Chapter 1

GARCH models with outliers

1.1 Introduction

Outliers can disturbe the dynamic structure of a time series considerably. Since the estimation of an outlier model requires a large number of parameters, it is difficult to estimate such model from a classical point of view. It was shown by Verdinelli and Wasserman (1991) that the location shift outlier model can be estimated in a straightforward way by the Gibbs sampler which was developed by Gelfand and Smith (1990) for a wide class of Bayesian models. Bayesian time series models have been developed by Chib (1993), Albert and Chib (1993), and McCulloch and Tsay (1994). In this paper we show how the Bayesian outlier model can be combined with the general autoregressive heteroskedasticity (GARCH) model. ARCH and GARCH models are used in financial economics to model volatile time series. The variances in GARCH model are time dependent, i.e. are a function of past squared observations or errors (ARCH) or includes past variances as well (GARCH). We will use a Metropolis step within Gibbs sampling for the distribution of the ARCH parameters. An example involving the USD/YEN exchange rate will demonstrate our approach.

1.2 The AR-GARCH(p,q) model

The univariate GARCH model with AR structure in the mean is written as

$$y_t \sim N\left[\mathbf{x}_t'\beta, h_t = \mathbf{z}_t'\gamma\right], \qquad t = 1, \ldots, T, \tag{1.1}$$

where $\mathbf{x}_t' = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})$ contains p lagged observations and the prior information is specified by

$$\beta \sim N[\mathbf{b}_*, \mathbf{H}_*] \quad \text{and} \quad \gamma \sim N[\gamma_*, \mathbf{R}_*], \tag{1.2}$$

where the ARCH parameters are $\gamma = (\gamma_0,\ \gamma_1,\ \ldots,\ \gamma_q)'$, and $\mathbf{z}_t = (1, h_{t-1}, \cdots, h_{t-q})'$ contains q past variances.

1.3 The AR model with outliers

The univariate location shift regression AR(p) model with possible outliers is formulated as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{D}_\delta \mathbf{a} + \varepsilon \tag{1.3}$$

or

$$\mathbf{y} \sim N[\mathbf{X}\beta + \mathbf{D}_\delta \mathbf{a}, \sigma^2 \mathbf{I}_T], \tag{1.4}$$

where $\mathbf{D}_\delta = diag(\delta_1, \delta_2, \cdots, \delta_T)$ is a diagonal matrix with location indicators $\delta_{\mathbf{t}}$ on its diagonal, $\mathbf{a}$ is the the $n \times 1$ vector of shift parameter. The location shift regression model can also be written as

$$f(y_t|\theta) = (1-\varepsilon)N[y_t|\mathbf{x}_t'\beta, \sigma^2] + \varepsilon N[y_t|\mathbf{x}_t'\beta + a_t, \sigma^2]. \tag{1.5}$$

Each observation has an ε chance of being an outlier. The location indicator ϑ_t follows a Bernoulli distribution with probability ε: $\vartheta_t \sim Ber(\varepsilon)$. When ϑ_t is 1, y_t is an outlier, and when ϑ_t is 0, y_t is not an outlier.

1.4 The AR-GARCH model with outliers

The location shift AR-GARCH outliers model is given by the following switching model for outliers

$$y_t \sim \begin{cases} N[\mathbf{x}_t'\beta, h_t = \mathbf{z}_t'\gamma] & \text{with prob. } \epsilon, \\ \\ N[\mathbf{x}_t'\beta + a_t, h_t = \mathbf{z}_t'\gamma] & \text{with prob. } 1-\epsilon. \end{cases} \tag{1.6}$$

The location shift AR-GARCH model has the joint density

$$\begin{aligned} p(\theta, \mathbf{y}) \quad = \quad & \prod_{t=1}^{T} \{N[y_t|\mathbf{x}_t'\beta + \delta_t a_t, h_t] \cdot Ber[\delta_t|\epsilon]\} \cdot N[\beta|\mathbf{b}_*, \mathbf{H}_*] & (1.7) \\ & \cdot N[\mathbf{a}|\mathbf{a}_*, \mathbf{G}_*] \cdot N[\gamma|\gamma_*, \mathbf{r}_*]. & (1.8) \end{aligned}$$

The parameter vector is $\theta = (\beta, \sigma^2, \delta, \mathbf{a})$, and the full conditional distributions (f.c.d.'s) for θ are the following:

a) The f.c.d. for the AR coefficients β

Let θ^c denote all parameters except the current argument, then the f.c.d. of the posterior distribution is given by

$$p(\beta|\theta^c, \mathbf{y}) \quad = \quad N\left[\mathbf{b}_{**}, \mathbf{H}_{**}\right], \tag{1.9}$$

is a normal distribution with the parameters

$$\mathbf{H}_{**}^{-1} \quad = \quad \mathbf{H}_{*}^{-1} + \mathbf{X}'\mathbf{D}_h^{-1}\mathbf{X}, \tag{1.10}$$

$$\mathbf{D}_h \quad = \quad diag(h_1, \ldots, h_T), \tag{1.11}$$

$$\mathbf{b}_{**} \quad = \quad \mathbf{H}_{**}\left[\mathbf{H}_*\mathbf{b}_* + \mathbf{X}'\mathbf{D}_h^{-1}\left(\mathbf{y} - \mathbf{D}_\delta \mathbf{a}\right)\right]. \tag{1.12}$$

b) The f.c.d. for the shift parameter a

$$p(\mathbf{a}|\theta^c, \mathbf{y}) \quad = \quad N\left[\mathbf{a}_{**}, \mathbf{G}_{**}\right], \tag{1.13}$$

is a normal distribution with the parameters

$$\mathbf{G}_{**}^{-1} \quad = \quad \mathbf{G}_{*}^{-1} + \mathbf{D}_\delta'\mathbf{D}_h^{-1}\mathbf{D}_\delta, \tag{1.14}$$

$$\mathbf{a}_{**} \quad = \quad \mathbf{G}_{**}\left[\mathbf{G}_*\mathbf{a}_* + \mathbf{D}_\delta'\mathbf{D}_h^{-1}\left(\mathbf{y} - \mathbf{X}\mathbf{b}\right)\right]. \tag{1.15}$$

c) The f.c.d. for the indicator variables δ_t

$$p(\delta_t|\theta^c, \mathbf{y}) \quad = \quad Ber\left[\epsilon_t = \frac{c_t}{c_t + d_t}\right], \tag{1.16}$$

$$\tag{1.17}$$

is a Bernoulli distribution with the components

$$c_t = \epsilon_* \Phi\left(\frac{y_t - \mathbf{x}_t'\beta - a_t}{h_t}\right) \quad \text{and} \quad d_t = (1 - \epsilon_*)\Phi\left(\frac{y_t - \mathbf{x}_t'\beta}{h_t}\right). \tag{1.18}$$

d) The f.c.d. for the ARCH parameter γ is

$$p(\gamma|\theta^c, \mathbf{y}) \quad \propto \quad \prod_{t=1}^{T}\left\{\frac{1}{2\sqrt{\mathbf{z}_t'\gamma}}\exp\left\{-\frac{(y_t - \mathbf{x}_t'\beta - \delta_t a_t)^2}{2\mathbf{z}_t'\gamma}\right\}\right\} \tag{1.19}$$

$$\cdot \exp\left\{-\frac{(\gamma - \gamma_*)'\mathbf{R}^{-1}(\gamma - \gamma_*)}{2}\right\}. \tag{1.20}$$

In the simulation we use for the proposal density in the Metropolis step

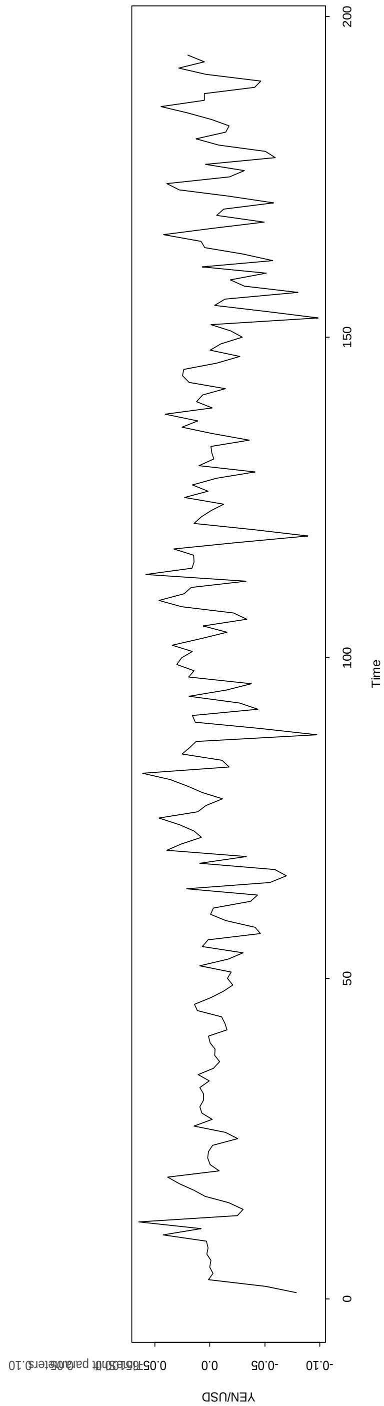
$$\gamma \sim N\left[\mathbf{1}/(q+1), \mathbf{I}/q\right]. \tag{1.21}$$

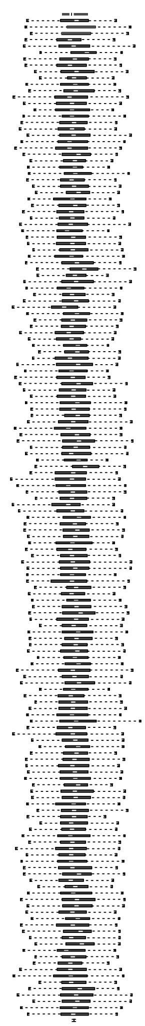
1.5 Example

We have analyzed the monthly exchange rate of the Japanese YEN to the US Dollar from January 1974 to February 1989 for ARCH effects and outliers.
The estimation result for the AR outlier model of order 1 are shown in Figure 1.1 to 1.4. In the first panel of Figure 1.1, we have displayed the time series, in the second the series of boxplots of the shift parameters a_i and in the last panel the probability that the i-th observation is an outlier. We can see 8 observations which have a posterior probability (the Bernoulli parameter ϵ_{i**}) larger then 1/2 of being an outlier. Figure 1.2 shows the same type of information for the AR-GARCH outlier model. Now we can see that the number of outliers is reduced to 4 with a 5th observation just hitting the 1/2-line. This reduction can be partly expected, since the GARCH outlier model can produce volatility clusters, a phenomenon which is similar to an outlier effect. A location shift in an outlier model can be explained by a data generating process which has a large variance at this location.
Figure 1.3 shows the unconditional residual variances for 3 models: the AR outlier model, the AR-GARCH, and the AR-GARCH outlier model. The variance of the GARCH model can be derived from the Gibbs output by $\sigma^2 = \gamma_0/(1-\gamma_1-\cdots-\gamma_q)$. We see the largest residual variance for the AR-outlier model and similier residual variance distribution for the AR-GARCH model with or without outliers. The mean and variance of the residual variance distribution is $(0.0035, 0.028^2)$ for the AR outlier model, while it is $(0.0026, 0.014^2)$ for the AR-GARCH model. The outlier AR-GARCH model is $(0.0028, 0.013^2)$ distributed, i.e. an increase in the mean and a decrease in the variance. This behaviour is confirmed by the behaviour of the predictive distributions which are shown in Figure 1.4. The predictive distribution for the AR-GARCH outlier model makes the tightest impression while the model without outliers has a small second mode at the the upper end of the distribution. The effect of the outlier model can be also seen in the estimated distribution of the GARCH coefficients which are less spread out.
A second example for the Deutsche Mark and US Dollar exchange rate has produced similar results.

1.6 Conclusion

This chapter has shown that outliers should be considered additionally in GARCH models since the estimation performance can be improved quite considerably. The Gibbs sampler with a Metropolis step was estimated by the software by Jin at al (1995). The approach might be extended to multivariate GARCH models (see Polasek and Jin (1994)) or semi-parametric ARCH models (see Kozumi and Polasek 1996).


YEN/USD
Time



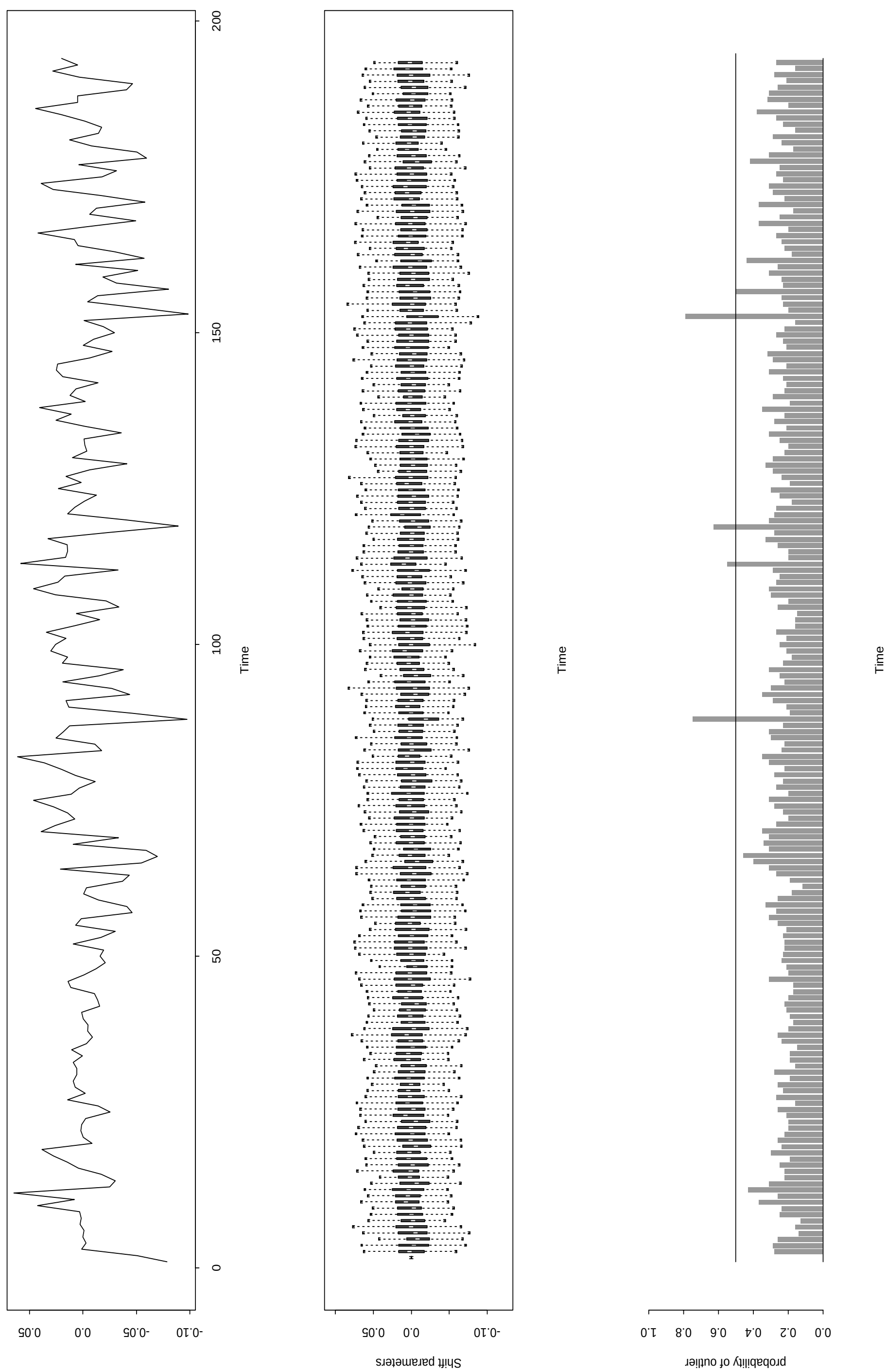


Figure 1.2: AR-ARCH-Outlier model with YEN/USD

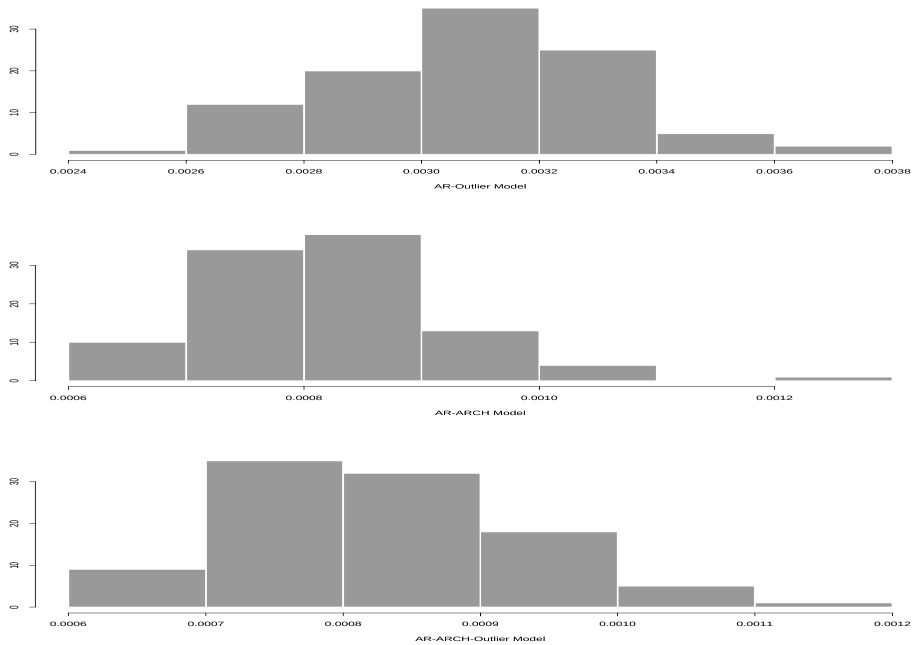


Figure 1.3: σ^2 for three models

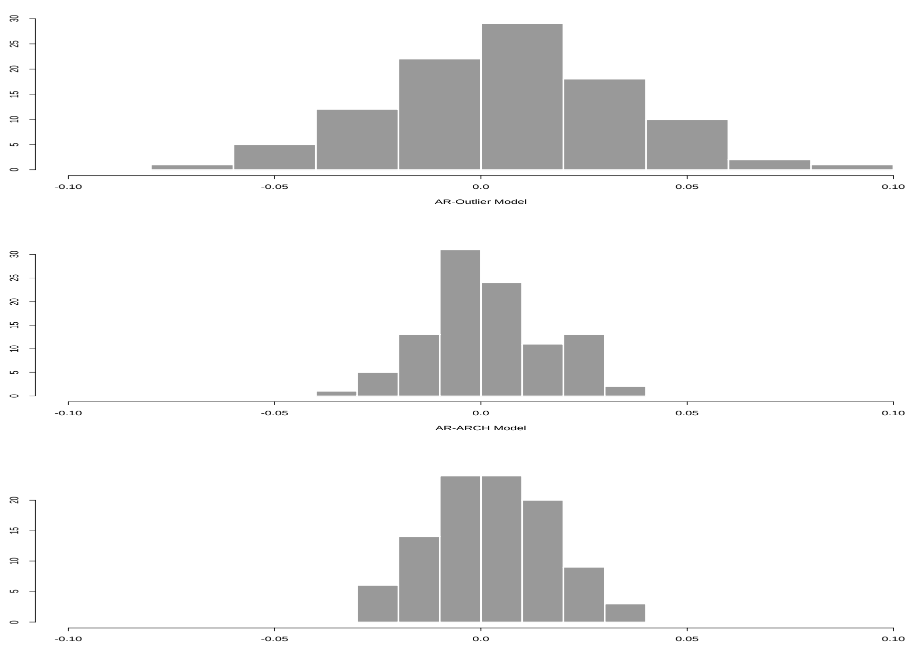


Figure 1.4: Predictive distribution