests for omitted variables) the degrees of freedom parameter decisively fails the Nyblom constancy test. The test statistic 1.79 is over twice the 1% critical value. This indicates that the model specification is not adequate.

As a first pass, we try a conditional student's t model, making the logistically transformed student's t parameter (bounded between 2.1 and 30) a linear function of e_{t-1} and e_{t-1}^2. These results are given in Table 7. The p-value for the increase in the likelihood is 10%, which cannot be taken as strong evidence for the augmented model, and the Nyblom stability test statistic still rejects the specification.

The student's t family is a fairly restrictive parametric family, only allowing for variation in the location, scale, and tail thickness of the density. To allow for a richer set of behaviors, we may need a more flexible family of probability densities. What would be desirable, I believe, is to use a density function which allows for skewness, but specializes to a shape similar to the student's t. In order to keep in the ARCH tradition, it is also important

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Variance Equation				
intercept	0.031	0.025	0.031	0.29
$e_{t-1}^2 - \sigma_t^2$	0.17	0.05	0.06	0.20
σ_{t-1}^2	1.01	0.02	0.02	0.35
Degrees of Freedom				
intercept	-1.07	0.73	0.79	1.59
e_{t-1}	-0.38	0.24	0.19	0.22
e_{t-1}^2	-0.08	0.07	0.06	0.33
Log L	1140.36			
Nyblom L	2.44			

Table 7: Exchange Rate: Conditional Student T Model

to have density functions which can be easily parameterized so that the innovations are mean zero and unit variance. Otherwise, it will be difficult to identify which fluctuations are in the mean and variance, and which are fluctuations in the shape of the conditional density.

For the following study, I have use the following density function, which is a simple generalization of the student's t density, and allows for skewness.

$$f(z \mid \eta, \lambda) = \begin{cases} bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1-\lambda}\right)^2\right)^{-(\eta+1)/2}, & z < -a/b, \\ bc\left(1 + \frac{1}{\eta-2}\left(\frac{bz+a}{1+\lambda}\right)^2\right)^{-(\eta+1)/2}, & z \geq -a/b, \end{cases} \tag{12}$$

where $2 < \eta < \infty$, and $-1 < \lambda < 1$. The constants a, b,and c are given by

$$a = 4\lambda c\left(\frac{\eta-2}{\eta-1}\right), \tag{13}$$

$$b^2 = 1 + e\lambda^2 - a^2, \tag{14}$$

and

$$c = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma\left(\frac{\eta}{2}\right)}. \tag{15}$$

In the appendix, we show that this is a proper density function with a mean of zero and a unit variance.

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Variance Equation				
intercept	0.032	0.0024	0.029	0.40
$e_{t-1}^2 - \sigma_t^2$	0.15	0.04	0.05	0.25
σ_{t-1}^2	1.00	0.02	0.02	0.43
Degrees of Freedom	8.1	2.7	2.5	1.60
Skew Parameter	-0.09	0.05	0.05	1.42
Log L	1141.2			
Nyblom L	3.1			

Table 8: Exchange Rate: Skewed Student T Model

Inspection of the density function reveals that the density is continuous, and has a single mode at $-a/b$, which is of opposite sign with the parameter λ. Thus if $\lambda > 0$, the mode of the density is to the left of zero and the variable is skewed to the right, and vice-versa when $\lambda < 0$. Figure 11 displays plots of the density for a few parameterizations.

Two estimated models are reported using the skewed student's t density function. The model estimated in Table 8 does not make the two shape parameters (η and λ) time-varying, and the model estimated in Table 9 allows both to be functions of e_{t-1}. As before, η_t is bounded between 2.1 and 30. λ_t is bounded between -.9 and .9, using the logistic function. Both logistically transformed variables are specified as quadratic functions of e_{t-1}.

The estimates in Table 8 for the variance equation and the degrees of freedom are essentially the same as before. The skewness parameter is negative, implying a density which is skewed to the left. The parametric and nonparametric estimates of the density function for z_t are very similar to those of figure 10, so our omitted.

To assess statistical significance, it is interesting to compare the four likelihoods of Tables 6-9. Simply allowing for η_t to be time-varying (Table 7) or the density to be skewed (Table 8) only produces a marginally significant change in the likelihood. But allowing for both effects simultaneously (Table 9) produces a LR test statistic (against the student's t model of Table 6) of 13.5 which has a p-value of 2% using a chi-square distribution with five degrees of freedom. This again provides strong evidence that parametrically-specified time-varying conditional densities are statistically important as descriptions of the time series properties of financial data.

Variables	Estimate	St. Error	Robust SE	Nyblom L_k
Variance Equation				
intercept	0.037	0.029	0.037	0.33
$e_{t-1}^2 - \sigma_t^2$	0.20	0.06	0.10	0.18
σ_{t-1}^2	1.03	0.03	0.04	0.41
Degrees of Freedom				
Intercept	-1.10	0.73	0.91	1.06
e_{t-1}	-0.54	0.21	0.20	0.23
e_{t-1}^2	-0.08	0.05	0.05	0.34
Skew Parameter				
Intercept	-0.06	0.14	0.14	0.95
e_{t-1}	-0.13	0.09	0.09	0.21
e_{t-1}^2	-0.10	0.05	0.07	0.07
Log L	1135.9			
Nyblom L	3.22			

Table 9: Exchange Rate: Conditionally Skewed Student T Model

Figure 12 displays the time series $\hat{\eta}_t$, and figure 13 displays $\hat{\eta}_t^{-1}$. From the latter it is clear that η_t is primarily hovering around 10, with occasional excursions into the more fat-tailed region. Figure 14 displays an estimate of the density for η_t. Figure 15 displays the estimates λ_t. The sequence is typically near zero, with the density becoming conditionally skewed after large squared innovations. Figure 16 displays a nonparametric estimate of the density of the process λ_t.

Unfortunately, the Nyblom stability test statistics for both the conditional degrees of freedom and skewness equations indicate misspecification. Attempts to rectify this problem by adding extra lags of e_{t-1} to the equations had no effect (the parameter estimates were very small and insignificant). It is also possible that these test statistics are revealing a nonstationary feature of the conditional distribution, which cannot be easily incorporated in an ARCH-type framework. This calls for further research.

5 Conclusion

This paper has generalized Engle's ARCH model to let shape parameters beyond the variance depend upon conditioning information. This is achieved simply by using a low-dimensional parametric family for the conditional density, and letting each parameter be a parametric function of the data. Two particular examples of this approach, using a conditional student t distribution and a new conditional skewed student t distribution, are developed and used to model the one-month excess holding yield on U.S. Treasury securities, and montly dollar/Franc exchange rate, respectively. The shape parameters of the conditional densities are found to be statistcally significant at the 5% level.

References

[1] Baillie, R.T. and T. Bollerslev (1992): "Prediction in dynamic models with time-dependent conditional variances," *Journal of Econometrics*, 52, 91-114.

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

[3] Engle, R.F. and T. Bollerslev (1986): "Modeling the persistence of conditional variances," *Econometric Reviews*, 5, 1-50.

[4] Engle, R.F. and G. Gonzalez-Rivera (1991): "Semiparametric ARCH models," *Journal of Business and Economic Statistics*, 9, 345-360.

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

[6] Gourieroux, C. and A. Monfort (1992): "Qualitative threshold ARCH models," *Journal of Econometrics*, 52, 159-200.

[7] Gallant, R.A., D.A. Hsieh and G.E. Tauchen (1991): "On fitting a recalcitrant series: The pound/dollar exchange rate, 1974-1983," in W. Barnett, J. Powell, and G. Tauchen eds., *Semiparametric and nonparametric methods in econometrics and statistics*, Cambridge University Press.

[8] Hansen, B.E. (1990): "Lagrange multiplier tests for parameter instability in non-linear models," unpublished paper, University of Rochester.

[9] Hentschel, L. "The Absolute value GARCH model and the volatility and U.S. stock returns," unpublished paper, Princeton University.

[10] Lee, S.W. (1992): "Asymptotic properties of the maximum likelihood estimator of the GARCH-M and IGARCH-M models," unpublished paper, University of Rochester.

[11] Lee, S.W. and B.E. Hansen (1992): "Asymptotic properties of the maximum likelihood estimator and test of the stability of parameters of the GARCH and IGARCH models," unpublished paper, University of Rochester.

[12] Lumsdaine, R.L. (1991): "Asymptotic properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models," unpublished paper, Princeton University.

[13] McCulloch, J.H. "Appendix II: U.S. Term Structure Data, 1946-87," in *Handbook of Monetary Economics*, forthcoming.

[14] Nelson, D.B. (1990): "Conditional heteroskedasticity in asset returns: A new approach," *Econometrica*, 59, 347-370.

[15] Nelson, F.D. and N.E. Savin (1990): "The danger of extrapolating asymptotic local power," *Econometrica*, 58, 977-982.

[16] Nyblom, J. (1989): "Testing the constancy of parameters over time," *Journal of the American Statistical Association*, 84, 223-230.

[17] Pagan, A.R. and Y.S. Hong (1991): "Non-parametric estimation and the risk premium," in: W. Barnett, J. Powell, and G. Tauchen, eds., *Semiparametric and nonparametric methods in econometrics and statistics*, Cambridge University Press.

[18] Pagan, A.R. and H.C.L. Sabau (1987): "Consistency tests for heteroskedastic and risk models," unpublished paper, Australian National University.

[19] White, H. (1982): "Maximum likelihood estimation of misspecified models," *Econometrica,* 50,1-25.

A Appendix

In this appendix we show that density (12) is a proper density with a mean of zero and unit variance. It will be convenient, however, to first analyze a random variable Z with density

$$g(y \mid \eta, \lambda) = \begin{cases} bc\left(1 + \frac{1}{\eta-2}\left(\frac{y}{1-\lambda}\right)^2\right)^{-(\eta+1)/2}, & y < 0, \\ bc\left(1 + \frac{1}{\eta-2}\left(\frac{y}{1+\lambda}\right)^2\right)^{-(\eta+1)/2}, & y \geq 0, \end{cases} \tag{16}$$

where the constants b and c are given in (14) and (15). Let $f(x \mid \eta)$ denote the student's t density normalized to have a unit variance, as in (10), which equals $g(x \mid \eta, 0)$. By the transformation $x = y/(1-\lambda)$ we see

$$\int_{-\infty}^{0} g(y \mid \eta, \lambda)dy = (1-\lambda)\int_{-\infty}^{0} f(x \mid \eta)dx = \frac{1-\lambda}{2},$$

and by the transformation $x = y/(1+\lambda)$ we find

$$\int_{0}^{\infty} g(y \mid \eta, \lambda)dy = (1+\lambda)\int_{0}^{\infty} f(x \mid \eta)dx = \frac{1+\lambda}{2}.$$

Thus

$$\int_{-\infty}^{\infty} g(y \mid \eta, \lambda)dy = \frac{1-\lambda}{2} + \frac{1+\lambda}{2} = 1$$

and $g(\cdot \mid \eta, \lambda)$ is a proper density.

Using the same set of transformations we find

$$\int_{-\infty}^{0} yg(y \mid \eta, \lambda)dy = (1-\lambda)^2\int_{-\infty}^{0} c\left(1 + \frac{x^2}{\eta-2}\right) dx = -c(1-\lambda)^2\left(\frac{\eta-2}{\eta-1}\right)$$

and

$$\int_{0}^{\infty} yg(y \mid \eta, \lambda)dy = (1+\lambda)^2\int_{0}^{\infty} c\left(1 + \frac{x^2}{\eta-2}\right) dx = c(1+\lambda)^2\left(\frac{\eta-2}{\eta-1}\right).$$

Thus

$$EY = \int_{-\infty}^{\infty} yg(y \mid \eta, \lambda)dy = c\left(\frac{\eta-2}{\eta-1}\right)\left[(1+\lambda)^2 - (1-\lambda)^2\right] = 4\lambda c\left(\frac{\eta-2}{\eta-1}\right) = a$$

(a is defined in equation (13)).

We also find that

$$\int_{-\infty}^{0} y^2 g(y \mid \eta, \lambda)dy = (1-\lambda)^3 \int_{-\infty}^{0} x^2 f(x \mid \eta)dx = \frac{(1-\lambda)^3}{2}$$

where the final inequality uses the fact that the density $f(x \mid \eta)$ is symmetric and has a variance of unity. Similarly,

$$\int_{0}^{\infty} y^2 g(y \mid \eta, \lambda)dy = \frac{(1+\lambda)^3}{2}.$$

Thus

$$EY^2 = \frac{(1-\lambda)^3}{2} + \frac{(1+\lambda)^3}{2} = 1 + 3\lambda^2 = b^2 + a^2$$

by definitions (13) and (14).

Now consider the random variable given by the transformation

$$Z = \frac{Y-a}{b}.$$

Its density is given by (12), which shows that this is a proper density. We can easily see that

$$EZ = \frac{EY-a}{b} = \frac{a-a}{b} = 0$$

and

$$EZ^2 = \frac{EY^2 - 2aEY + a^2}{b^2} = \frac{a^2 + b^2 - 2a^2 + a^2}{b^2} = 1,$$

which establishes that the density (12) has a mean of zero and unit variance, as desired.

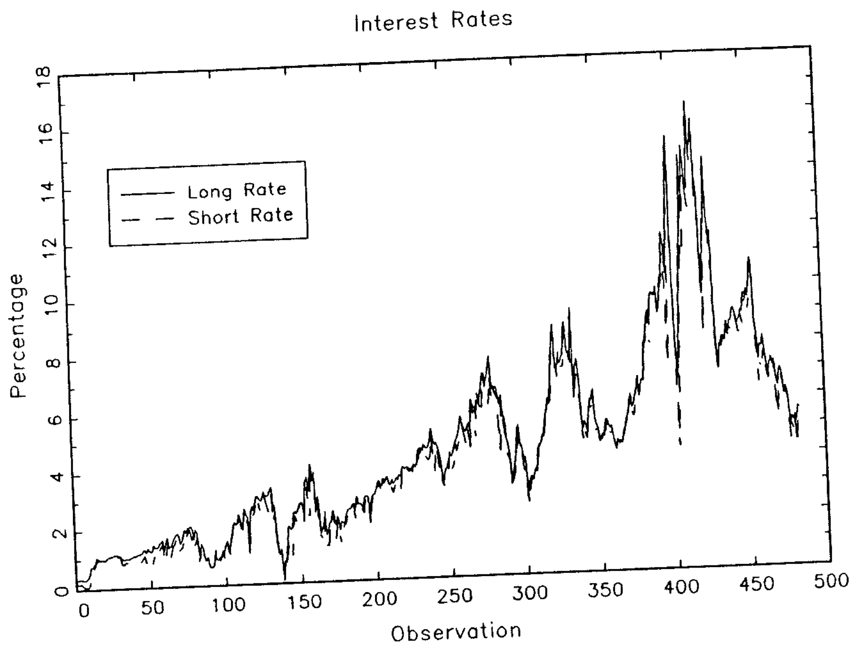

Figure 1

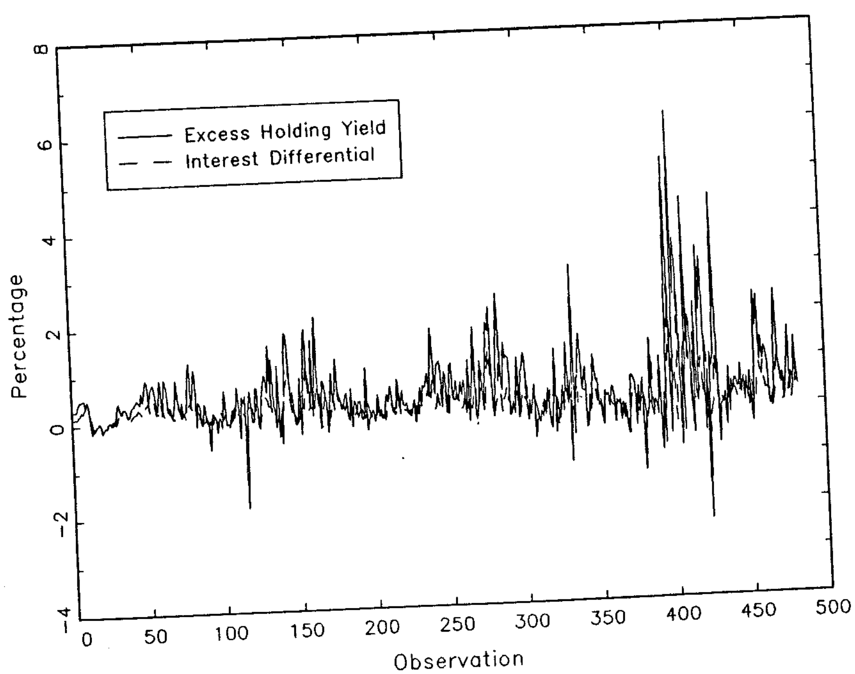

Figure 2

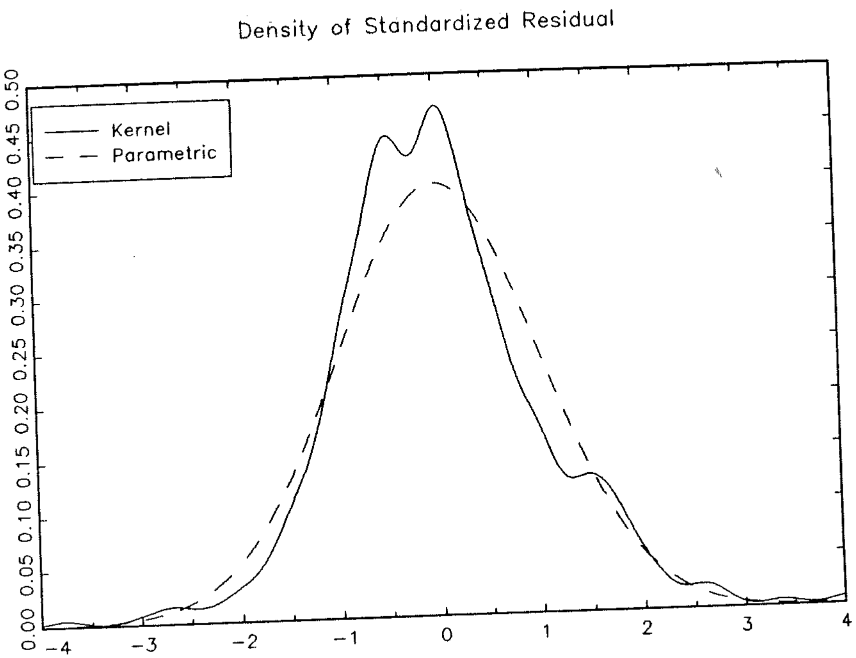

Figure 3

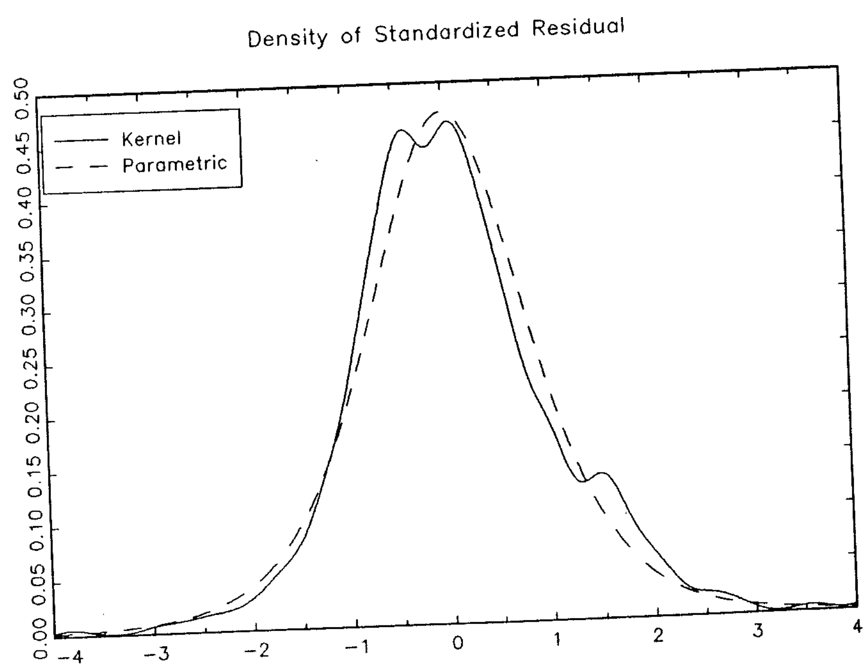

Figure 4

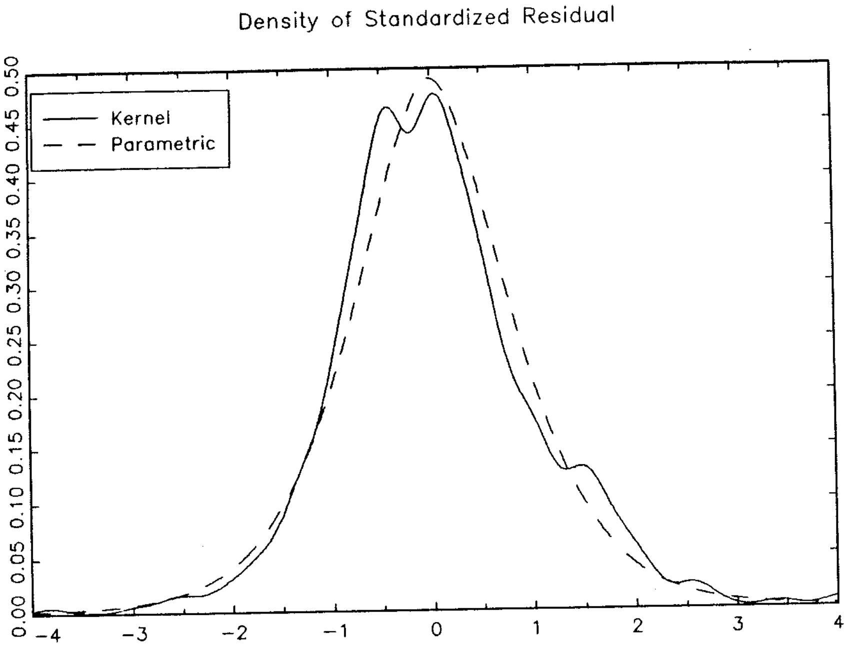

Figure 5

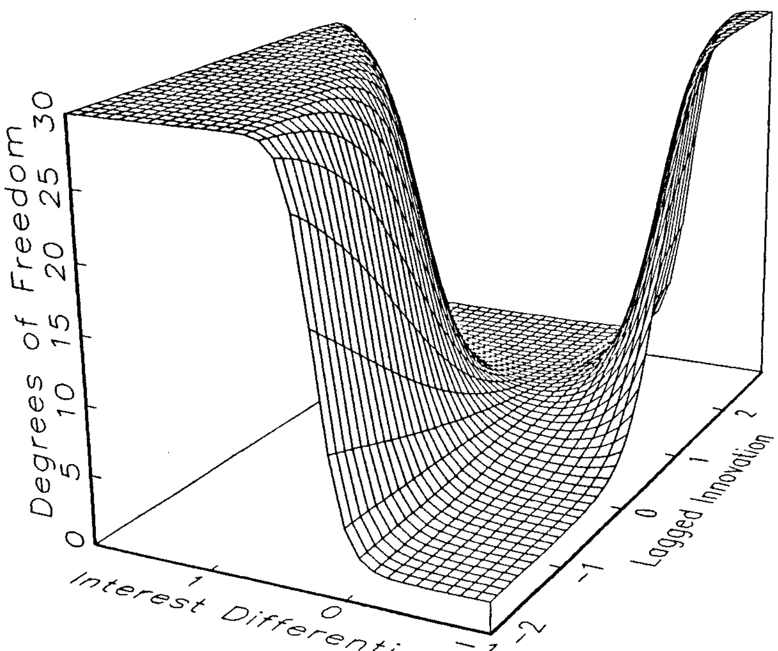

Figure 6

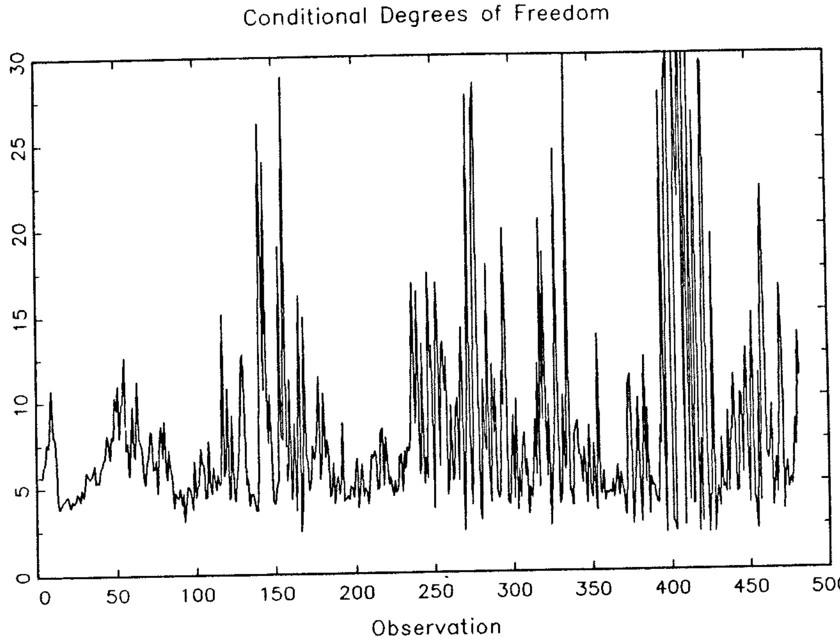

Figure 7

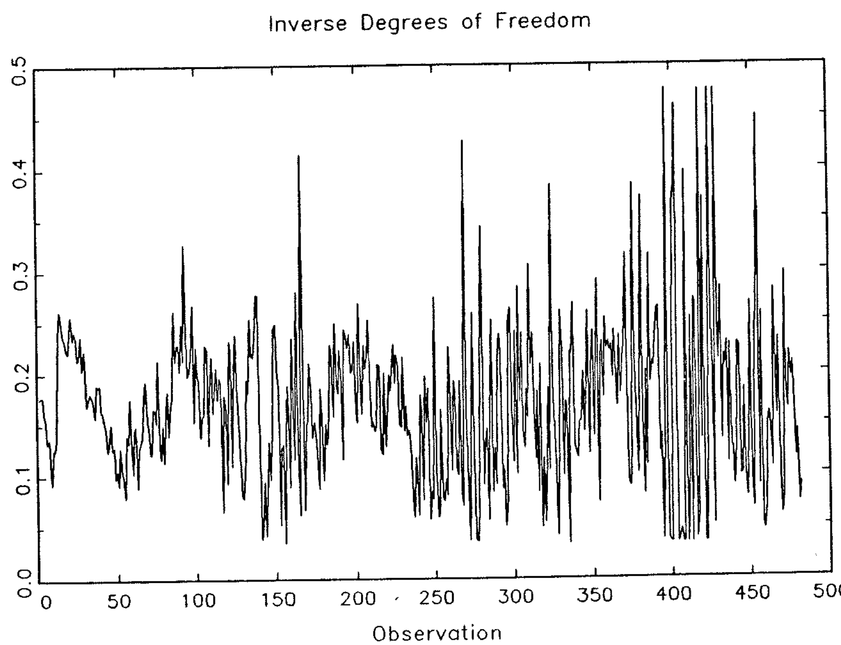

Figure 8

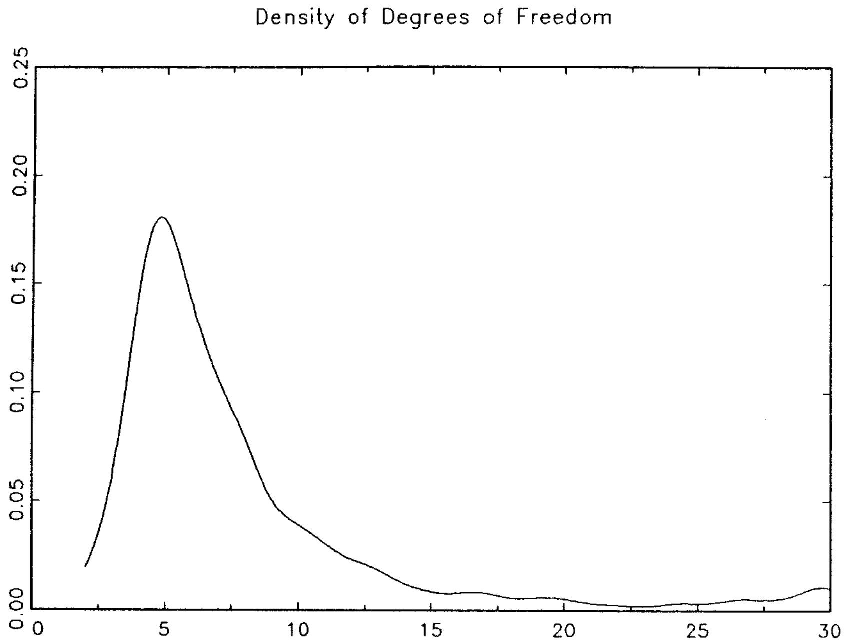

Figure 9

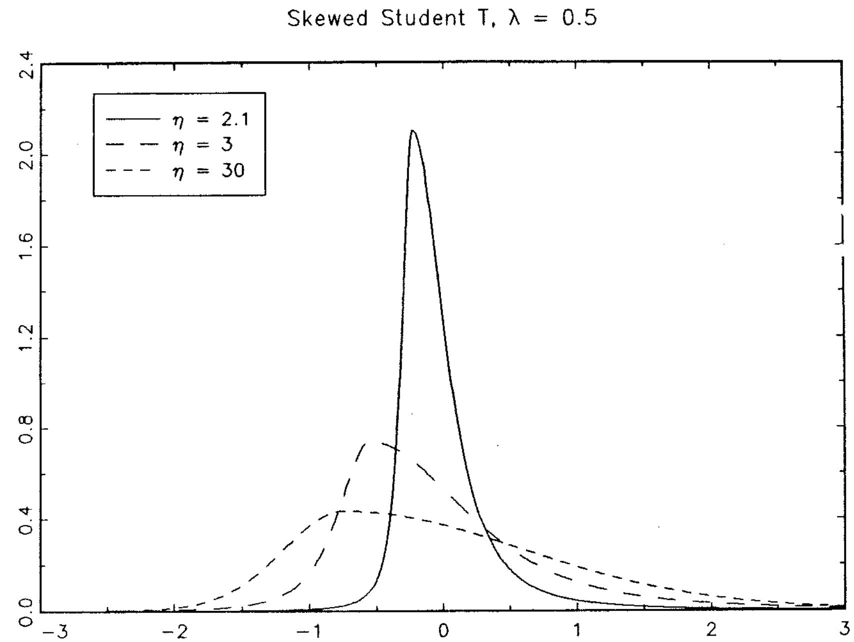

Figure 11

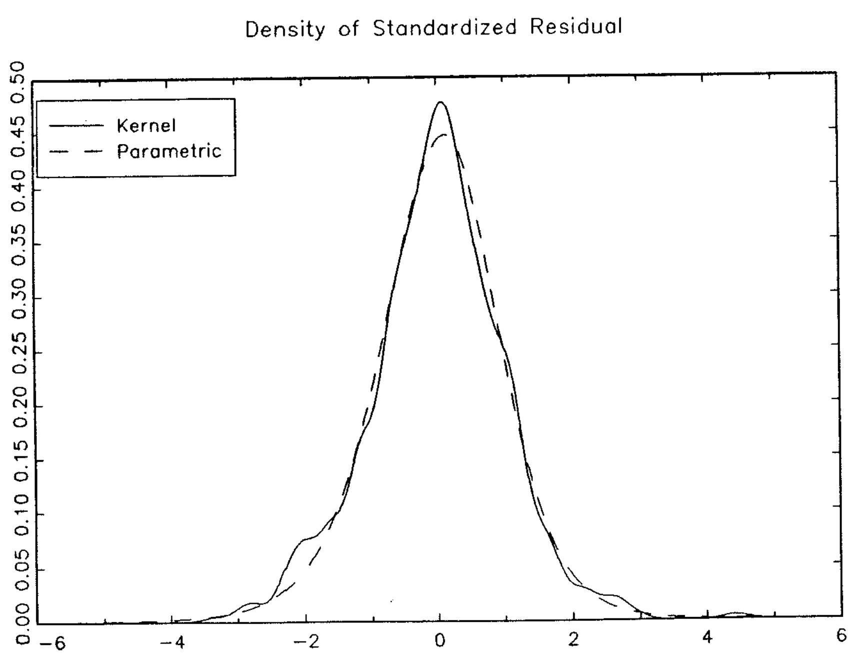

Figure 10

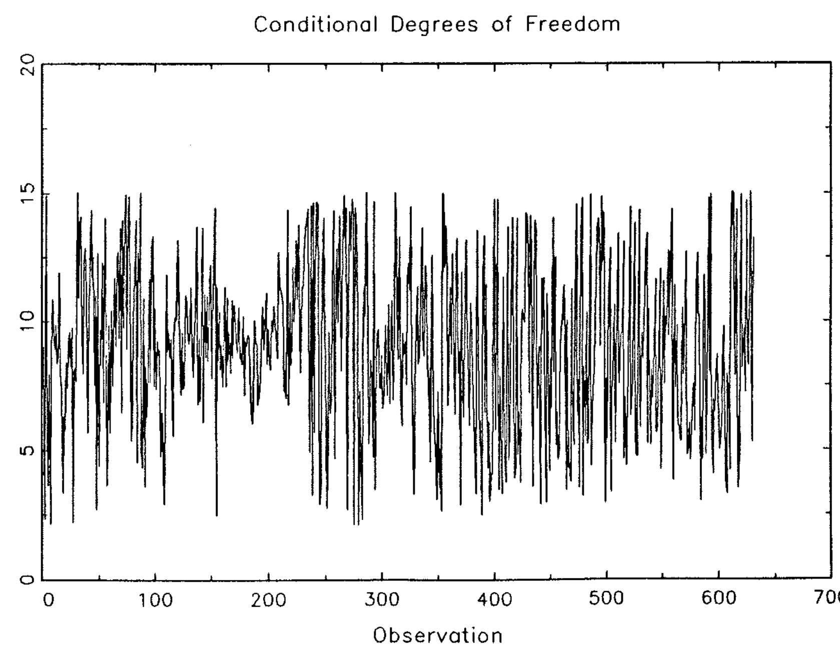

Figure 12

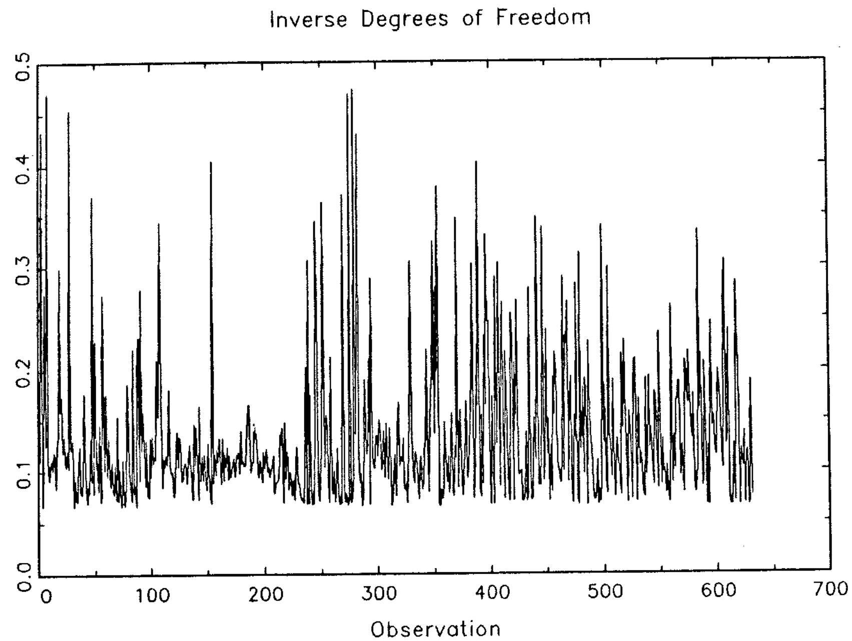

Figure 13

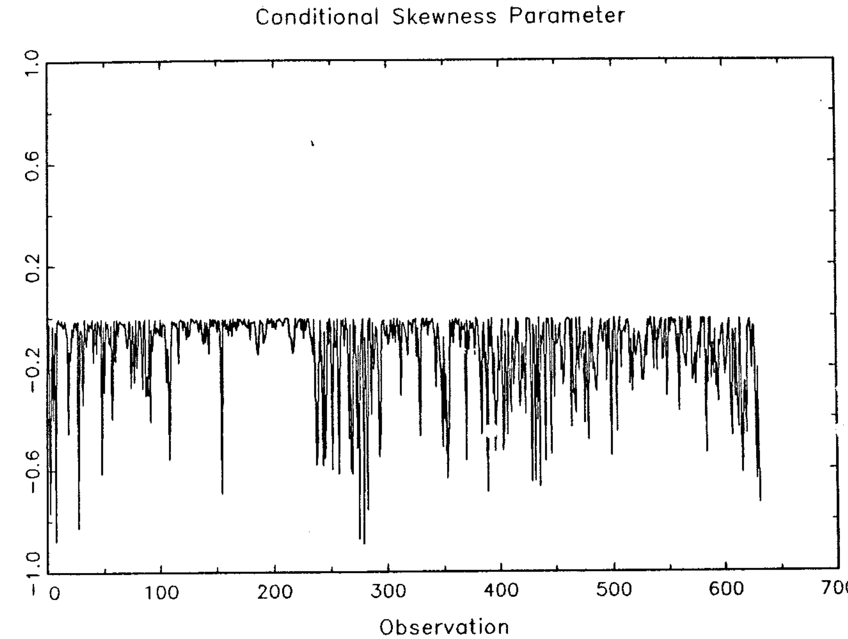

Figure 15

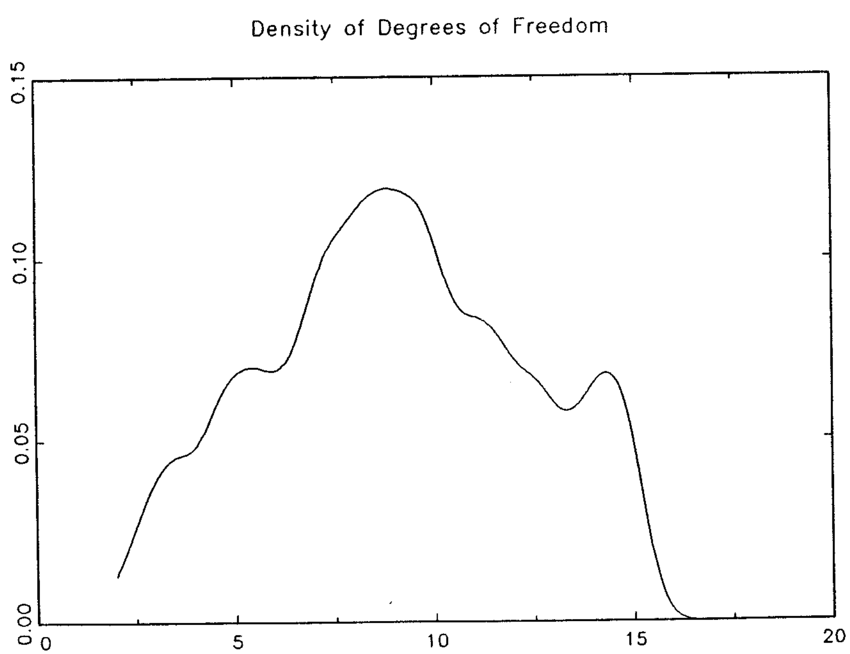

Figure 14

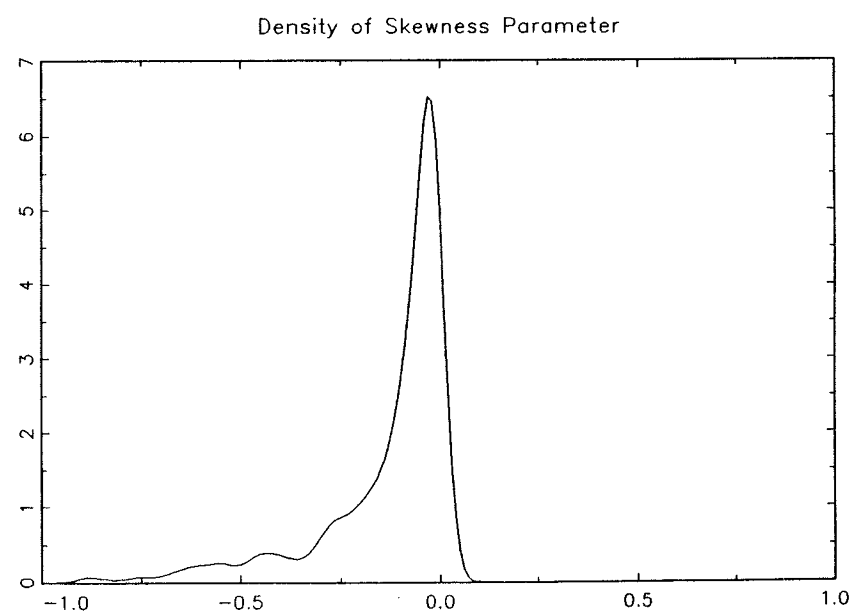

Figure 16