# Density forecasting performance of alternative GARCH and DCS models for daily financial returns

Błażej Mazur[1]

**Abstract**

The objective of the paper is to verify predictive performance (in terms of density forecasts) of some of recently proposed models for daily financial returns. We focus on Beta-Gen-*t*-EGARCH specification proposed by Harvey and Lange (2015,2016), which uses a generalized-*t* conditional distribution. We conduct an empirical comparison of out-of-sample predictive performance against some well-established univariate alternatives. We use daily S&P500 logarithmic returns and investigate sequences of 1525 density forecasts for horizons of 1 to 6 trading days ahead (3rd of January, 2011 till 30th of December, 2016) and the training sample begins on 3rd of January, 2006. We obtain the density forecasts using Bayesian methods. A general Beta-Gen-*t*-EGARCH model dominates other specifications in terms of criteria for density forecast comparison like LPS and CRPS. Moreover, it provides best estimates of Value-at-Risk. In general the results indicate the importance of asymmetric modelling of the news impact curve and of heavy-tailed conditional distribution, which is not unexpected; the results also suggest importance of tail asymmetry which is not addressed here.

***Keywords:*** *empirical finance, Value-at-Risk, generalized t distribution, dynamic conditional score, EGARCH, density forecasting, Bayesian inference*

***JEL Classification:*** G17, C58

## 1 Introduction

There exists a growing body of methods and models available for empirical financial analyses focusing on daily asset returns dynamics. An issue that has been receiving a lot of attention is that of risk evaluation. The related problem of time-series modelling is that of density forecasting. Crucial issues under consideration are those of adequate quantification of uncertainty and estimation of probability of extreme events. At the same time, the volume of available data seems to be growing very rapidly. Risk evaluation often requires costly computational methods and for real-time analysis it is important to balance advancement of computations against feasibility.

Here we focus on recent developments in the field of GARCH models that are popular in applied finance also because of their computational convenience. Well-established tools include extensions of the basic GARCH specification like EGARCH, GJR-GARCH among others. Moreover, simple conditional normality is often generalized to capture features of the

[1] Cracow University of Economics, Department of Econometrics and Operations Research, Rakowicka 27, 31-510 Kraków, Poland, e-mail: blazej.mazur@uek.krakow.pl.

conditional distribution like its heavy tails. A popular choice is that of *t* distribution, though some authors suggest the GED distribution instead.

An interesting class of dynamic models is based on the idea of Dynamic Conditional Score approach which was proposed by Harvey (2013). A related strand of research develops models under the label of Generalized Autoregressive Score (GAS), see Creal et al. (2011, 2013). The general idea of the approach is that features of the conditional sampling distribution are updated in a way that takes into account an autoregressive update and a score of the conditional sampling distribution. This leads to a class of dynamic models that is similar to GARCH models in the spirit – the models are not based on latent processes, though are likely to provide gains in terms of estimation feasibility.

In particular we focus on Beta-Gen-t-EGARCH model, see Harvey and Lange, (2015, 2016) and Harvey and Sucarrat (2014). The model not only makes use of the novel idea of dynamic updating. It also features an interesting (and flexible) form of the conditional sampling distribution, being a generalized *t* distribution. The purpose of the paper is to verify whether the model features lead to benefits in the applied work, in particular whether out-of-sample density predictive performance is competitive compared to some well-established specifications. An important point emphasized here is that – contrary to the work of Harvey and his co-authors – the paper makes use of Bayesian inference methods. This is because it is difficult to find any other modelling approach that allows for estimation uncertainty to be taken into account in a coherent way when constructing density forecasts.

The rest of the paper is organized as follows. Firstly, the types parametric of conditional distributions under consideration are reviewed. Secondly, basic GARCH models are recalled and Beta-Gen-*t*-EGARCH specification is introduced. Thirdly, some remarks on Bayesian model specification, estimation and prediction are considered. Fourthly, an empirical analysis (using daily S&P500 returns) is conducted, with out-of-sample predictive experiments on expanding sub-samples. The general predictive performance is examined using scoring rules that are appropriate for evaluation of density forecasts, namely the Log-Predictive Score (LPS) and Continuous Ranked Probability Score (CRPS) – see Gneiting and Raftery (2007). Moreover, adequacy of tail forecasts is examined in detail, as quality of Value-at-Risk estimates is discussed.

## 2 A generalized *t* distribution

There exists a number of generalizations of *t*-distribution that are of interest. Here we focus on a special case with the probability density function of the following form:

$$f(y)=\frac{1}{\sigma}K(\nu,\gamma)\left(1+\frac{1}{\nu}\left|\frac{y-\mu}{\sigma}\right|^{\gamma}\right)^{-(\nu+1)/\gamma}, \tag{1}$$

with location parameter $\mu$, scale parameter $\sigma$ and two shape parameters $\nu$ and $\gamma$ (see Theodossiou, 1998, Harvey and Lange, 2016). The normalizing constant is given by:

$$K(\nu,\gamma)=\frac{\gamma}{2\nu^{1/\gamma}}\frac{1}{B(\nu/\gamma,1/\gamma)}. \tag{2}$$

The distribution nests a number of interesting special cases. Firstly, as $\gamma=2$ it becomes a Student *t*-distribution with $\gamma$ degrees of freedom. Moreover, as $\gamma\to\infty$, the limiting case is the GED i.e. the Generalized Error Distribution, also known as the Exponential Power Distribution (EPD). Hence the generalized *t*-distribution also encompasses the Gaussian and the Laplace cases - see the discussion in Harvey and Lange (2016) and references therein for derivations of the limiting cases. The distribution seems appealing for financial applications since its tail behavior depends on two parameters (namely $\nu$ and $\gamma$). Though $\gamma$ controls also the shape of the distribution around the mode, it seems reasonable to verify whether such a general distribution can be used to improve performance of density forecasts and risk evaluation.

## 3 GARCH and DCS models under consideration

Consider the following dynamic specification for daily financial logarithmic returns $y_t$:

$$y_t=\delta_t+u_t \tag{3}$$

with

$$u_t=\rho u_{t-1}+\varepsilon_t \tag{4}$$

and

$$\varepsilon_t=\sqrt{h_t}z_t. \tag{5}$$

A standard AR(1)-GARCH(1,1) specification assumes that $z_t$ represents a sequence of independent and identically distributed zero-mean and unit-variance random variables. Moreover, $\delta_t=\delta$ and:

$$h_t=\alpha_0+\alpha_1\varepsilon_{t-1}^2+\beta_1 h_{t-1}. \tag{6}$$

In the GJR GARCH specification of Glosten et al. (1993) the parameter $\alpha_1$ takes different values depending on the sign of $\varepsilon_{t-1}$:

$$h_t=\alpha_0+\alpha_1^-\varepsilon_{t-1}^2 I_{\varepsilon_{t-1}<0}+\alpha_1^+\varepsilon_{t-1}^2 I_{\varepsilon_{t-1}>0}+\beta_1 h_{t-1}. \tag{7}$$

This allows for asymmetry of the news impact curve (in such a model it is possible that negative "surprises" lead to increased conditional volatility while positive surprises do not). The GJR effect is very often found empirically important. Another option is to consider so-called In-Mean effect that aims to capture the leverage effect:

$$\delta_t = \delta + \eta\sqrt{h_t} \,. \tag{8}$$

Parametric restrictions are of course necessary to ensure that the conditional variance $h_t$ is positive; further restrictions ensure stationarity of the model.

The simple GARCH-type models have certain disadvantages; an effort to deal with some of them has led to a development of the EGARCH model, where log-volatility is updated instead of volatility, see Nelson (1991). Here we are going to follow Harvey and Lange (2015, 2016) who derive similar model from the idea of Dynamic Conditional Score approach. It can be used to construct the following dynamic updating mechanism. Assume that $\lambda = \ln\sigma$ i.e. $\lambda$ denotes logarithm of scale of the conditional distribution of error terms. Let:

$$\lambda_t = \omega(1-\phi) + \phi\lambda_{t-1} + \kappa w_{t-1} \,. \tag{9}$$

where $w_{t-1}$ denotes score of the conditional distribution wrt. $\lambda$ at *t*-1 i.e. a derivative of log-(conditional) density at the actual outturn. Harvey and Lange (2016) consider also an asymmetric specification:

$$\lambda_t = \omega(1-\phi) + \phi\lambda_{t-1} + \kappa w_{t-1} + \kappa^* w_{t-1}^* \,. \tag{10}$$

with:

$$w_t^* = \mathrm{sgn}(\delta_t - y_t)(w_t + 1), \tag{11}$$

where $\omega$, $\phi$, $\kappa$ and $\kappa^*$ are model parameters. Similar approach can be taken for updating of conditional location parameter. Moreover, the expressions for the score functions (depending on the class of the conditional distribution) as well as the relationship between location and log-scale score is discussed in Harvey (2013) and Harvey and Lange (2016).

The above formulas define model classes considered in the paper. The first group of "traditional" models is that of AR-GARCH class. For the model class we assume that the conditional distribution is Student *t* with $\nu$ degrees of freedom. The models under consideration include a limiting case of conditional normality with $\nu \to \infty$. The volatility equation is either of GARCH (1,1) form or of GJR (1,1) form.

The other class of models that includes the "novel" specifications assume that the conditional distribution is of the generalized *t* form. Moreover, the conditional location and conditional scale is updated according to the DCS idea. Below a convention for labelling is used according to which the above EGARCH model is denoted by S/L-DCS(1,1)-Gt-AL-AS, where S/L-DCS(1,1) denotes DCS-updating of conditional scale (S) and location (L), Gt denotes the conditional distribution (being generalized *t* given by (1)), and AL and AS indicate that both scale and location equation feature asymmetric form of the type (10). Special (or limiting) cases as to the conditional distribution are labelled GED, t or N for GED, Student *t* and Gaussian distribution.

## 4 Bayesian model specification, estimation and prediction

Bayesian estimation of the model class outlined above requires certain prior assumptions as to model parameters. Here the assumptions are somewhat simple, since the number of observations is rather large, as it is usually the case with financial returns. Hence we assume that even if the prior structure is not fully justified from the theoretical point of view, it is not going to affect the predictive performance in a serious way, as any irrelevant prior information should be updated by the information coming from the training sample.

Priors used in the paper are independent and proper, the latter is important as it is well known that the use of improper priors for parameters controlling tail thickness might cause theoretical problems as to the existence of the proper posterior with finite moments. Throughout the paper all the autoregressive parameters feature flat priors do not exceeding the (-1,1) interval, flat priors are also assumed for $\alpha_1$ and $\beta_1$. The score parameters $\kappa$ are assumed to be a priori *t*-distributed. For $\gamma$ we assume a prior that is restricted to arbitrarily chosen range (1.05,6.05) covering empirically relevant cases and is practically uniform over the range. Prior for $\nu$ assumes that the parameter is restricted to the range (2.01,150) since it ensures that the conditional variance exists. The upper bound is used for numerical convenience only and the restriction has very limited empirical consequences.

The inference is conducted using MCMC techniques. Firstly, the posterior distribution is evaluated using a Metropolis-Hastings algorithm with a random walk proposal. Based on the results an Independent version of the Metropolis-Hastings algorithm is constructed. This is crucial for the predictive exercise conducted here. As all the models have likelihood that is fully sequential, if one can construct a proposal that is valid for all the subsamples within the expanding sample exercise, evaluation of the likelihood for all the samples is equivalent to its evaluation for the longest subsample. Hence the same proposal can be used for all the instances of the model (though the acceptance has to be addressed separately, so that each subsample has the associated MCMC chain). Within such a setup it is possible to undertake a sequential forecasting exercise of that complexity in a practical timespan.

## 5 Empirical model comparison: analysis of daily S&P500 returns 2006-2016

For the sake of the empirical comparison we analyse daily S&P500 logarithmic returns (based on closing prices), depicted in Figure 1. We use a sequence of expanding samples, the first observation corresponding to the 3rd of January, 2006. The shortest sample ends on 31st of December, 2010 (1259 observations) and it covers the period of the global financial crisis, with large negative returns of 15th of October and 1st of December 2008. The longest sample ends on 30th of December, 2016 (2751 observations).

| Model label | LPS | CRPS | std. dev. | $q_{0.95}$-$q_{0.05}$ | IQR | $q_{0.6}$-$q_{0.4}$ |
|---|---|---|---|---|---|---|
| AR(1)-N-GARCH(1,1) | -1873.59 | 0.486 | 0.92 | 3.04 | 1.24 | 0.47 |
| AR(1)-t-GARCH(1,1) | -1834.82 | 0.485 | 0.95 | 2.95 | 1.07 | 0.39 |
| AR(1)-t-IM-GARCH(1,1) | -1832.24 | 0.484 | 0.95 | 2.95 | 1.07 | 0.39 |
| AR(1)-N-GJR(1,1) | -1833.31 | 0.482 | 0.92 | 3.01 | 1.23 | 0.46 |
| AR(1)-t-GJR(1,1) | -1803.54 | 0.481 | 0.94 | 2.95 | 1.08 | 0.40 |
| AR(1)-t-IM-GJR(1,1) | -1796.75 | 0.480 | 0.94 | 2.96 | 1.08 | 0.40 |
| S/L-DCS(1,1)-Gt-AL-AS | -1795.70 | 0.480 | 0.94 | 3.09 | 1.06 | 0.37 |
| S/L-DCS(1,1)-Gt-AS | -1796.42 | 0.480 | 0.95 | 3.08 | 1.06 | 0.37 |
| S/L-DCS(1,1)-GED-AS | -1800.23 | 0.480 | 0.95 | 3.12 | 1.05 | 0.36 |
| S/L-DCS(1,1)-t-AS | -1798.89 | 0.482 | 0.94 | 3.00 | 1.11 | 0.41 |
| S/L-DCS(1,1)-Gt | -1837.52 | 0.485 | 0.93 | 3.05 | 1.03 | 0.36 |
| S/L-DCS(1,1)-t | -1843.27 | 0.487 | 0.93 | 2.95 | 1.08 | 0.40 |

**Table 1.** Overall predictive performance for $h = 1$; LPS is cumulated, CRPS is averaged, std. dev. denotes average standard deviation of the predictive distribution, the last three columns depict averaged distance between respective quantiles of predictive distributions.

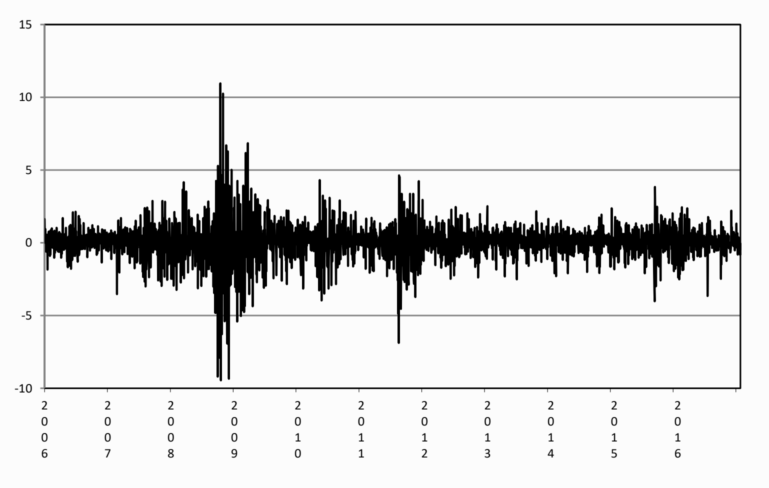


**Fig. 1.** S&P500 daily logarithmic returns, 2006-2016.

| Model label | LPS | CRPS | std. dev. | $q_{0.95}$-$q_{0.05}$ | IQR | $q_{0.6}$-$q_{0.4}$ |
|---|---|---|---|---|---|---|
| AR(1)-N-GARCH(1,1) | -1887.41 | 0.487 | 0.93 | 3.06 | 1.25 | 0.47 |
| AR(1)-t-GARCH(1,1) | -1844.5 | 0.487 | 0.96 | 2.98 | 1.07 | 0.39 |
| AR(1)-t-IM-GARCH(1,1) | -1841.27 | 0.485 | 0.96 | 2.98 | 1.07 | 0.39 |
| AR(1)-N-GJR(1,1) | -1843.15 | 0.484 | 0.92 | 3.03 | 1.23 | 0.46 |
| AR(1)-t-GJR(1,1) | -1815.43 | 0.483 | 0.95 | 2.97 | 1.08 | 0.40 |
| AR(1)-t-IM-GJR(1,1) | -1813.43 | 0.482 | 0.94 | 2.96 | 1.08 | 0.40 |
| S/L-DCS(1,1)-Gt-AL-AS | -1811.49 | 0.482 | 0.95 | 3.1 | 1.06 | 0.37 |
| S/L-DCS(1,1)-Gt-AS | -1814.38 | 0.482 | 0.95 | 3.09 | 1.05 | 0.37 |
| S/L-DCS(1,1)-GED-AS | -1817.22 | 0.483 | 0.95 | 3.12 | 1.05 | 0.36 |
| S/L-DCS(1,1)-t-AS | -1816.8 | 0.483 | 0.95 | 3.01 | 1.11 | 0.41 |
| S/L-DCS(1,1)-Gt | -1854.29 | 0.486 | 0.94 | 3.08 | 1.03 | 0.36 |
| S/L-DCS(1,1)-t | -1852.54 | 0.49 | 0.94 | 2.97 | 1.09 | 0.40 |

**Table 2.** Overall predictive performance for $h = 2$, LPS is cumulated, CRPS is averaged, std. dev. denotes average standard deviation of the predictive distribution, the last three columns depict averaged distance between respective quantiles of predictive distributions.

The data from the first two weeks of January 2017 are used for forecast evaluation, as we consider prediction horizons ranging from 1 to 6 trading days ahead. Hence the *ex-post* evaluation of predictive performance is based on sequences of 1510 realized forecasts for each horizon. All the models are reestimated with each new observation added. Basic results of the comparison are given in Table 1 and Table 2, for one-day-ahead and two-days-ahead forecasts, respectively. The results are similar for horizons up to 6 days ahead.

| **Model label** | **1%** | **2%** | **3%** | **4%** | **5%** | **6%** | **7%** | **8%** | **9%** | **10%** |
|---|---|---|---|---|---|---|---|---|---|---|
| AR(1)-N-GARCH(1,1) | 2.03 | 3.02 | 4 | 4.66 | 5.51 | 6.36 | 7.08 | 8.2 | 8.85 | 9.57 |
| AR(1)-t-GARCH(1,1) | 1.25 | 2.62 | 4 | 5.38 | 6.3 | 7.61 | 8.72 | 9.77 | 10.69 | 11.87 |
| AR(1)-t-IM-GARCH(1,1) | 1.25 | 2.62 | 3.93 | 5.11 | 6.3 | 7.87 | 8.59 | 9.77 | 10.69 | 12.07 |
| AR(1)-N-GJR(1,1) | 1.77 | 3.02 | 3.54 | 4.39 | 5.31 | 6.36 | 6.89 | 7.74 | 8.26 | 8.85 |
| AR(1)-t-GJR(1,1) | 1.11 | 2.75 | 3.8 | 5.05 | 6.16 | 7.34 | 8.2 | 8.98 | 9.77 | 10.75 |
| AR(1)-t-IM-GJR(1,1) | 1.11 | 2.75 | 3.8 | 4.85 | 6.16 | 7.28 | 8.13 | 8.98 | 9.9 | 11.02 |
| S/L-DCS(1,1)-Gt-AL-AS | 1.05 | 2.3 | 3.34 | 4.13 | 5.18 | 6.56 | 7.48 | 8.72 | 9.38 | 10.03 |
| S/L-DCS(1,1)-Gt-AS | 1.05 | 2.36 | 3.34 | 4.26 | 5.44 | 6.62 | 7.48 | 8.79 | 9.57 | 10.03 |
| S/L-DCS(1,1)-GED-AS | 1.05 | 2.16 | 3.21 | 4.13 | 5.38 | 6.43 | 7.41 | 8.72 | 9.51 | 10.03 |
| S/L-DCS(1,1)-t-AS | 1.31 | 2.75 | 3.67 | 4.79 | 6.03 | 6.89 | 7.8 | 8.72 | 9.31 | 10.43 |
| S/L-DCS(1,1)-Gt | 1.44 | 2.56 | 3.74 | 4.98 | 6.43 | 7.34 | 8.72 | 9.77 | 10.62 | 11.48 |
| S/L-DCS(1,1)-t | 1.64 | 2.82 | 4.26 | 5.7 | 7.08 | 7.87 | 9.11 | 9.9 | 10.82 | 11.87 |

**Table 3.** Lower tail predictive performance ($h$ = 1) – empirical fraction of realized forecasts that exceed quantile of order $\alpha$ = 1%,…,10% of the predictive distribution.

Values of LPS (computed with natural log, the higher the better) and CRPS (the lower the better, as it can be perceived as a generalization of *ex-post* absolute error) shown in Table 1 and Table 2 suggests that the model that dominates the comparison is the Beta-Gen-t-EGARCH with asymmetry in scale and location (S/L-DCS(1,1)-Gt-AL-AS), though it is followed by the AR(1)-GJR-GARCH model with conditional *t*-distribution and In-Mean effect. Asymmetry in the volatility equation (introduced via $\kappa^*$ or $\alpha_1^+$ parameters) seems to be the most important factor for goodness-of-fit measured by the LPS. Differences in the average predictive standard deviation or average length of predictive intervals (constructed from quantiles) convey some information as to differences in shape of the predictive distribution between the models implied by the differences in assumptions concerning the type of the conditional distribution. VaR estimates perform best for S/L-DCS(1,1)-Gt-AL-AS, followed by the model with GED distribution. Interestingly for $h = 1$ N-GARCH model performs badly for $0.01 < \alpha < 0.05$ and has good performance for $0.05 < \alpha < 0.1$.

## Conclusions

The objective of the paper is to verify predictive performance (in terms of density forecasts) of some of recently proposed models for daily financial returns. We focus on Beta-Gen-*t*-EGARCH specification of Harvey and Lange (2015, 2016). We conduct an empirical comparison of out-of-sample predictive performance of the model against some of alternatives that are well-established in empirical finance, like GJR-*t*-GARCH model. As the point of interest here is on density forecasting performance (in particular risk evaluation) it is important that estimation uncertainty problem is addressed. This is even more important as the model under consideration uses a complicated conditional sampling distribution, namely a Generalized *t*-distribution. The distribution nests a number of interesting special cases such as *t* distribution and GED distribution. Quantification of the estimation uncertainty for shape (in particular tail) parameters might be challenging though it has to be taken into account for the sake of density forecasting. This is why we make use of Bayesian inference methods which is an interesting contribution.

The empirical comparison conducted here uses daily S&P500 logarithmic returns data. We investigate sequences of 1525 density forecasts. The results confirm that the new specification is an interesting tool of empirical financial analyses, as Beta-Gen-t-EGARCH (with asymmetry in log-scale equation) dominates other specifications in terms of criteria for density forecast comparison like LPS or CRPS in all the horizons under consideration (though the difference against the best 'traditional' model is not large). Moreover, Beta-Gen-t-EGARCH provides best VaR estimates during the verification period. However, the gain over the AR(1)-t-GJR-In-Mean model is rather moderate. In general the results indicate the importance of asymmetric modelling of the news impact curve and of heavy-tailed conditional distribution, which is not unexpected. Moreover, the results suggest that tail asymmetry might be an important feature of such models. This is because all the models underestimate risk in the lower tail and overestimate risk in the upper tail (according to VaR results) which is a suggestion for further research.

**Acknowledgements**

The research was supported by the funds granted to the Faculty of Management, Cracow University of Economics within the framework of the subsidy for the maintenance of research potential.

**References**

Creal, D., Koopman, S. J. & Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics*, 29, 552-563.

Creal, D., Koopman, S. J. & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28, 777-795.

Glosten, L., Jagannathan, R. & Runkle, D. (1993). On the relation between expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.

Gneiting, T., & Raftery, A. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477).

Harvey, A.C. (2013). Dynamic models for Volatility and Heavy Tails: with applications to financial and economic time series. Econometric Society Monograps, Cambridge University Press.

Harvey, A.C., & Lange, J.-R. (2016). Volatility Modelling with a Generalized *t*-distribution. *Journal of Time Series Analysis*, 38, 175-190.

Harvey, A.C., & Lange, J.-R. (2015). Modelling the Interactions between Volatility and Returns. *Cambridge Working Papers in Economics 1518*.

Harvey, A.C. & Sucarrat, G. (2014). EGARCH Models with Fat Tails, Skewness and Leverage. *Computational Statistics and Data Analysis*, 26, 320–338.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59, 347-370.

Theodossiou, P. (1998). Financial data and the skewed generalized *t* distribution, *Management Science*, 44, 1650-1661.