Fitting Weibull ACD Models to high frequency transactions data : a semi-parametric approach based on estimating functions

Kok Haur Ng

David E. Allen

Shelton Peiris

Fitting Weibull ACD Models to High Frequency Transactions Data: A Semi-parametric Approach based on Estimating Functions

By

Kok Haur Ng[1], David E. Allen[2] and Shelton Peiris[3]

[1]Institute of Mathematical Sciences, University of Malaya

[2]School of Accounting, Finance and Economics, Edith Cowan University

[3]Department of Mathematics and Statistics, University of Sydney

School of Accounting, Finance and Economics & FEMARC Working Paper Series
Edith Cowan University
March 2009
Working Paper 0901

Correspondence author:

David E. Allen
School of Accounting, Finance and Economics
Faculty of Business and Law
Edith Cowan University
Joondalup, WA 6027
Australia
Phone: +618 6304 5471
Fax: +618 6304 5271
Email: d.allen@ecu.edu.au

ABSTRACT

Autoregressive conditional duration (ACD) models play an important role in financial modeling. This paper considers the estimation of the Weibull ACD model using a semi-parametric approach based on the theory of estimating functions (EF). We apply the EF and the maximum likelihood (ML) methods to a data set given in Tsay (2003, p203) to compare these two methods. It is shown that the EF approach is easier to apply in practice and gives better estimates than the MLE. Results show that the EF approach is compatible with the ML method in parameter estimation. Furthermore, the computation speed for the EF approach is much faster than for the MLE and therefore offers a significant reduction of the completion time.

Keywords: Weibull distribution; Autoregressive conditional duration; Estimating function; Maximum likelihood; Semi-parametric; High frequency data

1. Introduction

In financial modeling, one problem we face is the analysis of high frequency transaction data. The main characteristic of this type of data is that it is collected at irregular, short time intervals. A basic tool used to study such duration data is the use of autoregressive conditional duration (ACD) models given by Engle and Russell (1998).

The general class of ACD models adapts the AR and GARCH theory to study the dynamic structure of the adjusted durations $\{x_i\}$ $(x_i = t_i - t_{i-1})$, where t_i is the time at the ith transaction. A crucial assumption underlying the ACD model is that the time dependence is described by a function ψ_i, where ψ_i is the conditional expectation of the adjusted duration between the $(i-1)$th and the ith trades.

Let

$$\psi_i = E[x_i \mid x_{i-1}, \cdots, x_1] = E[x_i \mid F_{i-1}], \tag{1.1}$$

where F_{i-1} is the information set available at the $(i-1)$th trade.

The basic ACD model is defined as

$$x_i = \psi_i \varepsilon_i, \tag{1.2}$$

where $\{\varepsilon_i\}$ is a sequence of iid non-negative random variable's with density $f(.)$ and $E(\varepsilon_i) = 1$. Also note that ε_i is independent of F_{i-1}. From Equation (1.2) it is clear that a vast set of ACD model specifications can be defined by different distributions of ε_i and specifications of ψ_i.

Since the durations are non-negative variables, in practice, we use the distributions such as the Exponential, Gamma and Weibull to model ACD structures (see, Peiris et al (2008) for details). The Weibull distribution is more flexible and therefore plays an important role in ACD modelling. Since the Exponential and Gamma distributions are special cases of the Weibull distribution, below we give the corresponding Weibull density and other useful results for later reference.

The Weibull Distribution

A random variable X has a Weibull distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ if its cumulative distribution function (cdf) and probability density function (pdf) are given by

$$F(x \mid \alpha, \beta) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-(x/\beta)^{\alpha}} & \text{if } x \geq 0 \end{cases}$$

and

$$f(x \mid \alpha, \beta) = \begin{cases} \dfrac{\alpha}{\beta^{\alpha}} x^{\alpha - 1} e^{-(x/\beta)^{\alpha}} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.3}$$

respectively. When $\alpha = 1$, the Weibull distribution reduces to an exponential distribution.

The pdf of the standardized Weibull distribution is

$$f(y \mid \alpha) = \begin{cases} \alpha \left[\Gamma\left(1 + \dfrac{1}{\alpha}\right) \right]^{\alpha} y^{\alpha - 1} \exp\left\{ -\left[\Gamma\left(1 + \dfrac{1}{\alpha}\right) y \right]^{\alpha} \right\} & ; y \geq 0 \\ 0 & ; \text{otherwise} \end{cases} \tag{1.4}$$

Notice that the scale parameter β not appears in (1.4). It can be seen that $E(Y) = 1$ and

$$V = Var(Y) = \frac{\Gamma\left(1 + \dfrac{2}{\alpha}\right)}{\left[\Gamma\left(1 + \dfrac{1}{\alpha}\right) \right]^2} - 1. \tag{1.5}$$

The corresponding cdf is

$$F(y \mid \alpha) = \begin{cases} 1 - \exp\left\{ -\left[\Gamma\left(1 + \dfrac{1}{\alpha}\right) y \right]^{\alpha} \right\} & ; y \geq 0 \\ 0 & ; y < 0 \end{cases} \tag{1.6}$$

The Section 2 reviews the general ACD model and its basic properties for later reference.

2. A Review of the General ACD(m,q), $q \geq 0$ Model

Suppose that only the most recent m durations $(m \geq 1)$ influence the conditional duration ψ_i in (1.1) and consider the model satisfying

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} \ ,$$

where $\omega > 0$, $\alpha_j > 0$ and $\sum_{j=1}^{m} \alpha_j < 1$. This is called an ACD$(m)$ model.

If there is no limited-memory characteristic, then one can define a more general class called ACD(m,q), $q \geq 1$ model as given in Engle and Russell (1988)

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \ , \tag{2.1}$$

where $\omega > 0$, $\alpha_j, \beta_j > 0$. It is easy to see that $\eta_i = x_i - \psi_i$ is a martingale difference sequence and the model in (2.1) can be written as

$$x_i - \eta_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j (x_{i-j} - \eta_{i-j})$$

and consequently

$$x_i = \omega + \sum_{j=1}^{r} (\alpha_j + \beta_j) x_{i-j} - \sum_{j=1}^{q} \beta_j \eta_{i-j} + \eta_i \ , \tag{2.2}$$

where $r = \max(m,q)$ and $\sum_{j=1}^{r} (\alpha_j + \beta_j) < 1$.

This is in the form of an ARMA process with non-Gaussian innovations. This representation is used to obtain the unconditional mean and variance of the ACD model in (2.1). Notice that $\{x_i\}$ is weakly stationary provided the zeroes of $\phi(z) = 1 - \sum_{j=1}^{r} \delta_j z^j$ are outside the unit circle, where $\delta_j = \alpha_j + \beta_j$, $j = 1, \cdots, r$.

If the parameters in the model are not well-estimated, then the model is not adequate for describing the behavior of the data and the accuracy of forecasts will be affected. The most common method of estimating the parameters is the use of maximum likelihood (ML). For example, see Engle and Russell (1998), Bauwens and Giot (2000), Zhang, Russell and Tsay (2001). This paper applies an alternative method of parameter estimation that is based on the EF approach due to Godambe (1985). In their paper Thavaneswaran and Peiris (1996) used the EF approach for estimating some nonlinear time series models. Peiris and Ng (2008) used this EF approach in parameter estimation of autogressive models with non-stationary innovations. Recently, Peiris, Ng and Mohamed (2008) compared the performance of the EF and ML estimates of simple exponential ACD models and showed that the EF method is more efficient than the ML method. Using a large scale simulation study Allen, Peiris and Ng (2008) showed that the parameter estimates based on EF method outperforms the ML estimates in Weibull ACD models.

With that view in mind the section 3 reviews the MLE and EF estimation procedures in detail for ACD modelling.

3. Parameter Estimation

We first review the maximum likelihood (ML) approach.

3.1 The MLE Approach

For an ACD(m,q) model, let $i_0 = \max(m,q)$ and $\mathbf{x}_{\mathrm{N(T)}} = (x_1,\cdots,x_{\mathrm{N(T)}})'$, where $\mathrm{N(T)}$ is the sample size. The likelihood function of the durations $x_1,\cdots,x_{\mathrm{N(T)}}$ is

$$\begin{aligned} L(\mathbf{x}_{\mathrm{N(T)}} \mid \theta, \mathbf{x}_{i_0}) &= \prod_{i=1}^{\mathrm{N(T)}} f(x_i \mid F_{i-1}, \theta) \\ &= \left[\prod_{i=i_0+1}^{\mathrm{N(T)}} f(x_i \mid F_{i-1}, \theta) \right] f(\mathbf{x}_{i_0} \mid \theta), \end{aligned}$$

where θ denotes the vector of model parameters, $\mathbf{x}_{i_0} = (x_1,\cdots,x_{i_0})$ and $f(\mathbf{x}_{i_0} \mid \theta) = \prod_{i=1}^{i_0} f(x_i)$.

The impact of the marginal pdf $f(\mathbf{x}_{i_0} \mid \theta)$ on the likelihood function diminishes as the sample size N(T) increases and so the marginal density can be ignored, resulting in the conditional likelihood function

$$L(\mathbf{x}_{\mathrm{N(T)}} \mid \theta, \mathbf{x}_{i_0}) = \prod_{i=i_0+1}^{\mathrm{N(T)}} f(x_i \mid F_{i-1}, \theta). \tag{3.1}$$

Estimating the Weibull ACD model

In the Weibull ACD Model, the $\{\varepsilon_i\}$ follows the standardised Weibull distribution with $F_\varepsilon(\varepsilon) = 1 - \exp\left\{-\left[\Gamma\left(1+\frac{1}{\alpha}\right)\varepsilon\right]^\alpha\right\}$. From Equation (1.2), we have $F_E(\varepsilon) = 1 - \exp\left\{-\left[\Gamma\left(1+\frac{1}{\alpha}\right)\frac{x}{\psi}\right]^\alpha\right\}$.

The corresponding conditional log likelihood function is given by

$$L(\mathbf{x} \mid \alpha, \mathbf{x}_{i_0}) = \prod_{i=i_0+1}^{\mathrm{N(T)}} \frac{\alpha}{\psi_i} \Gamma\left(1+\frac{1}{\alpha}\right)^\alpha \left(\frac{x_i}{\psi_i}\right)^{\alpha-1} \exp\left\{-\left[\Gamma\left(1+\frac{1}{\alpha}\right)\left(\frac{x_i}{\psi_i}\right)\right]^\alpha\right\}$$

So taking logs

$$l(\mathbf{x} \mid \alpha, \mathbf{x}_{i_0}) = \sum_{i=i_0+1}^{\mathrm{N(T)}} \left\{\ln\left(\frac{\alpha}{x_i}\right) + \alpha \ln\left[\Gamma\left(1+\frac{1}{\alpha}\right)\right] + \alpha \ln\left(\frac{x_i}{\psi_i}\right) - \left[\frac{1}{\psi_i}\Gamma\left(1+\frac{1}{\alpha}\right)x_i\right]^\alpha\right\}. \tag{3.2}$$

see Tsay (2002). Further examples can be found in Peiris et.al. (2005).

Now we review the theory of estimating functions (EF) as an alternative semi-parametric approach in parameter estimation.

3.2 The EF Approach

Suppose that $\{y_1, y_2, \cdots\}$ is a discrete stochastic process. We are interested of fitting a suitable model for a sample of size n from this process. Let Θ be a class of probability distributions F on R^n and $\theta = \theta(F)$, $F \in \Theta$ be a vector of real parameters.

Let h_i be a real valued function of $y_1, y_2, \cdots, y_i$ and θ such that

$$E_{i-1,F}[h_i\{y_1, y_2, \cdots, y_i; \theta(F)\}] = 0, \qquad (i = 1,2,\cdots,n; F \in \Theta)$$

and

$$E(h_i h_j) = 0, \qquad (i \neq j),$$

where $E_{i-1,F}(.)$ denotes the expectation holding the first $i-1$ values $y_1, y_2, \cdots, y_{i-1}$ fixed and $E_{i-1,F}(.) \equiv E_{i-1}$, $E_{0,F}(.) \equiv E_F(.) \equiv E(.)$ (unconditional mean).

Estimating Functions

Any real valued function $g(.)$ of the random variates $y_1, y_2, \cdots, y_n$ and the parameter θ, that can be used to estimate θ is called an estimating function.

In addition, if $g(.)$ satisfies some regularity conditions (ie. (i) the first and the second derivatives of $g(.)$ ($g^{'}(.)$ and $g^{''}(.)$) exist and (ii) $E[g^2(.)]$ is non-zero) and

$$E[g(y_1, y_2, \cdots, y_n; \theta(F))] = 0$$

then $g(.)$ is called a regular unbiased estimating function.

Among all regular unbiased estimating functions g, g^* is said to be optimum if

$$\frac{E[g^2(y_1, y_2, \cdots, y_n; \theta(F))]}{\left\{E\left(\left[\frac{\partial g(y_1, y_2, \cdots, y_n; \theta(F))}{\partial \theta}\right]_{\theta=\theta(F)}\right)\right\}^2} \tag{3.3}$$

is minimized for all $F \in \Theta$ at $g = g^*$.

We then estimate θ by solving the optimum estimating equations

$$g^*(y_1, y_2, \cdots, y_n; \theta) = 0.$$

Main Results

We consider the class of linear estimating functions L generated by

$$g = \sum_{i=1}^{n} h_i a_{i-1}$$

where h_i are as defined before and a_{i-1} is a suitably chosen function of the random variates $y_1, y_2, \cdots, y_{i-1}$ and the parameter θ for all $i = 1,2,\cdots,n$. Clearly,

$$E(g) = 0,\ g \in L.$$

Now we state the following theorem due to Godambe (1985):

Theorem

In the class L of estimating functions g, the function g^* minimizing (3.3) is given by

$$g^* = \sum_{i=1}^{n} h_i a_{i-1}^*,$$

where

$$a_{i-1}^* = \frac{E_{i-1}\left[\dfrac{\partial h_i}{\partial \theta}\right]}{E_{i-1}[h_i^2]}.$$

Notes:

1. The function g^* is called the optimum estimating function.
2. An optimal estimate of θ (in the sense of Godambe(1985)) can be obtained by solving the equation(s) $g^* = 0$.

Estimation of ACD (1,1) Using the EF Approach

Let $\psi_i = E(x_i \mid x_{i-1}, x_{i-2}, \cdots, x_1)$. Consider the ACD (1,1) model given by

$$x_i = \psi_i \varepsilon_i, \tag{3.4}$$

with

$$\psi_i = \omega + a x_{i-1} + b \psi_{i-1}, \tag{3.5}$$

where $\{\varepsilon_i\}$ is a sequence of iid standard Weibull random variables with $E(\varepsilon_i) = 1$ & $Var(\varepsilon_i) = V$ and $\omega > 0$, $a, b > 0$ such that $a + b < 1$.

It is clear that the conditional distribution

$$x_i \mid \Omega_{i-1} \sim (\psi_i, \psi_i^2 V),$$

where Ω_{i-1} is the information set available at time $i-1$, $V = Var(\varepsilon_i)$, and V is given in (1.5).

Let $h_i = \psi_i - x_i$. Then clearly, h_i is an unbiased estimating function. Now we construct a linear unbiased estimating function such that

$$g = \sum_{i=1}^{n} h_i a_{i-1}^* ,$$

where n is the number of observations.

It can be seen that the optimal value of a_{i-1} in the sense of Godambe (1985) is given by

$$a_{i-1}^* = \frac{\dfrac{\partial \psi_i}{\partial \theta}}{\psi_i^2 V} ,$$

where θ is a parameter.

Solving the system of equations

$$\sum_{i=1}^{n} \frac{1}{\psi_i^2 V} \frac{\partial \psi_i}{\partial \theta} (\psi_i - x_i) = 0 \qquad (3.6)$$

for $\theta = (\omega, a, b)$ the corresponding optimal set of estimates can be obtained. The following derivatives under the conditions of second order stationarity can be used:

- $\dfrac{\partial \psi_i}{\partial \omega} = 1 + b \dfrac{\partial \psi_{i-1}}{\partial \omega}$ or $\dfrac{\partial \psi_i}{\partial \omega} = \dfrac{1}{1-b}$
- $\dfrac{\partial \psi_i}{\partial a} = x_{i-1} + \beta \dfrac{\partial \psi_{i-1}}{\partial a}$
- $\dfrac{\partial \psi_i}{\partial b} = \psi_{i-1} + b \dfrac{\partial \psi_{i-1}}{\partial b}$.

Since these equations do not estimate V, an estimate of α is obtained by solving

$$\frac{\Gamma\left(1 + \dfrac{2}{\alpha}\right)}{\left[\Gamma\left(1 + \dfrac{1}{\alpha}\right)\right]^2} = \frac{(1 - b^2 - 2ab)(\mathrm{var}(x) + [E(x)]^2)}{a^2\ \mathrm{var}(x) + [E(x)]^2 (1 - b^2 - 2ab)}$$

The Section 4 applies these two approaches for a real data set from Tsay (2002) and compares the corresponding EF and ML estimates.

4. An Application of ACD Modelling

The data set used in this paper is based on a sample of high frequency transactions data obtained for the US IBM stock on five consecutive trading days from November 1 to November 7, 1990 (see Tsay (2003, p203)). Focusing on positive transaction durations, we have 3534 observations. The series is then adjusted (see Tsay (2003, p195-197) such that we obtain 3534 positive adjusted durations. Figures 1 to 3 are respectively the series, the histogram of the series and the autocorrelation (ACF) of the series. Based on Figure 3, there exist some serial correlations in the adjusted durations. Now we fit the series with Weibull ACD (1,1) model as shown in Tsay (2003, p2003) and estimate the following two Weibull models.

Model 1 (based on ML method):

$$x_i = \psi_i \varepsilon_i, \qquad \psi_i = 0.2085 + 0.0693 x_{i-1} + 0.8679 \psi_i$$

$$\hat{\alpha} = 0.8781.$$

Model 2 (based on EF method):

$$x_i = \psi_i \varepsilon_i, \qquad \psi_i = 0.2296 + 0.0712 x_{i-1} + 0.8602 \psi_i$$

$$\hat{\alpha} = 0.7764.$$

where ε_i is follow the standardized Weibull distribution with parameter $\hat{\alpha}$.

To assess the performance of ML and EF methods given in Section (3.1) and (3.2) on this two models, the standard errors were computed. Standard errors of ω, a, b, α for the Model 1 are $0.0570, 0.0114, 0.0248$ and 0.0115 respectively. The standard errors of ω, a, b, α for the Model 2 are $0.0620, 0.0117, 0.0263$ and 0.0203. The EF method in general is comparable to the ML method in term of parameter estimates and standard errors. Furthermore, we note that if we use the ML method to find the estimates, the method needs to search for the maximum value under the maximum likelihood procedure. One the other hand, the EF approach is just solving the simultaneous equations to obtain the estimates. Thus, we would expect a reduction in computation time if we use EF method instead of that based on the ML method. The reason is that the EF method is only involved in solving the simultaneous nonlinear equations while the ML method needs to search for the maximum value of likelihood function. It is important to note that the EF method requires 9.5313 seconds in a Core 2 Duo 2.2 GHz computer to obtain the solution while the ML method requires 41.2187 seconds.

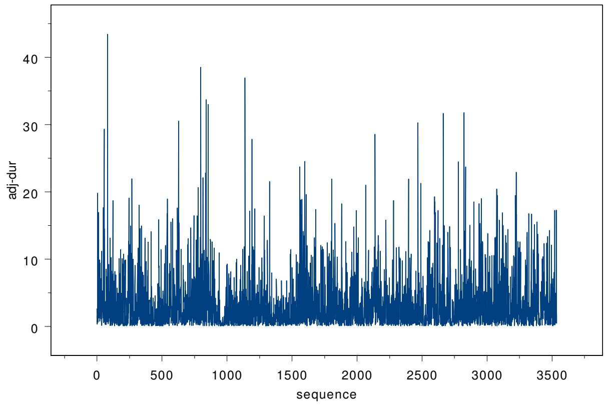

Figure 1: Time plots of durations for IBM stock traded in the first five trading days of November 1990: the adjusted series.

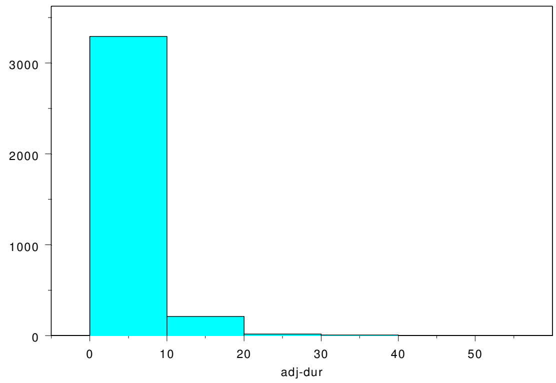

Figure 2: The histogram of the adjusted series.

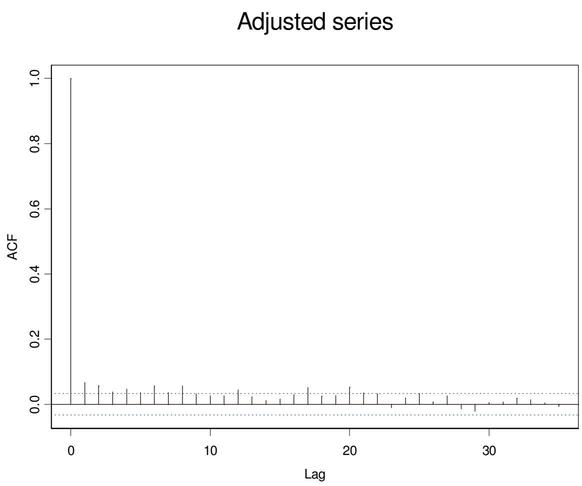

Figure 3: ACF of the adjusted series

5. Conclusion

This paper applied the EF approach in parameter estimation of Weibull ACD models and compared the properties with the corresponding ML estimates. Results show that the standard errors of the estimates using either EF or ML methods are comparable. However, the computation time for EF method is much shorter than that of the ML method.

References

[1] Allen, D., Chan, F., McAleer, M. and Peiris, M.S. (2008), Finite sample properties of the QMLE for the Log-ACD model: Application to Australian stocks, *Journal of Econometrics*, **147**, 163–185.

[2] Bauwens, L. and Giot, P. (2000), The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annales D'Economie et de Statistique* **60**, 117-145.

[3] Engle, R.F. and Russell, J.R. (1998), Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica*, **66**, 1127-1162.

[4] Engle.R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of variance of U.K. inflation, *Econometrica*, **31**, 987-1008.

[5] Godambe, V.P. (1985), The foundations of finite sample estimation in stochastic processes, *Biometrika*, **72**, 419-428.

[6] Peiris, M.S., Allen, D. and Yang, W. (2005), Some statistical models for durations and an application to news corporation stock prices, *Mathematics and Computers in Simulation*, **68**, 549-556.

[7] Peiris, M.S. , Ng, K.H., and Ibrahim, M. (2007), A review of Recent Developments of Financial Time Series: ACD Modelling using the Estimating Function Approach, *Sri Lankan Journal of Applied Statistics*, **8**:1-17.

[8] Peiris, M.S. and Ng, K.H. (2007), Optimal estimation of autoregressive models with non-stationary innovations: a simulation study. (*submmitted to International Journal of Modelling and Simulation*)

[9] Tsay, R.S. (2002), *Analysis of financial time series*, John Wiley & Inc.

[10] Peiris, M.S., Ng, K.H. and Allen, D. (2008), On estimation of Weibull ACD models using estimating functions: A simulation study. (submitted for publication)

[11] Thavaneswaran. A, Peiris, M.S.(1996), Nonparametric estimation for some nonlinear models, *Statistics and Probability Letters*, **28**, 227–233.

[12] Zhang, M.Y. Russell, J.R. and Tsay, R.S. (2001), A nonlinear autoregressive conditional duration model with applications to financial duration data, *Journal of Econometrics*, **104**, 179-207.