ANZIAM J. 55 (EMAC2013) pp.C153–C167, 2014

On quantiles of the temporal aggregation of a stable moving average process and their applications

A. W. Barker[1]

(Received 18 December 2013; revised 24 March 2014)

Abstract

A stochastic volatility model is proposed for the daily log returns of a financial asset based on conditional log quantile differences, assuming the availability of high frequency intraday log returns. Calculation of the conditional log quantile differences is performed with the assumption that the intraday log returns follow a stable moving average process. The use of conditional log quantile difference in the proposed model, rather than conditional variance in standard models, offers an increase in flexibility, with the potential for a different dependency structure at different parts of the conditional distribution. The proposed model makes use of high frequency intraday log returns which are generally neglected in standard models. Formulae for the calculation of the

http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/7826 gives this article, Published May 22, 2014, as part of the Proceedings of the 11th Biennial Engineering Mathematics and Applications Conference. ISSN 1446-8735.

conditional log quantile differences are provided and a method for their estimation is described The proposed model was applied to the ASX200 index from 2009 and 2010.

Contents

1 Introduction

It is generally accepted that the volatility of a financial market asset return displays a much stronger autocorrelation than the actual return [5]. A vast collection of models, referred to as conditional heteroscedastic models, have been proposed for modelling the volatility of financial market asset returns. The most popular of these models is the generalised autoregressive conditional heteroscedastic (GARCH) model [4] and its variations such as the IGARCH, EGARCH and EGARCH-M models [12]. Another class of models for modelling the volatility of financial market asset returns is the stochastic volatility (SV) model [11, 12, e.g.]. In this article we are mainly concerned with SV models.

Let P_d denote the closing price of a financial asset on day d and define the logarithm of the daily return on that asset by

$$Y_d = \log(P_d/P_{d-1}) \,. \tag{1}$$

The SV model for the sequence of log returns $\{Y_d\}$, for $d \in \mathbb{Z}$, is defined by

$$Y_d = \sigma_d \eta_d \,, \tag{2}$$

$$\log(\sigma_d^2) = \mu + \sum_{j=1}^{k} \phi_j \log(\sigma_{d-j}^2) + e_d + \sum_{j=1}^{m} \psi_j e_{d-j} \,, \tag{3}$$

where σ_d^2 is the conditional variance of the log return Y_d, $\{\eta_d\}$ is a sequence of independent and identically distributed (iid) $N(0,1)$ random variables, $\{e_d\}$ is a sequence of iid $N(0,\kappa^2)$ random variables independent of $\{\eta_d\}$ and with variance κ^2, μ is the location parameter, the ϕ terms are the autoregressive parameters with order k, and the ψ terms are the moving average parameters with order m.

The SV model contains four sequences of random variables: $\{Y_d\}$, $\{\sigma_d^2\}$, $\{\eta_d\}$ and $\{e_d\}$, of which only $\{Y_d\}$ is observable. The daily log return is the temporal aggregation of the intraday log returns, which are also observable. Given appropriate assumptions on the dependency structure of the intraday log return process, the availability of high frequency intraday log return data allows the estimation of σ_d^2 directly. Thus estimation of the parameters of (3) becomes possible, independent of (2).

However, empirical evidence from the ASX200 index of the Australian Stock Exchange, as shown in Section 5, suggests that generally the intraday log return processes are not finite variance processes. Under these conditions, it would require a somewhat contrived dependency structure for the temporal aggregation of the intraday log return process to have a finite variance, that is, for σ_d^2 to be finite. To overcome this problem, we propose using the log quantile difference, which exists for all stationary processes, as the measure of volatility.

We assume that the intraday log return process is a stable moving average process and show in Section 2 how to calculate the log quantile difference of the daily log return process under this assumption. A complete definition of our model and assumptions are given in Section 3 and some brief comments on estimation in Section 4. The results of applying our model to ASX200 index returns are shown in Section 5.

2 Definitions

2.1 Stable moving average processes

Let $\{X_t\}$ be the moving average process of order q,

$$X_t = \sum_{j=0}^{q} \theta_j e_{t-j}\,, \tag{4}$$

where t is the time, measured in units of the time between successive intraday price reports, $\theta_0 = 1$ and $\{e_t\}$ is an independently and identically distributed stable (iids) sequence of random variables such that

$$e_t \sim S^0_\alpha(\beta, \gamma, \delta)\,, \tag{5}$$

using Nolan's S^0 parameterisation of stable distributions [10]. The parameter α determines the heaviness of the tails of the distribution, with special cases at $\alpha = 1$ (Cauchy) and $\alpha = 2$ (Gaussian). The parameters β, γ and δ, respectively, determine the skewness, scale and location of the distribution. The $q + 1$ dimensional vector of moving average parameters is

$$\theta = (\theta_0, \ldots, \theta_q). \tag{6}$$

Using the properties of the S^0 parameterisation it can be shown [3] that the distribution of the stable moving average process defined in (4) is

$$X_t \sim S^0_\alpha(\beta^{(1)}, \gamma^{(1)}, \delta^{(1)})\,, \tag{7}$$

where

$$\gamma^{(1)} = \left(\sum_{j=0}^{q} |\theta_j|^\alpha\right)^{1/\alpha} \gamma\,, \tag{8}$$

$$\beta^{(1)} = \frac{\sum_{j=0}^{q} \operatorname{sgn}(\theta_j)|\theta_j|^\alpha}{\sum_{j=0}^{q} |\theta_j|^\alpha}\beta\,. \tag{9}$$

In this article we do not require the formula for $\delta^{(1)}$.

2.2 Temporal aggregation

The temporal aggregation of $\{X_t\}$ is

$$S_t^{(r)} = \sum_{i=0}^{r-1} X_{t-i}\,, \tag{10}$$

where r is the aggregation level. When $r = 1$,

$$S_t^{(1)} = X_t\,. \tag{11}$$

Henceforth, we refer to $\{S_t^{(r)}\}$ as the temporal aggregation of $\{X_t\}$ or the aggregated process, and $\{X_t\}$ is the base process.

The aggregated process is still a moving average process. For $r \geqslant q$,

$$S_t^{(r)} = \sum_{j=0}^{r+q-1} c_j e_{t-j}\,, \tag{12}$$

where

$$c_j = \begin{cases} \sum_{i=0}^{j} \theta_i \,, & j = 0, \ldots, q-1 \,, \\ \sum_{i=0}^{q} \theta_i \,, & j = q, \ldots, r-1 \,, \\ \sum_{i=j-r+1}^{q} \theta_i \,, & j = r, \ldots, r+q-1 \,. \end{cases} \tag{13}$$

Therefore, following the same argument used to derive the distribution in (7), for $r \geqslant q$ we obtain

$$S_t^{(r)} \sim S_\alpha^0(\beta^{(r)}, \gamma^{(r)}, \delta^{(r)}) \,, \tag{14}$$

where, after the substitution of (8) and (9),

$$\gamma^{(r)} = \left(\frac{\sum_{j=0}^{r+q-1} |c_j|^\alpha}{\sum_{j=0}^{q} |\theta_j|^\alpha} \right)^{1/\alpha} \gamma^{(1)}, \tag{15}$$

$$\beta^{(r)} = \frac{\sum_{j=0}^{r+q-1} \operatorname{sgn}(c_j)|c_j|^\alpha \sum_{j=0}^{q} |\theta_j|^\alpha}{\sum_{j=0}^{r+q-1} |c_j|^\alpha \sum_{j=0}^{q} \operatorname{sgn}(\theta_j)|\theta_j|^\alpha} \beta^{(1)} \,. \tag{16}$$

In this article we do not require the formula for $\delta^{(r)}$.

2.3 Log quantile differences

Let ξ_{p_i} denote the p_ith quantile of some distribution function. At quantile levels $p = (p_1, p_2)$, such that $p_1 \neq p_2$ and $0 < p_1, p_2 < 1$, we define the log quantile difference

$$\zeta_p = \log(|\xi_{p_2} - \xi_{p_1}|) \,. \tag{17}$$

We assume that any random variable on which a log quantile difference is calculated has a positive density at ξ_{p_1} and ξ_{p_2}. This assumption implies uniqueness of the quantiles ξ_{p_1} and ξ_{p_2} and that the log quantile difference is finite. Recall that stable distributions satisfy this condition.

The δ parameter of the S^0 parameterisation of the stable distribution acts as a location parameter. It can be shown [3] that the log quantile difference of any stable distribution, being a function of the difference of two quantiles, is independent of the δ parameter, that is, it is completely determined by the α, β and γ parameters.

3 Model description

Let $\{X_{d;t}\}$ and $\{S^{(r)}_{d;t}\}$ denote, respectively, the base and aggregated log return processes of a financial asset on day $d \in \mathbb{Z}$. Let $\zeta^{(r)}_{d;p}$ denote the log quantile difference of $\{S^{(r)}_{d;t}\}$ at quantile levels $p = (p_1, p_2)$. We define the log quantile difference stochastic volatility (LQDSV) model at aggregation level r and quantile levels p by

$$\zeta^{(r)}_{d;p} = \mu^{(r)}_p + \sum_{j=1}^{k} \phi^{(r)}_{j;p} \zeta^{(r)}_{d-j;p} + e^{(r)}_{d;p} + \sum_{j=1}^{m} \psi^{(r)}_{j;p} e^{(r)}_{d-j;p}\,, \tag{18}$$

where $\{e^{(r)}_{d;p}\}$ is a sequence of iid $N(0, \kappa^{(r)}_p)$ random variables with variance $\kappa^{(r)}_p$. As is the case in the SV model, the volatility term $\zeta^{(r)}_{d;p}$ in the LQDSV model (18) is an autoregressive moving average (ARMA) process.

The calculation of the quantiles of sums of general random variables is usually very difficult, so, in order to make the LQDSV model more tractable and allow the use of the theory in Section 2, we make the following assumptions on the base process.

1. The intraday log returns comprising the base process $\{X_{d;t}\}$ are calculated over the same constant time period for all $d \in \mathbb{Z}$.
2. The base process $\{X_{d;t}\}$ is a stable moving average process.

Assumption 1 is not too onerous; however, Assumption 2 may be disputed. A stable moving average process does provide the expected stylised facts of heavy tails, some skewness and an absence of long term autocorrelations of the asset returns. The implicit assumption that asset returns have the same distribution throughout a trading day is less likely. Financial asset returns commonly have higher volatility at the beginning of a trading day as overnight information is being absorbed, compared to the middle of the day when the traders are at lunch.

For an appropriate choice of aggregation level, r_T, such that there are r_T multiples of the base process time period in each trading day, the daily return is

$$Y_d = S_{d;t}^{(r_T)} . \tag{19}$$

4 Model estimation

Given estimators for θ_d , α_d , $\beta_d^{(1)}$ and $\gamma_d^{(1)}$, we use equations (15) and (16) to calculate estimators for $\beta_d^{(r)}$ and $\gamma_d^{(r)}$. Given estimators for θ_d , α_d , $\beta_d^{(r)}$ and $\gamma_d^{(r)}$, we use the definition of the log quantile difference in (17) to calculate an estimator for $\zeta_{d;p}^{(r)}$. A derivation of the asymptotic distribution of this estimator for $\zeta_{d;p}^{(r)}$ is beyond the scope of this article.

4.1 Stable distribution parameter estimation

McCulloch [9] proposed asymptotically normal estimators for the parameters of a stable distribution from an independent sample, based on functions of the empirical quantile estimators. Analytic formulae for the asymptotic distribution were not derived, although a method for numerically calculating the asymptotic distribution was shown.

The empirical quantile estimators from an S-mixing process sample are consistent and asymptotically normal [8]. The class of S-mixing processes includes stable moving average processes. It follows that the McCulloch stable distribution parameter estimators can also be used on samples from a stable moving average process, although the asymptotic distribution of these estimators is different to the asymptotic distribution of estimators from an independent sample [2].

4.2 Infinite variance moving average process parameter estimation

Many of the common methods used for the estimation of finite variance ARMA processes, for example Gaussian maximum likelihood, least squares and Hannan–Rissanen, are not valid for the estimation of infinite variance ARMA processes. Several alternative methods have been proposed, but the estimators from many of these methods have complicated asymptotic distributions which make them difficult to use [6, e.g.]. However, the self-weighted least absolute deviation (SLAD) method was shown under various conditions to provide estimators which are both consistent and asymptotically normal [13].

To estimate the order of a stable moving average process, we use the sample autocorrelation function, which is a well-defined statistic even for infinite variance processes [7]. However, the rate of convergence to the asymptotic distribution is very slow, so for this analysis we estimate the order of a stable moving average process to be the highest lag, up to a maximum of eight, of the sample autocorrelation function which lies outside the Cauchy limit 95% confidence interval [1].

5 Application to ASX200 returns

In this section we apply the LQDSV model to returns of the ASX200 index. The ASX200 index is derived from the market capitalisation of the leading 200 companies listed on the Australian Stock Exchange (ASX). Normal trading days are Monday to Friday from 10 am to 4 pm. The ASX200 index is reported every 30 seconds throughout each trading day. Thus on a normal trading day there are 721 values of the ASX200 index reported, from which we calculate 720 intraday log returns over 30 second intervals. The intraday log returns over 30 second intervals are used as the base process data in this analysis.

The data used in this analysis is from the calendar years 2009 and 2010, containing 507 trading days. The data was obtained from the Thomson–Reuters Tick History service.

Unfortunately, on some trading days, there are less than 721 values of the ASX200 index reported. On Christmas Eve and New Years Eve, the ASX closes early at 2 pm, so there are only 481 values of the ASX200 index reported on those days. On other days, there are periods where the data is not reported (missing data) or is reported at the same value for several minutes (frozen data). Instances of missing and frozen data are not frequent and generally only last for a few minutes. For the purposes of this analysis, the estimation of $\zeta_{d;p}^{(r)}$ on trading days containing less than 721 reported values of the ASX200 index, is performed assuming that all reported values are from consecutive time intervals. Estimation of the log quantile difference of the daily return was always done using an aggregation level of 720, regardless of the number of intraday log returns over 30 second intervals that were reported on that day.

Estimates of α, $\beta^{(1)}$ and $\gamma^{(1)}$ were calculated for each of the 507 trading days in 2009 and 2010. Partial results for α and $\beta^{(1)}$ estimates are displayed in Figure 1. Estimates of α were within the range $(1.14, 2.00]$ with a median value of 1.55 and eight trading days had an estimate of 2.00, that is, a Gaussian distribution. The estimates of $\beta^{(1)}$ had a median value of 0.00 and approximately 93% of the estimates were within the range $(-0.60, 0.60)$.

As shown in Table 1, on almost half the trading days the base process is estimated to have order zero, that is, they are considered to be iid. In general, even on those trading days where the base process is estimated to have a high order, the sample autocorrelation functions exceed the Cauchy limit 95% confidence interval at only two or three lags. Thus an autoregressive model for these base processes would not be appropriate.

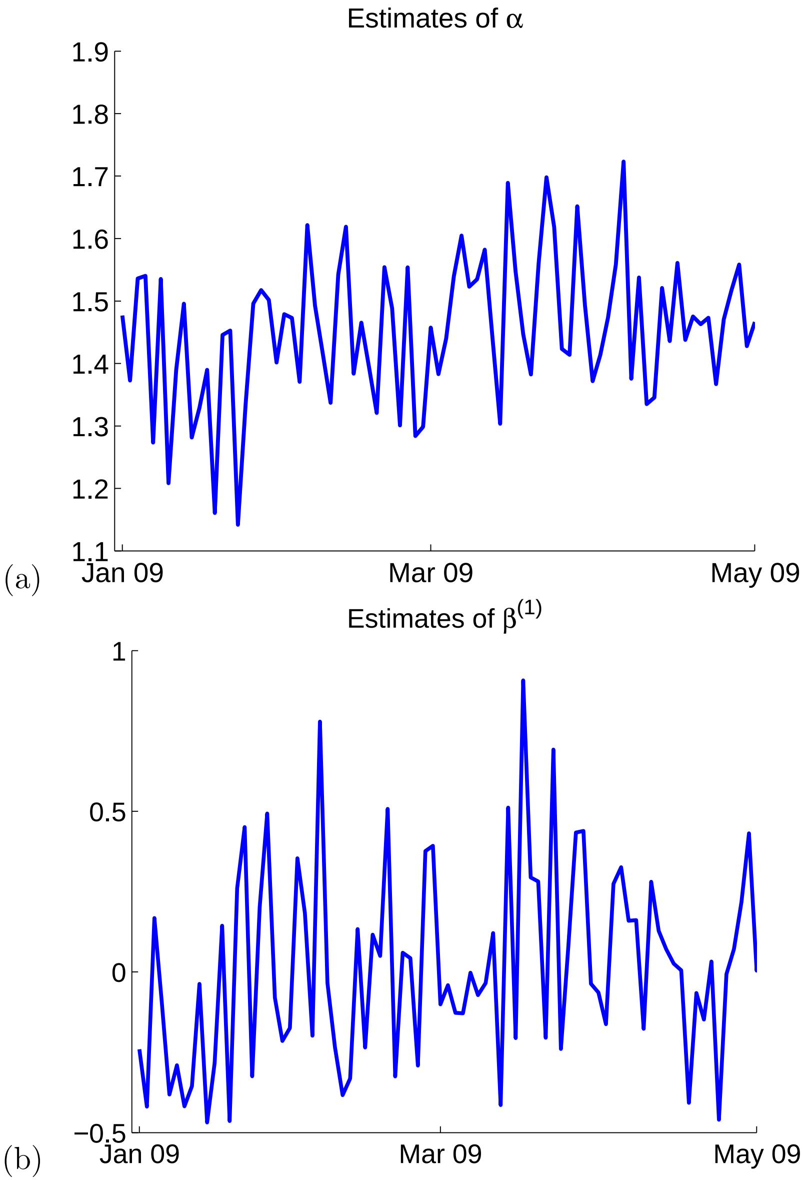
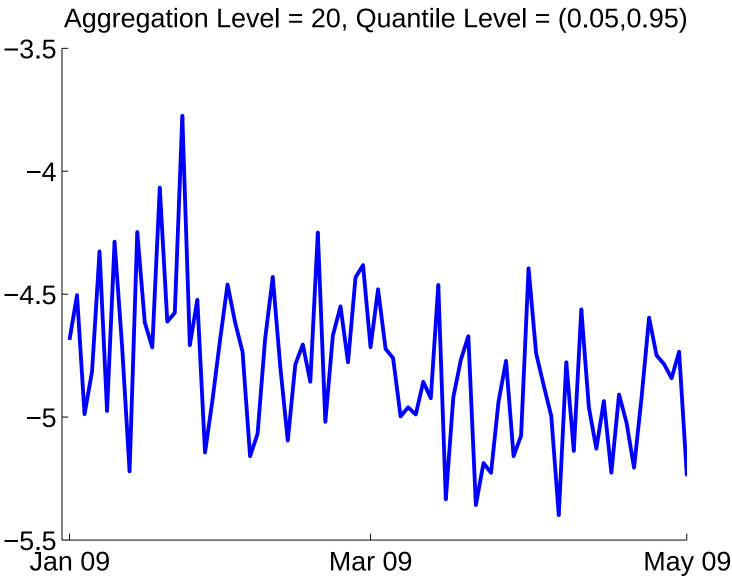


Figure 1: Estimates for (a) α and (b) $\beta^{(1)}$ from the base process of 30 second intraday log returns of ASX200 index data over the period 1-Jan-2009 to 30-April-2009

(a) 

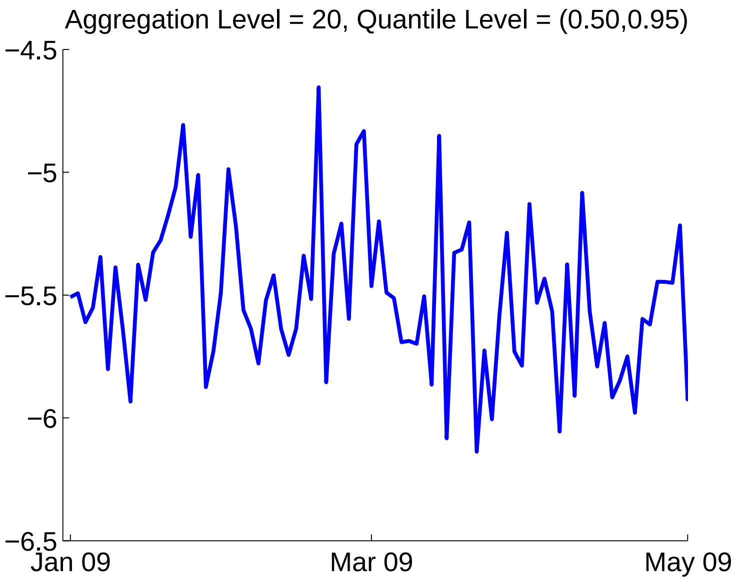
(b) 

Figure 2: Estimates for $\zeta_p^{(r)}$ from the base process of 30 second intraday log returns of ASX200 index data over the period 1-Jan-2009 to 30-April-2009 for (a) $r = 20$ and $p = (0.05, 0.95)$; and (b) $r = 20$ and $p = (0.50, 0.95)$.

Table 1: Categorisation of the 507 trading days in 2009 and 2010 according to the estimated order of the base process stable moving average model.

Order	0	1	2	3	4	5	6	7	8
No. Trading Days	238	46	27	5	72	90	7	6	16

Table 2: Parameter estimates of an ARMA$(1,1)$ model fitted to the estimates of $\zeta_p^{(r)}$ for various aggregation and quantile levels.

Quantile Level	Aggregation Level	$\mu_p^{(r)}$	$\phi_{1;p}^{(r)}$	$\psi_{1;p}^{(r)}$	$\kappa_p^{(r)}$
$(0.05, 0.95)$	20	-0.173	0.968	-0.656	0.092
	720	-0.111	0.964	-0.696	0.230
$(0.50, 0.95)$	20	-0.189	0.969	-0.687	0.103
	720	-0.133	0.965	-0.713	0.238

Using the methods described in Section 4, estimates were calculated for $\zeta_p^{(r)}$ at various quantile and aggregation levels (Figure 2). An ARMA$(1,1)$ model was fitted to the time series of $\zeta_p^{(r)}$ at various quantile and aggregation levels using Gaussian maximum likelihood (Table 2). The parameters $\phi_{1;p}^{(r)}$ have similar values for each of the quantile and aggregation levels. The difference values for $\psi_{1;p}^{(r)}$ and $\kappa_p^{(r)}$ are largely due to the increased measurement error associated with the estimation of $\zeta_p^{(r)}$ at higher aggregation levels. Similar results were seen for other quantile and aggregation levels not shown here. The values of $\phi_{1;p}^{(r)}$ are close to 1, which suggests an autoregressive integrated moving average (ARIMA) model for this data might be appropriate.

Acknowledgements The author thanks Dr N. Kordzakhia for her helpful comments on this article.

References

[1] R. J. Adler, R. E. Feldman, and C. Gallagher. Analysing stable time series. In R. J. Adler, R. E. Feldman, and M. S. Taqqu, editors, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser, 1998. C161

[2] A. W. Barker. Estimation of stable distribution parameters from a dependent sample. Technical report, 2014. http://arxiv.org/abs/1405.0374 C161

[3] A. W. Barker. On the log quantile difference of the temporal aggregation of a stable moving average process. Technical report, 2014. http://arxiv.org/abs/1404.6875 C157, C158

[4] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31:307–327, 1986. doi:10.1016/0304-4076(86)90063-1 C154

[5] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Financ.*, 1:223–236, 2001. doi:10.1080/713665670 C154

[6] R. A. Davis. Gauss–Newton and M-estimation for ARMA processes with infinite variance. *Stoch. Proc. Appl.*, 63:75–95, 1996. doi:10.1016/0304-4149(96)00063-4 C161

[7] R. A. Davis and S. Resnick. Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Stat.*, 14:533–558, 1986. http://www.jstor.org/stable/2241234 C161

[8] Y. Dominicy, S. Hörmann, H. Ogata, and D. Veredas. On sample marginal quantiles for stationary processes. *Stat. Prob. Lett.*, 83:28–36, 2013. doi:10.1016/j.spl.2012.07.016 C160

[9] J. McCulloch. Simple consistent estimators of stable distribution parameters. *Commun. Stat. Simulat.*, 15:1109–1136, 1986. doi:10.1080/03610918608812563 C160

[10] J. P. Nolan. Parameterizations and modes of stable distributions. *Stat. Prob. Lett.*, 38:187–195, 1998. doi:10.1016/S0167-7152(98)00010-8 C156

[11] S. J. Taylor. Modelling stochastic volatility: A review and comparative study. *Math. Financ.*, 4:183–204, 1994. doi:10.1111/j.1467-9965.1994.tb00057.x C155

[12] R. S. Tsay. *Analysis of Financial Time Series*. Wiley, 3rd edition, 2010. C155

[13] K. Zhu and S. Ling. The global weighted LAD estimators for finite/infinite variance ARMA(p,q) models. *Economet. Theor.*, 28:1065–1086, 2012. doi:10.1017/S0266466612000059 C161

Author address

1. **A. W. Barker**, Department of Statistics, Macquarie University, North Ryde, New South Wales 2109, Australia.
mailto:adrian.barker@students.mq.edu.au