European University Institute

Max Weber Programme

EUI Working Papers

MWP 2007/32

Sample Kurtosis, GARCH-t and the Degrees of Freedom Issue

Maria S. Heracleous

EUROPEAN UNIVERSITY INSTITUTE
MAX WEBER PROGRAMME

Sample Kurtosis, GARCH-t and the Degrees of Freedom Issue

MARIA S. HERACLEOUS

EUI Working Paper MWP No. 2007/32

ISSN 1830-7728

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

Abstract

Econometric modeling based on the Student's t distribution introduces an additional parameter -- the degree of freedom. In this paper we use a simulation study to investigate the ability of (i) the GARCH-t model (Bollerslev, 1987) to estimate the true degree of freedom parameter and (ii) the sample kurtosis coefficient to accurately determine the implied degrees of freedom. Simulation results reveal that the GARCH-t model and the sample kurtosis coefficient provide biased and inconsistent estimates of the degree of freedom parameter. Moreover, by varying σ^2, we find that only the constant term in the conditional variance equation is affected, while the other parameters remain unaffected.

Keywords

Student's *t* distribution, Degree of freedom, Kurtosis coefficient, GARCH *t* model

JEL Classification: C15, C16, C22

Sample Kurtosis, GARCH-t and the Degrees of Freedom Issue*

Maria S. Heracleous†

1 Introduction

Most theoretical and applied research using the linear regression model assumes that the error term follows a Normal distribution. In many real world applications, however, there is substantial evidence showing that the distribution of errors has thicker tails than the Normal. One area where this distinction becomes particularly relevant is in modeling speculative price data where thick tails and volatility clustering are well documented features. In a seminal paper, Mandelbrot (1963) already pointed out these empirical regularities and proposed replacing the Normality assumption with the Pareto-Levy (Stable) family of distributions in an attempt to capture the leptokurticity and infinite variance in the distribution of returns. Fama (1965) also makes similar suggestions. Alternatives have also been proposed by Praetz (1972) and Blattberg and Gonedes (1974) among others, where continuous mixtures of Normal distributions leading to Student's t errors for modeling stock price indices are

*I would like to thank Anya McGuirk, Aris Spanos, Anindya Banerjee and Amos Golan for their comments and suggestions. I also thankfully acknowledge financial support from the Max Weber Program.

†Max Weber Fellow (2006-2007), EUI and Department of Economics, American University, Washington DC, 20016, USA. <*email*: heracleo@american.edu>

used.

By the end of the 1970s, however, it was largely recognized that existing volatility models based on the Pareto family were unable to account for the volatility clustering present in speculative price data. This gave rise to a new line of research which began with the introduction of the Autoregressive Conditional Heteroskedastic Model (ARCH) in a classic paper by Engle (1982). The original ARCH(p) model assumes a conditional error distribution that is Normal but expresses the conditional variance as a p^{th} order weighted average of past (squared) disturbances, and is thus able to explicitly capture volatility clustering in financial series. Following this, an enormous body of research has focused on extending and generalizing the ARCH model, primarily by suggesting alternative functional forms for the conditional variance.

In 1986 an important contribution to this literature occurred when Bollerslev proposed the Generalized ARCH (GARCH) model as a more parsimonious way to capture volatility dynamics. In order to better account for the observed leptokurtosis, Bollerslev (1987) further extended the GARCH specification to allow for the conditional Student's t distribution as an alternative to the Normal. In addition to these models, alternatives to the GARCH family utilizing the multivariate Student's t distribution have been proposed by Spanos (1994), McGuirk, Robertson and Spanos (1993) and Heracleous and Spanos (2005). These authors begin by assuming a multivariate Student's t distribution for the observables and impose appropriate probabilistic reduction assumptions to derive the conditional statistical model. Conditional model specifications obtained using this approach naturally accommodate static and dynamic heteroskedasticity as well as non-linear dependence.

Econometric modeling based on the Student's t distribution, however, introduces an additional parameter $-$ the degree of freedom parameter, which measures the extent of leptokurtosis in the data. One can also interpret this as a measure of the extent of departure from the Normal distribution. This in turn raises an estimation issue, since Zellner (1976) shows that there do not exist maximum likelihood estimates for the linear regression coefficients, the dispersion parameter and the degree of freedom parameter. Consequently, to use maximum likelihood, it is necessary to assign a degree of freedom parameter that reflects the distributional properties of the error term. One commonly proposed technique for selecting the degree of freedom is by using the kurtosis coefficient as a guide to solve for the implied degrees of freedom. Most studies using the Student's t distribution, however just assume that the degree of freedom parameter is known so that the standard maximum likelihood approach can be used. Bollerslev (1987) is an exception where the degree of freedom parameter as well as the GARCH parameters are all estimated by maximum likelihood methods.

In this paper we conduct a simulation study to examine two questions that naturally arise in the context of the Student's t distribution. The first relates to the reliability of the estimated degree of freedom parameter from Bollerslev's (1987) GARCH t model as well as its consequences. Secondly, we investigate the usefulness of using the sample kurtosis coefficient for determining the appropriate degree of freedom parameter for the Student's t distribution.

The plan of this paper is as follows. In the next section the Student's t distribution is introduced and the theoretical questions to be investigated are discussed. In section 3 we then provide a detailed description of the simulation set up. Section 4 presents the simulation results. The final section contains a summary of results and some concluding remarks.

2 Theoretical Background

We begin this section by defining and stating some of the properties of this distribution. Let $\mathbf{y} = (y_1, y_2, \ldots y_n)$ be a $n \times 1$ vector which has a multivariate Student's t distribution denoted by $\mathbf{y} \sim St_n\left(\nu, 0, \sigma^2\mathbf{I}\right)$. Its probability density function is given by:

$$f(\mathbf{y}/\nu, \sigma^2) = \frac{\Gamma\left[\frac{1}{2}\left(\nu + n\right)\right]}{(\pi\nu)^{\frac{n}{2}}\,\sigma^{\frac{n}{2}}\Gamma\left[\frac{1}{2}\nu\right]}\left[1 + \frac{\mathbf{y}'\mathbf{y}}{\nu\sigma^2}\right]^{-\frac{1}{2}(\nu+n)} \tag{1}$$

where ν and σ^2 are respectively the degree of freedom parameter and the dispersion parameter. The degree of freedom parameter is also referred to as the shape parameter because the peakedness of the density function and thickness of the tails in equation(1) depend on its value. It is well known that as $\nu \longrightarrow \infty$ the t-distribution approaches the Normal but for small values of ν the t-distribution is more sharply peaked and has thicker tails than the Normal. The student's t distribution is symmetric around 0 has the following properties:

$$E(y_i) = \mathbf{0}, \qquad E(y_i^2) = \frac{\nu}{\nu - 2}\sigma^2, \; and \qquad E(y_i^4) = \frac{3\nu^2}{(\nu - 2)\,(\nu - 4)}\sigma^4, \tag{2}$$

where $\nu > 4$.

As already mentioned, financial data such as stock prices, interest rates and exchange rates seem to have a distribution which is much closer to the Student's t. Thus, an important decision to be made when using the Student's t distribution for modeling such data is to choose the appropriate value of the degree of freedom parameter. Some authors have used density estimates (Spanos 1994), standardized Student's t P-P plots (Heracleous and Spanos 2005) as well as the kurtosis coefficient as guides for the an initial value of ν. In these papers however the final choice of

ν is made on statistical adequacy grounds. By contrast, Bollerslev (1987) does not prespecify the degree of freedom but treats it as a parameter to be estimated. This will be discussed in detail in the following section.

2.1 Bollerslev's GARCH - t model

In view of the fact that the Gaussian GARCH model could not explain the leptokurtosis exhibited by asset returns, Bollerslev (1987) suggested replacing the assumption of conditional Normality of the error with the conditional Student's t distribution. He argued that this formulation would permit us to distinguish between conditional leptokurtosis and conditional heteroskedasticity as plausible causes of the unconditional kurtosis observed in the data. This model can be specified in terms of its first two conditional moments. The conditional mean is constant as follows:

$$y_t = \mu + u_t \qquad\qquad u_t/\mathcal{F}_{t-1} \sim St_\nu(0, h_t^2) \tag{3}$$

where $\mathcal{F}_{t-1}$ represents the past history of the dependent variable. The GARCH (p,q) conditional variance, h_t^2 for this model takes the form:

$$h_t^2 = \omega + \sum_{i=1}^{p} a_i u_{t-i}^2 + \sum_{j=1}^{q} \gamma_j h_{t-j}^2, \qquad p \geq 1, \qquad q \geq 1 \tag{4}$$

where the parameter restrictions $\omega > 0$, $a_i \geq 0$, $\gamma_j \geq 0$ ensure that the conditional variance is always positive. Moreover, $\sum_{i=1}^{p} a_i + \sum_{j=1}^{q} \gamma_j < 1$ is required for the convergence of the conditional variance. The distribution of the error term according to Bollerslev (1987) takes the form:

$$f(u_t/Y_{t-1}^p; \theta_1) = \frac{\Gamma\left[\frac{1}{2}(\nu+1)\right]}{\pi^{\frac{1}{2}}\Gamma\left[\frac{1}{2}\nu\right]} \left[(\nu-2)\,h_t^2\right]^{-\frac{1}{2}} \left[1 + \frac{u_t^2}{(\nu-2)\,h_t^2}\right]^{-\frac{1}{2}(\nu+1)} \tag{5}$$

However, McGuirk et al. (1993) argue that the above distribution in equation (5) can be obtained by substituting the conditional variance h_t^2 in the functional form of the *marginal* Student's t distribution and re-arranging the dispersion (scale) parameter. This would be indeed the correct strategy for the Normal distribution, where one does not have to be concerned about the degree of freedom parameter. For the Student's t distribution however, the degree of freedom in the conditional distribution change depending on the number of the conditioning variables. In fact McGuirk et al. (1993) further show that if we derive the conditional distribution from the joint distribution of the observables $f(y_t, y_{t-1,...}y_{t-p}; \psi)$ we get the expression shown in equation (6) below:

$$f(u_t/Y_{t-1}^p; \theta_2) = \frac{\Gamma\left[\frac{1}{2}(\nu + p + 1)\right]}{\pi^{\frac{1}{2}}\Gamma\left[\frac{1}{2}\nu + p\right]} \left[\nu\sigma^2 h_t^2\right]^{-\frac{1}{2}} \left[1 + \frac{u_t^2}{\nu\sigma^2 h_t^2}\right]^{-\frac{1}{2}(\nu+p+1)} \tag{6}$$

Note that the parameter in the gamma function in the two equations are different. More importantly however, we observe that σ^2 does not appear in equation (5). This suggests that estimation of the degree of freedom parameter, ν from the GARCH- t model à la Bollerslev (1987), that ignores σ^2 will give an incorrect mixture of both ν and σ^2. To investigate this issue further we allow σ^2 to vary and examine its effect on all estimated parameters of the GARCH- t model.

2.2 Sample Kurtosis Coefficient

The sample kurtosis coefficient introduced by Pearson (1895) measures of the peakedness in relation to the tails of the distribution. It is defined as the standardized fourth central moment as follows: $\alpha_4 = \frac{\mu_4}{(\mu_2)^2}$. In the case of the Student's t distribution the sample kurtosis coefficient is related to the degree of freedom parameter in the following way: $\alpha_4 = 3+\frac{6}{\nu-4}$. This provides one way of choosing the degrees of freedom parameter ν. The Normal distribution, with a value of $\alpha_4 = 3$, is often used as a benchmark. Distributions with $\alpha_4 > 3$ are called leptokurtic and have a sharper peak and fatter tails than the Normal. Typically, for financial data the kurtosis coefficient is well above 3 indicating possible non-Normality.

Even though the sample kurtosis coefficient is widely used in quantitative finance, it has been criticized in the statistics literature as a "vague concept"(Mosteller and Tuckey, 1977). Also Ballanda and MacGillivray (1988) point out that "although moments play an important role in statistical inference they are very poor indicators of distributional shape". Its usefulness has also been questioned since it is based on sample averages, which are sensitive to outliers. This effect can be amplified easily since they are raised to the fourth power. Fisher pointed out the weakness of using moments beyond the second as early as 1922. Given these criticisms it is not apparent that the kurtosis coefficient is a useful way to compute the implied degrees of freedom. Hence we use simulations to investigate the sampling distribution of the kurtosis coefficient and the sampling distribution of the implied degrees of freedom parameter.

3 Simulation Set Up

Data for this simulation were generated in the following way. First, a raw series of Student's t random numbers with mean 0 and variance 1 is generated. The degree of freedom parameter was allowed to vary in different simulations according to the needs of the study. The raw Student's t numbers were generated using the algorithm proposed by Dagpunar (1988). This algorithm uses numbers from a uniform distribution as an input, allowing the user to control the seed. The following procedure was used in order to enable easy reproduction of the data from any given run. The initial seed was set to $2^{11} - 1$ and series of Student's t numbers were generated. The generated data were tested for skewness and kurtosis using tolerance levels specified in Paczkowski (1997). Only samples that met the standards were used in the simulation study. Maximum allowable skewness was set to ± 0.1. The tolerance for kurtosis was set to ± 0.5 around the value implied by the degree of freedom parameter according to the relationship shown below:

$$\alpha_4 = 3 + \frac{6}{\nu - 4}; \nu > 4 \tag{7}$$

Note that for $\nu = 4$ the above relationship breaks down and the allowable kurtosis range was set to 7.5–8.5.

The true structure of the data is given by:

$$\mathbf{Y}_t^1 \sim \mathbf{S}t_t\left(1_t\mu, \mathbf{\Sigma}_t^1, \nu\right), \qquad t = 1, \ldots, T,$$

where $\mathbf{Y}_t^1 = (Y_1, Y_2, \ldots, Y_t)'$, $\mu = 0$, and $(\mathbf{\Sigma}_t^1)^{-1}$ is

$$
(\Sigma_t^1)^{-1} = \begin{bmatrix}
1 & -0.4 & -0.2 & & & & & & & \\
-0.4 & 1.16 & -0.32 & -0.2 & & & & & & \\
-0.2 & -0.32 & 1.20 & -0.32 & -0.2 & & & & & \\
& -0.2 & -0.32 & 1.20 & -0.32 & -0.2 & & & & \\
& & . & . & . & . & & & & \\
& & & & . & . & . & . & & \\
& & & & -0.2 & -0.32 & 1.20 & -0.32 & -0.20 \\
& & & & & -0.2 & -0.32 & 1.16 & -0.4 \\
& & & & & & -0.2 & -0.4 & 1
\end{bmatrix} \tag{8}
$$

Using $\mu = 0$ and equation (8) the remaining parameters can be computed as follows:

$$
\boldsymbol{\beta} = \begin{pmatrix} \beta_2 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.4 \end{pmatrix}, \qquad \sigma^2 = 1, \qquad \beta_0 = 0.
$$

Next, we use the Cholesky factorization as shown below to impose the necessary dependence structure on the generated data.

$$
\mathbf{Y}_T^1 = chol\left(\frac{\nu}{\nu - 2}\boldsymbol{\Sigma}_T^1\right)\mathbf{J} \tag{9}
$$

where $\mathbf{J}$ is the series of raw Student's t numbers and $\boldsymbol{\Sigma}_T^1$ is the scale matrix, and $chol(.)$ refers to the Cholesky factorization. Using this method we generate the joint

Student's t distribution directly.

To study the questions of interest mentioned in the previous section we consider a number of different scenarios. Sample sizes (n) of 50, 100, 500 and 1000 were chosen since financial data series can be available annually, monthly, weekly or daily. Three different degrees of freedom $\nu = 4$, 6, and 8 were used in this study. The true value of α_4 is 6 when $\nu = 6$ and 4.5 when $\nu = 8$. For $\nu = 4$, α_4 is undefined. Most financial data are leptokurtic and thus can be described by degrees of freedom in the range of $4-8$. Moreover, as ν increases above 8 the distribution looks much closer to the Normal and it would not be pertinent to consider larger degrees of freedom. For each combination of the sample size and ν, 1000 data sets were generated. Also for sample size 500 we allowed σ^2 to vary. It was allowed to take the values 1, 0.25 and 4. In each of these instances the value of σ^2 affects only the scale matrix $\mathbf{\Sigma}_t^1$, leaving the other parameters unchanged. For $\sigma^2 = 1, (\mathbf{\Sigma}_t^1)^{-1}$ takes the values given in equation (8). In general for $\sigma^2 = k$, $k > 0$, the inverse of the scale matrix in equation (8) is multiplied by the factor of $\frac{1}{k}$.[1]

4 Results

In this section we present the results of the simulation study. In the first part we present the results for the sample kurtosis coefficient α_4 and the implied degrees of freedom parameter (inu) computed using α_4. In the second part we focus on the degrees of freedom parameter as estimated in Bollerslev's (1987) model. By allowing σ^2 to take different values we examine its impact on all estimated parameters.

[1]The computer code for this simulation exercise was written in GAUSS 4.0 and the Student's t GARCH model was estimated using the FANPAC toolbox which is part of the GAUSS package.

Parameter	DF	Sample size	Mean	Empirical s.e	Skewness	Kurtosis
$\hat{\alpha}_4$	$\nu = 4$	50	5.2981	.9064	.4637	3.2697
		100	5.5234	.9170	.3247	3.0346
		500	5.6626	.6454	.6828	4.3656
		1000	5.6577	.5183	.6625	4.4543
	$\nu = 6$	50	4.3069	.7529	.5666	3.7797
		100	4.4364	.7100	.5382	3.7072
		500	4.5507	.4824	.5406	3.6502
		1000	4.5317	.3960	.5921	3.3306
	$\nu = 8$	50	3.6416	.6341	.4985	3.2237
		100	3.7211	.5630	.6295	3.7385
		500	3.7202	.3554	.7278	4.0377
		1000	3.7350	.2914	.7100	4.0557
inu	$\nu = 4$	50	6.7046	10.0026	-20.6339	458.6991
		100	6.8047	2.0034	2.8237	95.6585
		500	6.3920	.6194	1.3633	7.1254
		1000	6.3441	.4664	.9253	5.5797
	$\nu = 6$	50	10.3436	21.9843	-.2979	74.1571
		100	14.6619	145.1031	30.0049	929.4971
		500	8.3120	1.6467	2.4282	14.7343
		1000	8.1886	1.1193	0.9077	4.2954
	$\nu = 8$	50	-1.4269	314.2906	-22.8729	631.6351
		100	992.27.62	30878	31.5739	997.9412
		500	15.5592	26.7089	6.0128	316.9994
		1000	13.1225	25.5438	-28.1416	864.1377

Table 1: Descriptive Statistics for Sample Kurtosis and inu

4.1 Estimates of α_4 and the Implied Degrees of Freedom

In Table 1 we report descriptive statistics for the empirical distribution of the estimates. We observe that the sample kurtosis coefficient, α_4, is relatively stable around 5.6 for $\nu = 4$. Similarly it is stable around 4.5 for $\nu = 6$, and around 3.7 for $\nu = 8$. Also note that as ν increases, α_4 decreases as expected. However, once we use α_4 to derive the implied degrees of freedom (inu), we find some interesting features. In all three cases ($\nu = 4, 6, 8$) the implied degrees of freedom consistently exceed the true ones. For $\nu = 4$, inu is around 6 (starting at 6.7 for a sample size of 50 and going down to 6.3 for sample size of 1000). Also note that the standard error decreases dramatically from a value of 10 ($n = 50$) to a value of 0.46 ($n = 1000$). For $\nu = 6$, at sample size 50, $inu = 10$, but stabilizes around 8 for sample sizes of 500 and 1000. Interestingly, for sample size 100, it jumps to 14 and the standard error is high as well. For $\nu = 8$ we find evidence of erratic behavior especially for sample sizes of 50 and 100 where $inu = -1.4$ and $inu = 992$ respectively.

As the sample size increases the implied degrees of freedom parameter decreases to 13 for $n = 1000$ and possibly could decrease even more for sample sizes beyond 1000. This erratic behavior at small sample sizes can largely be explained by the fact that a few bad draws can affect the (mean) estimates for small sample sizes. Also as mentioned before, for $\nu = 6$ and $\nu = 8$, we find that α_4 is around 4 and 3 respectively. Recall that $inu = 4+\frac{6}{\alpha_4-3}$. Hence, for $\nu = 6$ and $\nu = 8$, there is a higher probability of getting unusually large values for inu. This can explain the erratic behavior observed above. Overall, the results suggest that the sample kurtosis coefficient is not a good measure of the true degrees of freedom. This result is not surprising. In fact, Wang and Ip (2003) have shown theoretically that the moment estimate of the degree of freedom parameter of the multivariate Student's t distribution for the disturbance in the linear regression model is inconsistent.

4.2 Estimates of the GARCH- t parameters

Table 2 provides descriptive statistics for the empirical distribution of the degree of freedom parameter (nu), estimated by Bollerslev's (1987) model. The results suggest that the GARCH- t model consistently overestimates the true degree of freedom parameter. For instance, when $\nu = 4$, the estimated value is around 8. Also note

Parameter	DF	Sample size	Mean	Empirical s.e	Skewness	Kurtosis
nu	$\nu = 4$	50	7.8161	12.4646	5.2276	31.7047
		100	8.2640	10.9403	5.3855	35.1996
		500	7.9918	2.7387	3.0789	22.0161
		1000	8.1036	1.8096	1.8848	9.4895
	$\nu = 6$	50	13.0683	20.7030	3.0020	10.9406
		100	12.2761	16.1543	3.5597	15.7681
		500	10.9033	4.1220	1.9758	8.9713
		1000	11.0866	2.8603	1.2313	5.0127
	$\nu = 8$	50	20.7026	27.1638	1.8269	4.9891
		100	20.3718	24.1092	2.0239	6.0065
		500	17.7961	10.3940	2.9042	16.7992
		1000	15.4479	4.5214	1.9255	13.8620

Table 2: Descriptive Statistics for Estimated nu

that for small sample sizes the empirical standard error of the parameters is larger than their estimated value. Even if sample size increases the standard error is quite large giving rise to imprecise estimates. For $\nu = 4$ estimated nu is stable around the value of 8. For $\nu = 6$ and $\nu = 8$, as the sample size increases, we observe a downward trend towards the true value of ν. However, the final estimates are never close enough to the true value.

Next we present Table 3 which shows how the estimates of the conditional variance from the GARCH- t model vary for different values of σ^2. We chose to use a sample size of 500 in this case to avoid any problems with small sample sizes. Recall that σ^2 is not a free parameter that can be estimated in Bollerslev's (1987) formulation (see equation 5). Therefore we can interpret this as the $\sigma^2 = 1$ case that serves as the benchmark for our simulation.

Interestingly, we find that the only parameter which varies with σ^2 is ω $-$ the constant term in the conditional variance equation. Moreover, we observe that there is a proportional relationship between $\sigma^2$and ω. For example when $\nu = 4$ and $\sigma^2 = 4$ the estimate of ω is 4.5016 which is roughly four times the value of ω when $\sigma^2 = 1$. This relationship holds for $\nu = 6$ and 8 (see Table 4). We can also see from Table 3 that the estimated degree of freedom (nu), and the GARCH and ARCH parameters in the conditional variance remain unchanged as σ^2 varies. Thus, these results suggest that the effect of the σ^2 (the missing parameter) in Bollerslev's formulation is fully absorbed by the constant in the conditional variance equation.

We also present Normal kernel density estimates of the empirical distribution of nu for various sample sizes and $\nu = 6$ in figures 1-6. The dashed lines in these figures represent the contour of the Normal density, with the same mean and variance as the data whose distribution is shown in the graph. Figures 1-4 show the kernel density for $\nu = 6$, $\sigma^2 = 1$ and sample sizes $n = 50, 100, 500$ and 1000 respectively. In figures 5-6 we let σ^2 take the values of 4 and 0.25 while $\nu = 6$, and $n = 500$. It is easy to see from figures $1 - 6$ that the distribution is leptokurtic, skewed to the left and the mode and the mean are far from the true value ($\nu = 6$). Graphs for $\nu = 4$ and $\nu = 8$ exhibit similar patterns and hence have been omitted.

Parameter	DF	σ^2	Mean	Empirical s.e	Skewness	Kurtosis
		1	7.9918	2.7387	3.0789	22.0161
nu	$\nu = 4$	0.25	7.8868	2.5311	3.0377	24.2803
		4	7.8574	2.5787	3.2067	27.3063
		1	1.1164	.3846	.4161	3.1568
ω	$\nu = 4$	0.25	.2778	.0959	.3616	3.0933
		4	4.5016	1.5929	.4569	3.1791
		1	.2226	.2074	-.0394	2.8308
γ_1	$\nu = 4$	0.25	.2264	.2072	.0119	2.8227
		4	.2196	.2121	-.0699	2.8588
		1	.2229	.0642	.3804	3.2621
a_1	$\nu = 4$	0.25	.2219	.0642	.3653	3.1943
		4	.2219	.0641	.3451	3.0918

Note:
1. γ_1 is the GARCH parameter, a_1 is the ARCH parameter and ω is the constant in the Student's t GARCH model (Bollerslev, 1987)

Table 3: Descriptive Statistics for the Student's t GARCH parametes, *n=500*

Parameter	DF	σ^2	Mean	Empirical s.e	Skewness	Kurtosis
		1	1.1164	.3846	.4161	3.1568
ω	$\nu = 4$	0.25	.2778	.0959	.3616	3.0933
		4	4.5016	1.5929	.4569	3.1791
		1	.9111	.4027	.6434	4.0782
ω	$\nu = 6$	0.25	.2246	.0987	.5478	3.6875
		4	3.6325	.1896	.5383	3.7684
		1	.9728	.5244	.7333	3.2792
ω	$\nu = 8$	0.25	.2454	.1310	.7034	3.2997
		4	3.8037	2.0297	.6705	3.2348

Table 4: Descriptive Statistics for the constant in the Student's t GARCH model, *n=500*

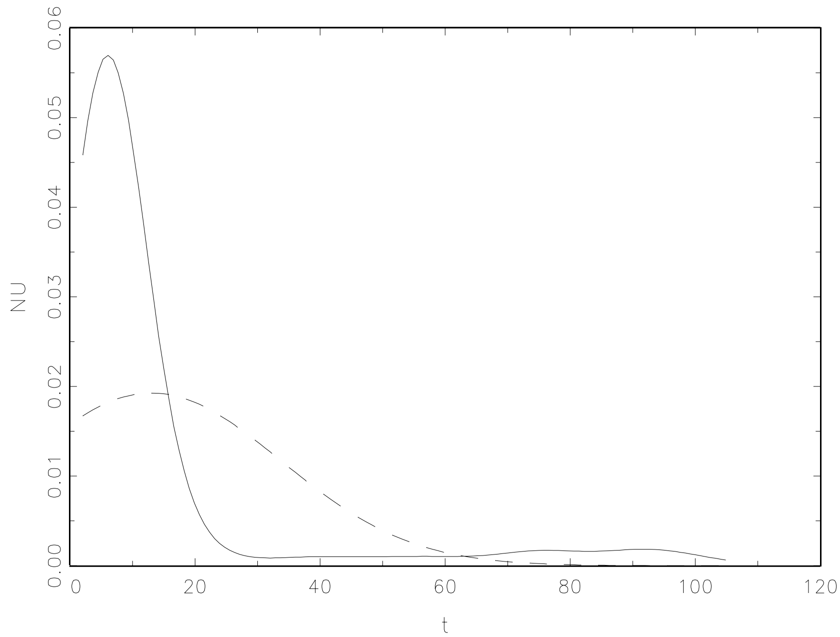

Figure 1: Kernel density for nu, $\nu = 6$, $\sigma^2 = 1, n = 50$

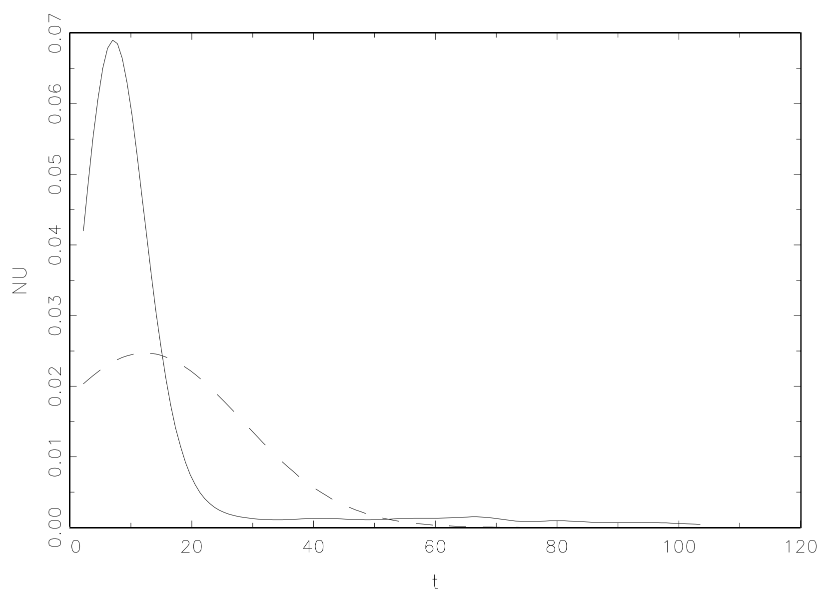

Figure 2: Kernel density for nu, $\nu = 6$, $\sigma^2 = 1, n = 100$

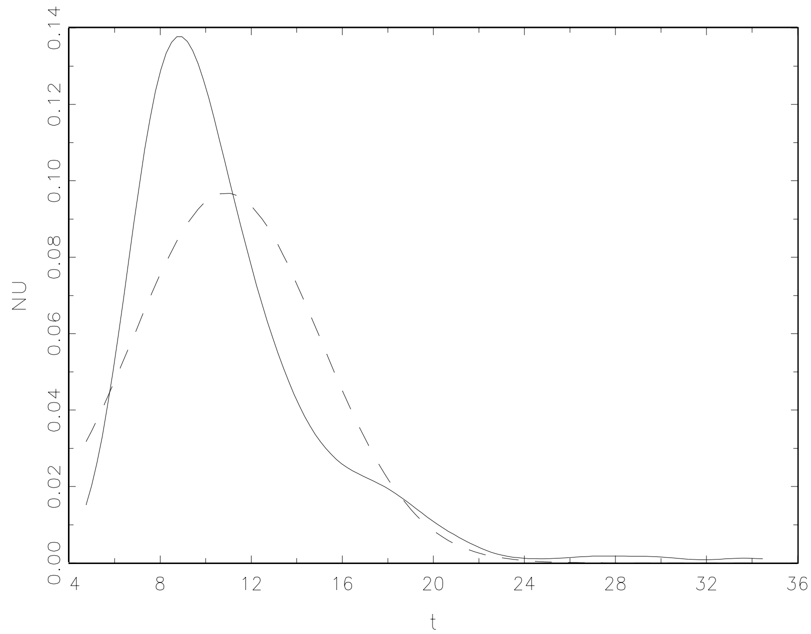

Figure 3: Kernel density for nu, $\nu = 6$, $\sigma^2 = 1, n = 500$

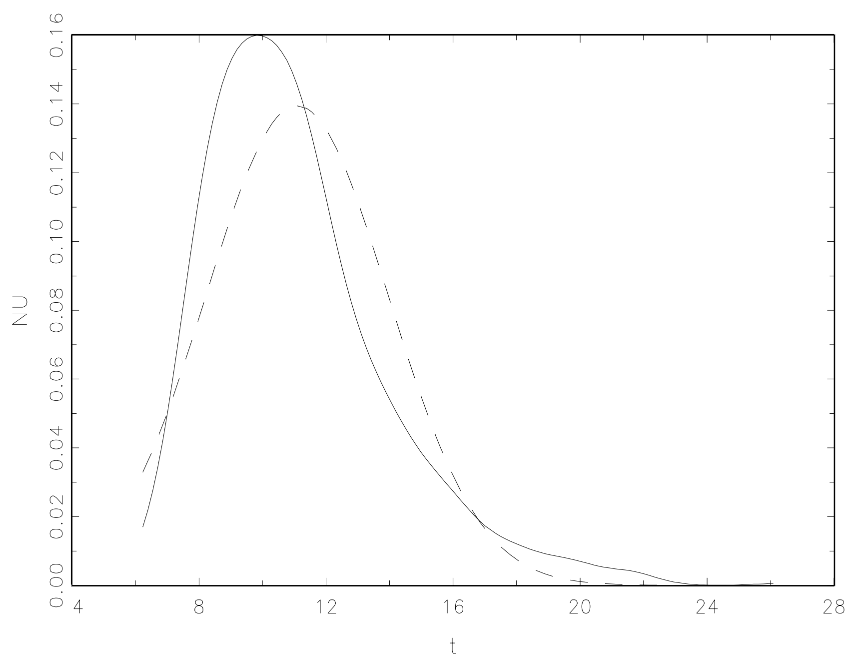

Figure 4: Kernel density for nu, $\nu = 6$, $\sigma^2 = 1, n = 1000$

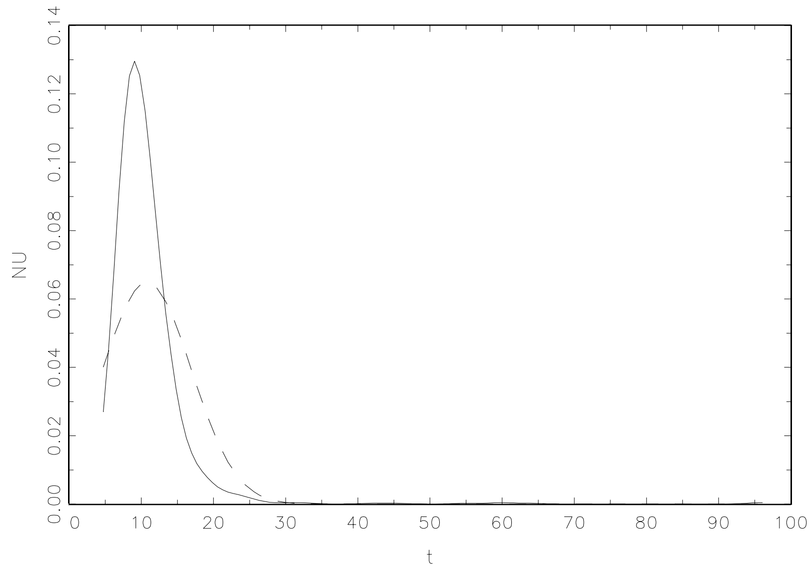

Figure 5: Kernel density for nu, $\nu = 6$, $\sigma^2 = 0.25$, $n = 500$

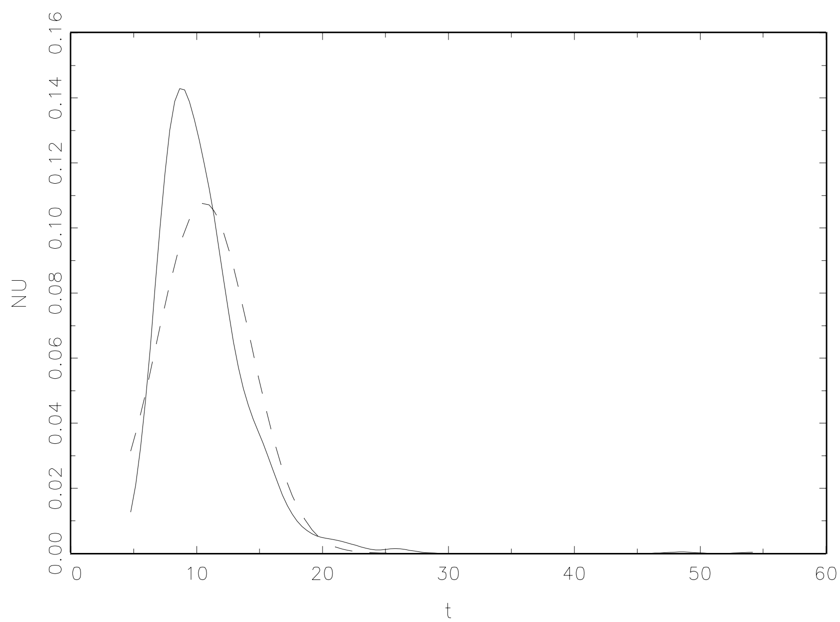

Figure 6: Kernel density for nu, $\nu = 6$, $\sigma^2 = 4$, $n = 500$

5 Conclusion

The Student's t distribution provides a useful alternative for modeling financial data. Relative to the Normal distribution, it has an additional variable – the degree of freedom parameter for capturing the leptokurtosis in the data. However, it also raises an additional question – how do we ascertain the appropriate degrees of freedom? In this paper we evaluate this question in the context of the sample kurtosis coefficient which is often used to determine the implied degree of freedom, and for Bollerslev's GARCH- t model (1987) where the degree of freedom parameter is also estimated.

Our simulation results reveal that the sample kurtosis coefficient provides a biased and inconsistent estimator of the degree of freedom parameter. Our simulations further show that the GARCH- t model also provides biased and inconsistent estimates. This is mitigated by the fact that the conditional mean parameters, the ARCH and GARCH coefficients as well as the degree of freedom parameter in the Student t GARCH model are not affected when the dispersion parameter, σ^2, is allowed to vary. However, we do find that the constant term in the conditional variance equation is affected when σ^2 varies. This will certainly have consequences for estimating and predicting volatility. Finally note that there seems to be a proportional relationship between the change in σ^2 and the effect on the constant term in the volatility equation.

References

[1] Aptech Systems (2002). GAUSS 4.0 Manual, Maple Valley, Aptech Systems Inc.

[2] Balanda, K.P. and H.L. MacGillivray (1988). "Kurtosis: A Crirical Review", *The American Statistician*, **42**, 111-119.

[3] Blattberg, R.C., N.J. Gonedes (1974). "A Comparison of the Stable and the Student Distributions as Statistical Models for Stock Prices", *Journal Of Business*, **47**, 244-280.

[4] Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, **31**, 307-327.

[5] Bollerslev, T. (1987). "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics*, **69**, 542-547.

[6] Dagpunar, J. (1988). *Principles of Random Variate Generation*, Oxford: Oxford University Press.

[7] Engle, R.F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, **50**, 987-1008.

[8] Engle, R.F. and T. Bollerslev (1986). "Modeling the Persistence of Conditional Variances", *Econometric Reviews*, **5**, 1-50 (with discussion).

[9] Fama, E.F. (1965). "The Behaviour of Stock-Market Prices", *Journal of Business*, **38**, 34-105.

[10] Fisher, R.A. (1922). "On the Mathematical Foundations of Theoretical Statistics", *Philosophical Transactions of the Royal Society A*, **222**, 309-368.

[11] Heracleous, M.S. and A. Spanos (2005). "The Student's t Dynamic Linear Regression: Re-examining Volatility Modeling". Advances in Econometrics, **20A**, 289-319.

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

[13] McGuirk, A., J. Robertson and A. Spanos (1993). "Modeling Exchange Rate Dynamics: Non-Linear Dependence and Thick Tails", *Econometric Reviews*, **12**, 33-63.

[14] Mosteller, F. and J.W. Tuckey (1977). *Data Analysis and Regression.* Reading, MA: Addison - Wesley.

[15] Paczowski, R. (1997). *Monte Carlo Examination of Static and Dynamic Student's t Regression Models*, Ph.D Dissertation, Virginia Polytechnic Institute and State University.

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

[17] Praetz, P.D. (1972). "The Distribution of Share Price Changes", *Journal of Business*, **45**, 49-55.

[18] Spanos, A. (1994). "On Modeling Heteroskedasticity: The Student's t and Elliptical Linear Regression Models", *Econometric Theory*, **10**, 286-315.

[19] Spanos, A. (1999). *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*, Cambridge: Cambridge University Press.

[20] Wang, S. and W. Ip (2003). "Inconsistency of estimate of the degree of freedom of multivariate student-t disturbances in linear regresssion models", *Economics Letters*, **80**, 283-285.

[21] Zellner, A. (1976). "Bayesian and non-Bayesian analysis of the regression model with multvariate student-t error terms", *Journal of American Statistical Association*, **71**, 400-405.