Lecture 15: Autoregressive Conditional Duration Models

Bus 41910, Time Series Analysis, Mr. R. Tsay

Note: This handout is based on a chapter I wrote for *Handbook in Econometrics II* by Palgrave Publishing Company in 2007.

1 Introduction

The autoregressive conditional duration (ACD) model was proposed by Engle and Russell (1998) to model irregularly spaced financial transaction data. It has attracted much interest among researchers and practitioners ever since, and has found many applications outside of modelling transaction data. Duration is commonly defined as the time interval between consecutive events, e.g., the time interval between two transactions of a stock in the New York Stock Exchange or the difference between arrival times of two customers at a service station. The duration between two consecutive transactions in finance is important, for it may signal the arrival of new information concerning the underlying asset. A cluster of short durations corresponds to active trading and, hence, an indication of the existence of new information.

Since duration is necessarily non-negative, the ACD model has also been used to model time series that consist of positive observations. An example is the daily range of the log price of an asset. The range of an asset price during a trading day can be used to measure its price volatility, e.g., Parkinson (1980). Therefore, studying range can serve as an alternative approach to volatility modeling. Chou (2005) considers a conditional autoregressive range (CARR) model and shows that his CARR model can improve volatility forecasts for the weekly log returns of the Standard and Poor's 500 index over some commonly used volatility models. The CARR model is essentially an ACD model.

In this chapter, we shall introduce the ACD model, discuss its properties, and address issues of statistical inference concerning the model. We then demonstrate its applications via some real examples. We also consider some extensions of the model, including nonlinear duration models and intervention analysis. Using the daily range of the log price of Apple stock, our ACD application shows that adopting the decimal system for U.S. stock prices on January 29, 2001 significantly reduces the volatility of the stock price.

2 Duration models

Duration models in finance are concerned with time intervals between trades. For a given asset, longer durations indicate lack of trading activities, which in turn signify a period of no new information. On the other hand, arrival of new information often results in heavy trading and, hence, leads to shorter durations. The dynamic behavior of durations thus contains useful information about market activities. Furthermore, since financial markets typically take a period of time to uncover the effect of new information, active trading is likely to persist for a period of time, resulting in clusters of short durations. Consequently, durations might exhibit characteristics similar to those of asset volatility. Considerations like this lead to the development of duration models. Indeed, to model the durations of intraday trading, Engle and Russell (1998) use an idea similar to that of the generalized autoregressive conditional heteroscedastic (GARCH) models to propose an autoregressive conditional duration (ACD) model and show that the model can successfully describe the evolution of time durations for (heavily traded) stocks. Since intraday transactions of a stock often exhibit certain diurnal patterns, adjusted time durations are used in ACD modeling. We shall discuss methods for adjusting the diurnal pattern later. Here we focus on introducing the ACD model.

Let t_i be the time, measured with respect to some origin, of the ith event of interest with t_0 being the starting time. The ith duration is defined as

$$x_i = t_i - t_{i-1}, \quad i = 1, 2, \ldots.$$

For simplicity, we ignore, at least for now, the case of zero durations so that $x_i > 0$ for all i. The ACD model postulates that x_i follows the model

$$x_i = \psi_i \epsilon_i \tag{1}$$

where $\{\epsilon_i\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with $E(\epsilon_i) = 1$ and positive support, and ψ_i satisfies

$$\psi_i = \alpha_0 + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{v=1}^{q} \beta_v \psi_{i-v}, \tag{2}$$

where p and q are non-negative integers and α_j and β_v are constant coefficients. Since x_i is positive, it is common to assume that $\alpha_0 > 0$, $\alpha_j \geq 0$ and $\beta_v \geq 0$ for $j \in \{1, \ldots, p\}$ and $v \in \{1, \ldots, q\}$. Furthermore, the zeros of the polynomial $\alpha(L) = 1 - \sum_{j=1}^{g} (\alpha_j + \beta_j) L^j$ are outside the unit circle, where L denotes the lag operator, $g = \max\{p, q\}$, and $\alpha_j = 0$ for $j > p$ and $\beta_j = 0$ for $j > q$.

Let F_h be the σ-field generated by $\{\epsilon_h, \epsilon_{h-1}, \ldots\}$. It is easy to see that $E(x_i | F_{i-1}) = \psi_i E(\epsilon_i | F_{i-1}) = \psi_i$. Thus, ψ_i is the conditional expected duration of the next transaction given F_{i-1}. Since ϵ_i has a positive support, it may assume the standard exponential distribution. This results in an exponential ACD model. For ease of reference, we shall refer to the model in (1)-(2) as an EACD(p, q) model when ϵ_i follows the standard exponential distribution.

2.1 Properties of EACD model

We start with the simple EACD(1,1) model

$$x_i = \psi_i \epsilon_i, \quad \psi_i = \alpha_0 + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1}. \tag{3}$$

Taking expectation of the model, we obtain

$$E(x_i) = E(\psi_i \epsilon_i) = E[\psi_i E(\epsilon_i | F_{i-1})] = E(\psi_i),$$

$$E(\psi_i) = \alpha_0 + \alpha_1 E(x_{i-1}) + \beta_1 E(\psi_{i-1}).$$

Under the weak stationarity assumption, $E(x_i) = E(x_{i-1})$, so that

$$\mu_x \equiv E(x_i) = E(\psi_i) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Consequently, $0 \leq \alpha_1 + \beta_1 < 1$ for a weakly stationary process $\{x_i\}$. Next, making use of the fact that $E(\epsilon_i) = 1$ and $E(\epsilon_i^2) = 2$, we have $E(x_i^2) = 2E(\psi_i^2)$. Again, under weak stationarity,

$$E(\psi_i^2) = \frac{\mu_x^2[1 - (\alpha_1 + \beta_1)^2]}{1 - 2\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}, \tag{4}$$

$$\text{Var}(x_i) = \frac{\mu_x^2(1 - \beta_1^2 - 2\alpha_1\beta_1)}{1 - 2\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}. \tag{5}$$

From these results, for the EACD(1,1) model to have a finite variance, we need $1 > 2\alpha_1^2 + \beta_1^2 + 2\alpha_1\beta_1$. Similar results can be obtained for the general EACD(p, q) model, but the algebra involved becomes tedious.

Forecasts from an EACD model can be obtained using a procedure similar to that of a GARCH model, which in turn is similar to that of a stationary autoregressive moving-average (ARMA) model. Again, consider the simple EACD(1,1) model and suppose that the forecast origin is $i = h$. For a 1-step ahead forecast, the model states that $x_{h+1} = \psi_{h+1}\epsilon_{h+1}$ with $\psi_{h+1} = \alpha_0 + \alpha_1 x_h + \beta_1 \psi_h$. Let $x_h(1)$ be the 1-step ahead forecast of x_{h+1} at the origin h. Then,

$$x_h(1) = E(x_{h+1}|F_h) = E(\psi_{h+1}\epsilon_{h+1}) = \psi_{h+1},$$

which is known at the origin $i = h$. The associated forecast error is $e_h(1) = x_{h+1} - x_h(1) = \psi_{h+1}(\epsilon_{h+1} - 1)$. The conditional variance of the forecast error is then ψ_{h+1}^2. For multi-step ahead forecasts, we use $x_{h+j} = \psi_{h+j}\epsilon_{h+j}$ so that, for $j = 2$,

$$\begin{aligned} \psi_{h+2} &= \alpha_0 + \alpha_1 x_{h+1} + \beta_1 \psi_{h+1} \\ &= \alpha_0 + (\alpha_1 + \beta_1)\psi_{h+1} + \alpha_1 \psi_{h+1}(\epsilon_{h+1} - 1). \end{aligned}$$

Consequently, the 2-step ahead forecast is

$$x_h(2) = E(\psi_{h+2}\epsilon_{h+2}) = \alpha_0 + (\alpha_1 + \beta_1)\psi_{h+1} = \alpha_0 + (\alpha_1 + \beta_1)x_h(1),$$

and the associated forecast error is

$$e_h(2) = \alpha_0(\epsilon_{h+2} - 1) + \alpha_1\psi_{h+1}(\epsilon_{h+2}\epsilon_{h+1} - 1) + \beta_1\psi_{h+1}(\epsilon_{h+2} - 1).$$

In general, we have

$$x_h(m) = \alpha_0 + (\alpha_1 + \beta_1)x_h(m-1), \quad m > 1.$$

This is exactly the recursive forecasting formula of an ARMA(1,1) model with AR polynomial $1 - (\alpha_1 + \beta_1)L$. By repeated substitutions, we can rewrite the forecasting formula as

$$x_h(m) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{m-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{m-1}x_h(1).$$

Since $\alpha_1 + \beta_1 < 1$, we have

$$x_h(m) \to \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \quad \text{as} \quad m \to \infty,$$

which says that, as expected, the long-term forecasts of a stationary series converge to its unconditional mean as the forecast horizon increases.

Let $\eta_j = x_j - \psi_j$. It is easy to show that $E(\eta_j) = 0$ and $E(\eta_j \eta_t) = 0$ for $t \neq j$. The variables $\{\eta_j\}$, however, are not identically distributed. Using $\psi_j = x_j - \eta_j$, we can rewrite the EACD(p, q) model in Eq. (2) as

$$x_i = \alpha_0 + \sum_{j=1}^{g}(\alpha_j + \beta_j)x_{i-j} + \eta_i - \sum_{j=1}^{q}\beta_j\eta_{i-j},$$

where $g = \max\{p, q\}$ and it is understood that $\alpha_j = 0$ for $j > p$ and $\beta_j = 0$ for $j > q$. This is in the form of an ARMA(g, q) model with AR polynomial $1 - \sum_{j=1}^{g}(\alpha_j + \beta_j)L^j$. Consequently, some properties of EACD models can be inferred from those of ARMA models.

2.2 Estimation of EACD models

Suppose that $\{x_1, \ldots, x_n\}$ represents a realization of an EACD(p, q) model. The parameter $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)'$ can be estimated by the conditional likelihood method. Again, let $g = \max\{p, q\}$. The likelihood function of the data is

$$f(\boldsymbol{x}_n|\boldsymbol{\theta}) = f(\boldsymbol{x}_g|\boldsymbol{\theta}) \times \prod_{i=g+1}^{n} f(x_i|\boldsymbol{x}_{i-1}, \boldsymbol{\theta})$$

where $\boldsymbol{x}_j = (x_1, \ldots, x_j)'$. Since the joint distribution of $\boldsymbol{x}_g$ is complicated and its influence on the overall likelihood function is diminishing as n increases, we adopt the conditional likelihood method by ignoring $f(\boldsymbol{x}_g|\boldsymbol{\theta})$. This results in using the conditional likelihood estimates. Since $f(x_i|F_{i-1}, \boldsymbol{\theta}) = \frac{1}{\psi_i}\exp(-x_i/\psi_i)$, the conditional log likelihood function of the data then becomes

$$\ell(\boldsymbol{\theta}|\boldsymbol{x}_n) = -\sum_{i=t_o+1}^{n}[\ln(\psi_i) + \frac{x_i}{\psi_i}]. \tag{6}$$

The usual asymptotics of maximum likelihood estimates apply when the process $\{x_i\}$ is weakly stationary.

2.3 Additional ACD models

The EACD model has several nice features. For instance, it is simple in theory and in ease of estimation. But the model also encounters some weaknesses. For example, the use of the exponential distribution implies that the model has a constant hazard function. In the statistical literature, the hazard function (or intensity function) of a random variable X is defined by

$$h(x) = \frac{f(x)}{S(x)},$$

where $f(x)$ and $S(x)$ are the probability density function and the survival function of X, respectively. The survival function of X is given by

$$S(x) = P(X > x) = 1 - P(X \leq x) = 1 - \text{CDF}(x), \quad x > 0,$$

which gives the probability that a subject, which follows the distribution of X, survives at the time x. Under the EACD model, the distribution of the innovations is standard exponential so that the hazard function of ϵ_i is 1. As mentioned before, transaction duration in finance is inversely related to trading intensity, which in turn depends on the arrival of new information, making it hard to justify that the hazard function of duration is constant over time.

To overcome this weakness, alternative innovational distributions have been proposed in the literature. Engle and Russell (1998) entertain the Weibull distribution for ϵ_i and Zhang, Russell and Tsay (2001) consider the generalized Gamma distribution. The probability density function of a standardized Weibull random variable X is

$$f(x|\alpha) = \begin{cases} \alpha \left[\Gamma\left(1+\frac{1}{\alpha}\right)\right]^{\alpha} x^{\alpha-1} \exp\left\{-\left[\Gamma\left(1+\frac{1}{\alpha}\right) y\right]^{\alpha}\right\}, & \text{if } x \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where the α is referred to as the shape parameter and $\Gamma(.)$ is the usual Gamma function. The mean and variance of X are $E(X) = 1$ and $\text{Var}(X) = \Gamma(1+2/\alpha)/[\Gamma(1+1/\alpha)]^2 - 1$. The hazard function of X is

$$h(x|\alpha) = \alpha \left[\Gamma\left(1+\frac{1}{\alpha}\right)\right]^{\alpha} x^{\alpha-1}.$$

Consequently, if $\alpha > 1$, the hazard function is a monotonously increasing function of x. If $0 < \alpha < 1$, then the hazard function is a monotonously decreasing function of x.

The probability density function of a generalized Gamma random variable X with $E(X) = 1$ is

$$f(x|\alpha,\kappa) = \begin{cases} \frac{\alpha x^{\kappa\alpha-1}}{\lambda^{\kappa\alpha}\Gamma(\kappa)} \exp\left[-\left(\frac{x}{\lambda}\right)^{\alpha}\right], & \text{if } x > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $\lambda = \Gamma(\kappa)/\Gamma(\kappa + 1/\alpha)$ with $\alpha > 0$ and $\kappa > 0$. Both α and κ are shape parameters so that the hazard function of X becomes more flexible than that of a Weibull distribution.

If ϵ_i of a duration model follows the standardized Weibull distribution with probability density function $f(x|\alpha)$ in Eq. (7), the conditional density function of x_i given F_{i-1} is

$$f(x, \alpha) = \alpha \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{\alpha} \frac{x_i^{\alpha-1}}{\psi_i^{\alpha}} \exp\left\{-\left[\frac{\Gamma\left(1+\frac{1}{\alpha}\right) x_i}{\psi_i}\right]^{\alpha}\right\} \tag{9}$$

which can be used to obtain the conditional log likelihood function of the data for estimation.

If ϵ_i of a duration model follows the generalized Gamma distribution with $E(\epsilon_i) = 1$ in Eq. (8), the conditional density function of x_i given F_{i-1} is

$$f(x_i|\alpha, \kappa) = \frac{\alpha x_i^{\kappa\alpha-1}}{(\psi_i \lambda)^{\kappa\alpha}\Gamma(\kappa)} \exp\left[-\left(\frac{x_i}{\psi_i \lambda}\right)^{\alpha}\right], \tag{10}$$

where, again, $\lambda = \Gamma(\kappa)/\Gamma(\kappa + 1/\alpha)$. This density function can be used to perform conditional maximum likelihood estimation of the model.

In what follows, we refer to the duration model in Eqs. (1)-(2) as the WACD(p, q) or GACD(p, q) model if the innovation ϵ_i follows the standardized Weibull or generalized Gamma distribution, respectively.

2.4 Quasi maximum likelihood estimates

In real applications, the true distribution function of the innovation ϵ_i of a duration model is unknown. One may, for simplicity, employ the conditional likelihood function of an EACD model in Eq. (6) to perform parameter estimation. The resulting estimates are called the quasi maximum likelihood estimates (QMLE). Engle and Russell (1998) show that, under some regularity conditions, QMLE of a duration model are consistent and asymptotically normal. They are, however, not efficient when the innovations are not exponentially distributed.

2.5 Model checking

Let $\hat{\psi}_i$ be the fitted value of the conditional expected duration of an ACD model. We define $\hat{\epsilon}_i = x_i/\hat{\psi}_i$ as the *standardized innovation* or *standardized residual* of the model. If the fitted ACD model is adequate, then $\{\hat{\epsilon}_i\}$ should behave as an i.i.d. sequence of random variables with the assumed distribution. We can use this standardized residual series to perform model checking. In particular, if the fitted model is adequate, both series $\{\hat{\epsilon}_i\}$ and $\{\hat{\epsilon}_i^2\}$ should have no serial correlations. The Ljung-Box statistics can be used to check the serial correlations of these two series. Large values of the Ljung-Box statistics indicate model inadequacy.

In addition, the quantile-to-quantile (QQ) plot of the standardized residuals against the assumed distribution of the innovations can be used to check the validity of the distributional assumption. For instance, under the WACD models, $\hat{\epsilon}_i$ should be close to the standardized

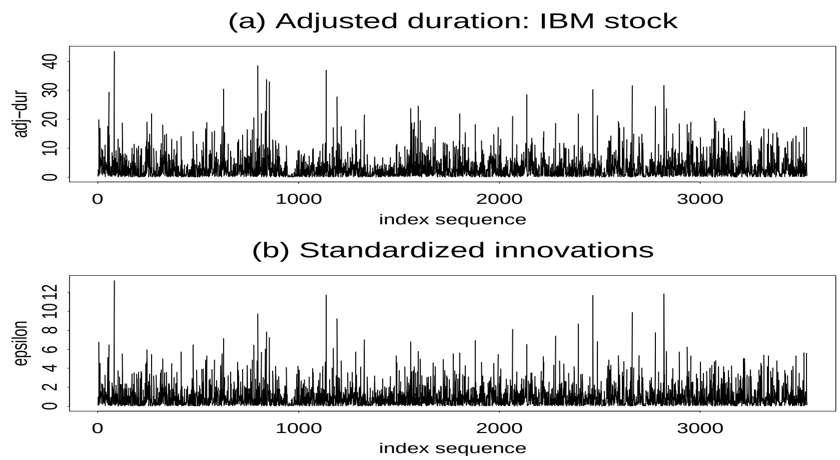

Figure 1: Time plots of the IBM transaction durations from November 1 to November 7, 1990: (a) adjusted durations, and (b) standardized innovations of a WACD(1,1) model.

Weibull distribution with shape parameter $\hat{\alpha}$. A deviation from the straight line of the QQ-plot suggests that the distributional assumption needs further improvement.

3 Some Simple Examples

In this section, we demonstrate the application of ACD models by considering two real examples.

Example 1. Consider the adjusted transaction durations of the IBM stock from November 1 to November 7, 1990. The original durations are time intervals between two consecutive trades measured in seconds. Overnight intervals and zero durations were ignored. The adjustment is made to take care of the diurnal pattern of daily trading activities. The series consists of 3534 observations and was used in Example 5.4 of Tsay (2005). Figure 1(a) shows the adjusted durations and Figure 2(a) gives the sample autocorrelation functions of the data. The autocorrelations are not large in magnitude, but they clearly indicate serial dependence in the data.

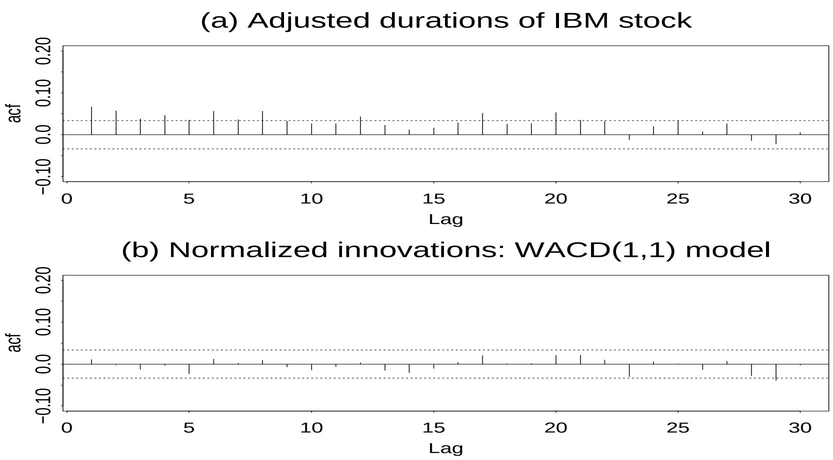

Figure 2: The sample autocorrelation function of IBM transaction durations from November 1 to November 7, 1990: (a) ACF of the adjusted durations, and (b) ACF of the standardized residual series of a WACD(1,1) model.

Table 1: Estimation results of EACD(1,1), WACD(1,1) and GACD(1,1) models for the IBM transaction durations of Example 1. The adjusted durations are from November 1 to November 7, 1990 with 3534 observations. The standard errors of the estimates are in parentheses. The p-values of the Ljung-Box statistics are also in parentheses with $Q(10)$ and $Q^*(10)$ for standardized residual series and its squared process, respectively.

Model	Parameters					Checking	
	α_0	α_1	β_1	α	κ	$Q(10)$	$Q^*(10)$
EACD	0.129	0.056	0.905			4.55	5.48
	(0.037)	(0.009)	(0.018)			(0.92)	(0.86)
WACD	0.125	0.056	0.906	0.880		3.85	5.51
	(0.040)	(0.010)	(0.019)	(0.012)		(0.92)	(0.85)
GACD	0.111	0.056	0.912	0.407	4.016	4.62	5.53
	(0.040)	(0.010)	(0.019)	(0.040)	(0.730)	(0.92)	(0.85)

For illustration, we entertain EACD(1,1), WACD(1,1) and GACD(1,1) models for the IBM transaction durations. The estimated parameters of the three models are given in Table 1. The estimates of the ACD equation are rather stable for all three models, consistent with the theory that the estimates based on the exponential likelihood function are QMLE. Figure 1(b) shows the standardized innovations and Figure 2(b) gives the sample autocorrelation function of the standardized innovations for the fitted WACD(1,1) model. The innovations appear to be random and their ACFs fail to indicate any serial dependence. Indeed, the Ljung-Box statistics for the standardized innovations and the squared innovations are insignificant, so that the fitted models are adequate in describing the dynamic dependence of the adjusted durations.

Figure 3 shows the QQ-plot of the standardized residuals versus a Weibull distribution with shape parameter 0.88 and scale parameter 1. The quantiles of the Weibull distribution are generated using a random sample of 30,000 observations. A straight line is imposed on the plot to aid interpretation. From the plot, except for a few large residuals, the assumption of a Weibull distribution seems reasonable. In this particular example, the GACD(1,1) model also fits the data well. We chose the WACD(1,1) model for its simplicity.

Finally, for the WACD(1,1) model, the estimated shape parameter α is less than one, indicating that the hazard function of the adjusted durations is monotonously decreasing. This seems reasonable for the adjusted durations of the heavily traded IBM stock.

Example 2. In this example, we apply the ACD model to stock volatility modeling. Consider the daily range of the log price of Apple stock from January 4, 1999 to November 20, 2007. The data are obtained from Yahoo Finance and consist of 2235 observations. The range has been used in the literature as a robust alternative to volatility modeling; see Chou (2005) and the references therein. Apple stock had two-for-one splits on June 21, 2000 and February 28, 2005 during the sample period, but for simplicity we make no

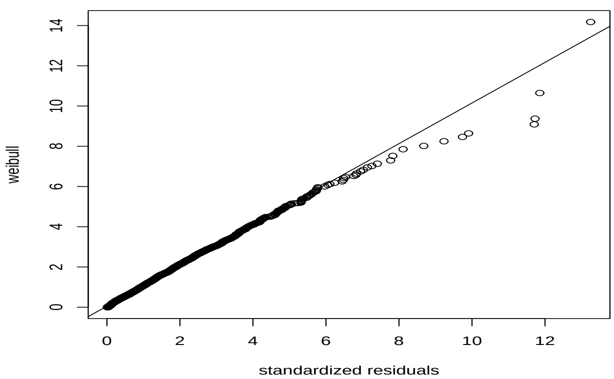

Figure 3: Quantile-to-quantile plot of the standardized residuals of the WACD(1,1) model versus a Weibull distribution. The Weibull quantiles are generated from a random sample of 30,000 observations using the shape parameter 0.88 and scale parameter 1.0.

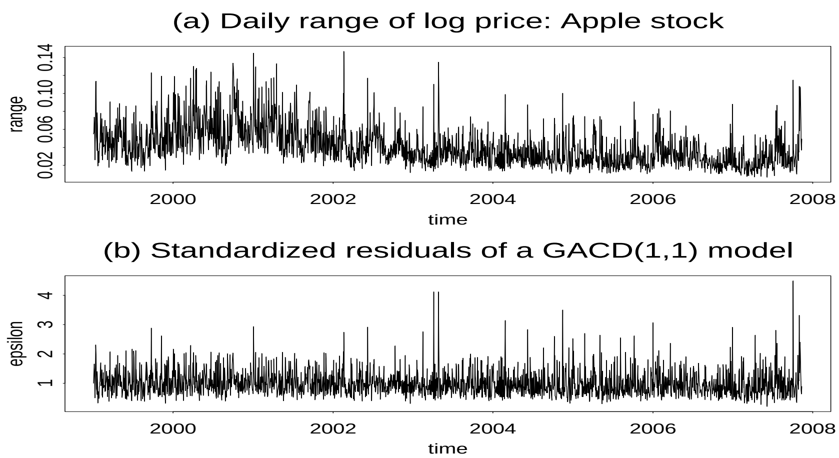

Figure 4: Time plots of the daily range of log price of Apple stock from January 4, 1999 to November 20, 2007: (a) Observed daily range and (b) standardized residuals of a GACD(1,1) model

adjustments for the splits. Also, stock prices in the U.S. markets switched from the tick size 1/16 of a dollar to the decimal system on January 29, 2001. Such a change affected the daily range of stock prices. We shall return to this point later. The sample mean, standard deviation, minimum and maximum of the range of log prices are 0.0407, 0.0218, 0.0068 and 0.1468, respectively. The sample skewness and excess kurtosis are 1.3 and 2.13, respectively. Figure 4(a) shows the time plot of the range series. The volatility seems to be increasing from 2000 to 2001, then deceasing to a stable level after 2002. It seems to increase somewhat at the end of the series. Figure 5(a) shows the sample ACF of the daily range series. The sample ACFs are highly significant and decay slowly.

Again, we fit the EACD(1,1), WACD(1,1), and GACD(1,1) models to the daily range series. The estimation results, along with the Ljung-Box statistics for the standardized residual series and its squared process, are given in Table 2. Again, the parameter estimates for the duration equation are stable for all three models, except for the constant term of the EACD model, which appears to be statistically insignificant at the usual 5% level. Indeed, in this particular instance, the EACD(1,1) model fares slightly worse than the other two

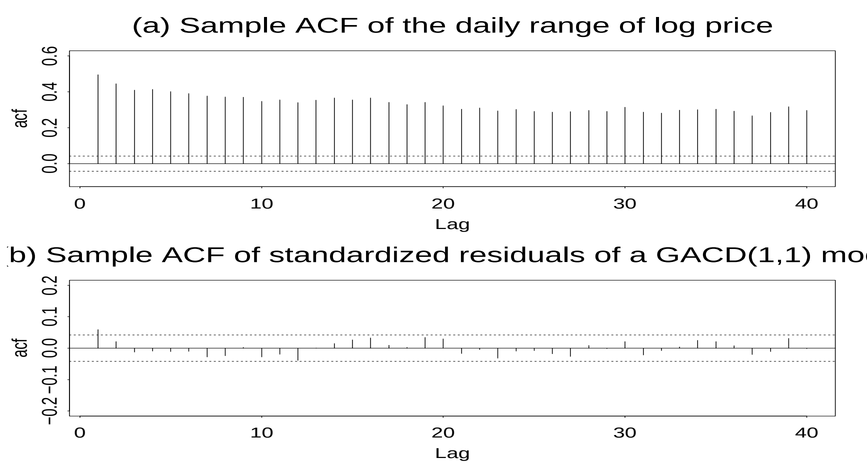

Figure 5: The sample autocorrelation function of the daily range of log price of Apple stock from January 4, 1999 to November 20, 2007: (a) ACF of daily range and (b) ACF of the standardized residual series of a GACD(1,1) model.

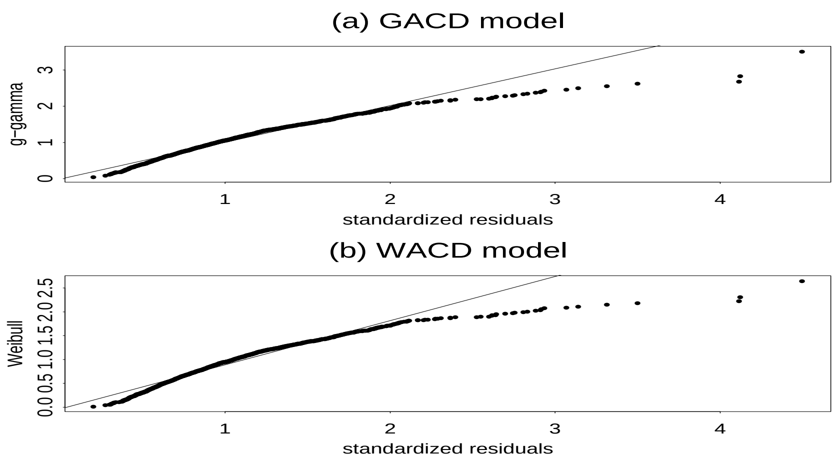

Figure 6: Quantile-to-quantile plots for the standardized residuals of ACD models for the daily range of log price of Apple stock from January 4, 1999 to November 20, 2007: (a) GACD(1,1) model and (b) WACD(1,1) model.

ACD models. Between the WACD(1,1) and GACD(1,1) models, we slightly prefer the GACD(1,1) model, because it fits the data better and is more flexible. Figure 6 shows the QQ-plots of the standardized residuals versus the assumed innovation distribution for the GACD(1,1) and WACD(1,1). The plots indicate that further improvement in the distributional assumption is needed for the daily range, but they support the preference of the GACD(1,1) model.

Figure 5(b) shows the sample ACFs of the standardized residuals of the fitted GACD(1,1) model. From the plot, the standardized residuals do not have significant serial correlations, even though the lag-1 sample ACF is slightly above its two standard-error limit. We shall return to this point later when we introduce nonlinear ACD models. Figure 4(b) shows the time plot of the standardized residuals of the GACD(1,1) model. The residuals do not show any pattern of model inadequacy. The mean, standard deviation, minimum and maximum of the standardized residuals are 0.203, 4.497, 0.999, and 0.436, respectively.

It is interesting to see that the estimates of the shape parameter α are greater than 1 for both WACD(1,1) and GACD(1,1) models, indicating that the hazard function of the daily

Table 2: Estimation results of EACD(1,1), WACD(1,1) and GACD(1,1) models for the daily range of log price of Apple stock from January 4, 1999 to November 20, 2007. The sample size is 2235. The standard errors of the estimates are in parentheses. The p-values of the Ljung-Box statistics are also in parentheses with $Q(10)$ and $Q^*(10)$ for standardized residual series and its squared process, respectively.

Model	Parameters					Checking	
	α_0	α_1	β_1	α	κ	$Q(10)$	$Q^*(10)$
EACD	0.0007	0.133	0.849			16.65	12.12
	(0.0005)	(0.036)	(0.044)			(0.082)	(0.277)
WACD	0.0013	0.131	0.835	2.377		13.66	9.74
	(0.0003)	(0.015)	(0.021)	(0.031)		(0.189)	(0.464)
GACD	0.0010	0.133	0.843	1.622	2.104	14.62	11.21
	(0.0002)	(0.015)	(0.019)	(0.029)	(0.040)	(0.147)	(0.341)

range is monotonously increasing. This is consistent with the idea of volatility clustering, for large volatility tends to be followed by another large volatility. This phenomenon is different from that of the transaction durations in Example 1 for which $\hat{\alpha}$ is less than 1.

4 Diurnal Pattern

In this section, we discuss a simple method to adjust the diurnal pattern of intradaily trading activities. Figure 7(a) shows the trade durations of General Motors (GM) stock from December 1 to December 5, 2003. Again, for simplicity, zero durations are ignored. Figure 7(b) shows the time intervals from the market opening (9:30 am Eastern time) to the transaction time. The four vertical drops of the intervals signify the five trading days. From parts (a) and (b) of the figure, the diurnal pattern of trading activities is clearly seen. Specifically, except for a few outliers, the trade durations exhibit a cap-shape pattern within a trading day, namely the durations are in general shorter at the beginning and closing of the market, and longer around the middle of a trading day. One must consider such a diurnal pattern in modeling the transaction durations.

There are many ways to remove the diurnal pattern of transaction durations. Engle and Russell (1998) and Zhang, Russell and Tsay (2001) use some simple exponential functions of time and Tsay (2005) constructs some deterministic functions of time of the day to adjust the diurnal pattern. Let $f(t_i)$ be the mean value of the diurnal pattern at time t_i, measured from midnight. Then, define

$$x_i = \frac{z_i}{f(t_i)}, \tag{11}$$

be the adjusted duration, where z_i is the observed duration between the i-th and $(i-1)$th

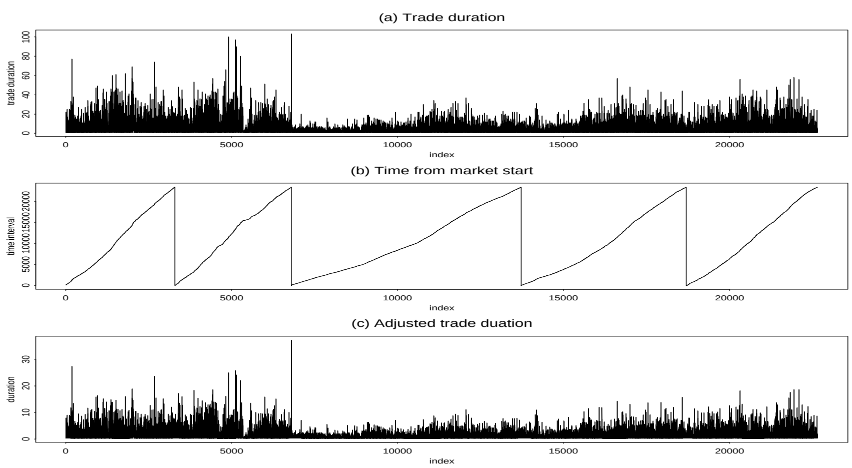

Figure 7: Time plots of durations for the General Motors stock from December 1 to December 5, 2003. (a) Observed trade durations (positive only), (b) Transaction times measured in seconds from midnight, (c) Adjusted trade durations.

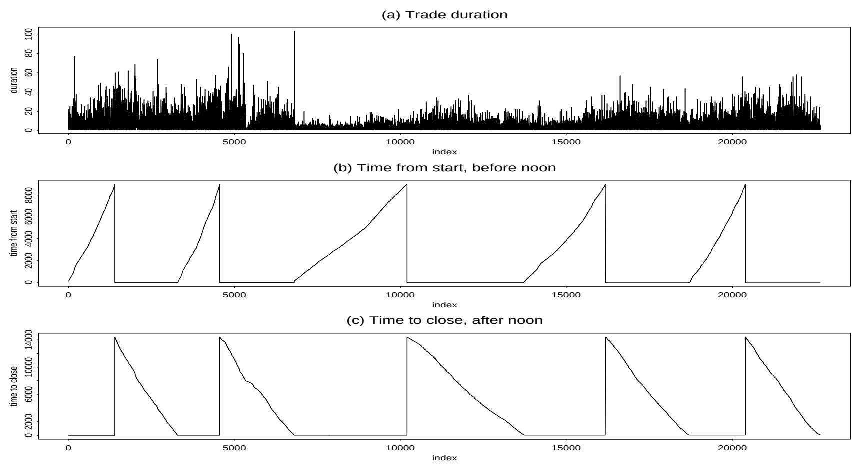

Figure 8: Time plots of durations for the General Motors stock from December 1 to December 5, 2003. (a) Observed trade durations (positive only), (b) and (c) the time function $O(t_i)$ and time function $C(t_i)$ of Eq. (12).

transactions. We construct $f(t_i)$ using two simple time functions. Define

$$O(t_i) = \begin{cases} t_i - 34200 & \text{if } t_i < 43200 \\ 0 & \text{otherwise,} \end{cases} \qquad C(t_i) = \begin{cases} 57600 - t_i & \text{if } t_i \geq 43200 \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where t_i is the time of the ith transaction measured in seconds from midnight and 34200, 43200, and 57600 denote, respectively, the market opening, noon, and market closing times measured in seconds. Figure 8(b) and (c) show the time plots of $O(t_i)$ and $C(t_i)$ of the GM stock transactions. Figure 8(a) shows the observed trade durations as in Figure 7(a). From the plots, the use of $O(t_i)$ and $C(t_i)$ is justified.
Consider the multiple linear regression

$$\ln(z_i) = \beta_0 + \beta_1 o(t_i) + \beta_2 c(t_i) + e_i, \tag{13}$$

where $o(t_i) = O(t_i)/10000$ and $c(t_i) = C(t_i)/10000$. Let $\hat{\beta}_i$ be the ordinary least squares estimates of the above linear regression. The residual is then given by

$$\hat{e}_i = \ln(z_i) - \hat{\beta}_0 - \hat{\beta}_1 o(t_i) - \hat{\beta}_2 c(t_i).$$

The adjusted durations then become

$$\hat{x}_i = \exp(\hat{e}_i). \tag{14}$$

For the GM stock transactions, the estimates of the β_i are 1.015(0.012), 0.133(0.028) and 0.313(0.016), respectively, where the numbers in parentheses denote standard errors. All estimates are statistically significant at the usual 1% level. Note that the residuals of the regression in Eq. (13) are serially correlated. Thus, the standard errors shown above underestimate the true ones. A more appropriate estimation method of the standard errors is to apply the Newey and West (1987) correction. The adjusted standard errors are 0.018, 0.044 and 0.027, respectively. These standard errors are larger, but all estimates remain statistically significant at the 1% level.
Figure 7(c) shows the time plot of the adjusted durations for the GM stock. Compared with part (a), the diurnal pattern of the trade durations is largely removed.

5 Nonlinear Duration Models

The linear duration models discussed in the previous sections are parsimonious in their parameterization and useful in many situations. However, in financial applications, the sample size can be large and the linearity assumption of the model might become an issue. Indeed, our limited experience indicates that some nonlinear characteristics are often observed in transaction durations and daily ranges of log stock prices. For instance, Zhang, Russell and Tsay (2001) showed that simple threshold autoregressive duration models can improve the analysis of stock transaction durations. In this section, we consider some simple nonlinear duration models and demonstrate that they can improve upon the linear ACD models.

5.1 Threshold autoregressive duration model

A simple nonlinear duration model is the threshold autoregressive conditional duration (TACD) model. The nonlinear threshold autoregressive (TAR) model was proposed in the time series literature by Tong (1978) and has been widely used ever since. See, for instance, Tong (1990) and Tsay (1989). A simple two-regime TACD(2;p, q) model for x_i can be written as

$$x_i = \begin{cases} \psi_i \epsilon_{1i} & \text{if } x_{t-d} \le r, \\ \psi_i \epsilon_{2i} & \text{if } x_{t-d} > r, \end{cases} \tag{15}$$

where d is a positive integer, x_{t-d} is the threshold variable, r is a threshold, and

$$\psi_i = \begin{cases} \alpha_{10} + \sum_{v=1}^{p} \alpha_{1v} x_{i-v} + \sum_{v=1}^{q} \beta_{1v} \psi_{i-v} & \text{if } x_{t-d} \le r, \\ \\ \alpha_{20} + \sum_{v=1}^{p} \alpha_{2v} x_{i-v} + \sum_{v=1}^{q} \beta_{2v} \psi_{i-v} & \text{if } x_{t-d} > r, \end{cases}$$

where $\alpha_{j0} > 0$ and α_{jv} and β_{jv} satisfy the conditions of the ACD model stated in Eq. (2) for $j = 1$ and 2. Here j denotes the regime. The innovations $\{\epsilon_{1i}\}$ and $\{\epsilon_{2i}\}$ are two independent *iid* sequences. They can follow the standard exponential, standardized Weibull, or standardized generalized Gamma distribution as before. For simplicity, we shall refer to the resulting models as the TEACD, TWACD, and TGACD model, respectively. The TACD model is a piecewise linear model in the space of x_{i-d}, and it is nonlinear when some of the parameters in the two regimes are different. The model can be extended to have more than 2 regimes. In what follows, we assume $p = q = 1$ in our discussion, because ACD(1,1) models fare well in many applications.

The TACD model appears to be simple, and it is indeed easy to use. However, its theoretical properties are very involved. For instance, the stationarity condition stated in Eq. (15) is only sufficient. The necessary condition of stationarity would depend on d and the parameters and deserves further investigation.

A key step in specifying a TACD model for a given time series is the identification of the threshold variable and the threshold, i.e., specifying d and r. The choice of d is relatively simple because $d \in \{1, \ldots, d_0\}$ for some positive integer d_0. For stock transaction durations, $d = 1$ is a reasonable choice as trading activities tend to be highly serially correlated. For the threshold r, a simple approach is to use empirical quantiles. Let $x_{<q>}$ be the q-th quantile of the observed durations $\{x_i | i = 1, \ldots, n\}$. We assume that $r \in \{x_{<q>} | q = 60, 65, 70, \ldots, 95\}$. For each candidate $x_{<q>}$, estimate the TACD(2;1,1) model

$$\psi_i = \begin{cases} \alpha_{10} + \alpha_{11} x_{i-1} + \beta_{11} \psi_{i-1} & \text{if } x_{t-1} \leq x_{<q>}, \\ \\ \alpha_{20} + \alpha_{21} x_{i-1} + \beta_{21} \psi_{i-1} & \text{otherwise,} \end{cases}$$

and evaluate the log likelihood function of the model at the maximum likelihood estimates. Denote the resulting log likelihood value by $\ell(x_{<q>})$. The threshold is then selected by

$$\hat{r} = x_{<q_o>} \quad \text{such that} \quad \ell(x_{<q_o>}) = \max_q \{\ell(x_{<q>}) | q = 60, 65, 70, \ldots, 95\}.$$

5.2 Example

In this subsection, we revisit the series of daily ranges of the log price of Apple stock from January 4, 1999 to November 20, 2007. The standardized innovations of the GACD(1,1) model of Section 3 have a marginally significant lag-1 autocorrelation. This serial correlation also occurs for the EACD(1,1) and WACD(1,1) models. Here we employ a two-regime threshold WACD(1,1) model to improve the fit. Preliminary analysis of the TWACD models indicates that the major difference in the parameter estimates between the two regimes is the shape parameter of the Weibull distribution. Thus, we focus on a TWACD(2;1,1) model with different shape parameters for the two regimes.

Table 3 gives the maximized log likelihood function of a TWACD(2;1,1) model for $d = 1$ and $r \in \{x_{<q>} | q = 60, 65, \ldots, 95\}$. From the table, the threshold 0.04753 is selected, which is the 70th percentile of the data. The fitted model is

$$x_i = \psi_i \epsilon_i, \quad \psi_i = 0.0013 + 0.1539 x_{i-1} + 0.8131 \psi_{i-1},$$

Table 3: Selection of the threshold of a TWACD(2;1,1) model for the daily range of the log price of Apple stock from January 4, 1999 to November 20, 2007. The threshold variable is x_{i-1}.

Quantile	60	65	70	75	80	85	90	95
$r \times 100$	4.03	4.37	4.75	5.15	5.58	6.16	7.07	8.47
$\ell(r) \times 10^3$	6.073	6.076	6.079	6.076	6.078	6.074	6.072	6.066

where the standard errors of the coefficients are 0.0003, 0.0164 and 0.0215, respectively, and ϵ_i follows the standardized Weibull distribution as

$$\epsilon_i \sim \begin{cases} W(2.2756) & \text{if } x_{i-1} \leq 0.04753, \\ W(2.7119) & \text{otherwise,} \end{cases}$$

where the standard errors of the two shape parameters are 0.0394 and 0.0717, respectively. Figure 9(a) shows the time plot of the conditional expected duration for the fitted TWACD(2;1,1) model, i.e. $\hat{\psi}_i$, whereas Figure 9(b) gives the residual ACFs for the fitted model. All residual ACFs are within the two-standard-error limits. Indeed, we have $Q(1) = 4.01(0.05)$, $Q(10) = 9.84(0.45)$ for the standardized residuals and $Q^*(1) = 0.83(0.36)$ and $Q^*(10) = 9.35(0.50)$ for the squared series of the standardized residuals, where the number in parentheses denotes p-value. Note that the threshold variable x_{i-1} is also selected based on the value of the log likelihood function. For instance, the log likelihood function of the TWACD(2;1,1) model assumes the value 6.069×10^3 and 6.070×10^3, respectively, for $d = 2$ and 3 when the threshold is 0.04753. These values are lower than that when $d = 1$.

6 Use of Explanatory Variables

High-frequency financial data are often influenced by external events, e.g., an increase or drop in interest rates by the U.S. Federal Open Market Committee or a jump in the oil price. Applications of ACD models in finance are often faced with the problem of outside interventions. To handle the effects of external events, the intervention analysis of Box and Tiao (1975) can be used. In this section, we consider intervention analysis in ACD modeling. We use the daily range series of Apple stock as an example. Here the intervention is the change in tick size of the U.S. stock markets.

On January 29, 2001, all stock prices on the U.S. markets switched to the decimal system. Before the switch, tick sizes of U.S. stocks went through several transitions, from 1/8 to 1/16 to 1/32 of a dollar. The observed daily range is certainly affected by the tick size.

Let t_o be the time of intervention. For the Apple stock, $t_o = 522$, which corresponds to January 26, 2001, the last trading day before the change in tick size. Since more

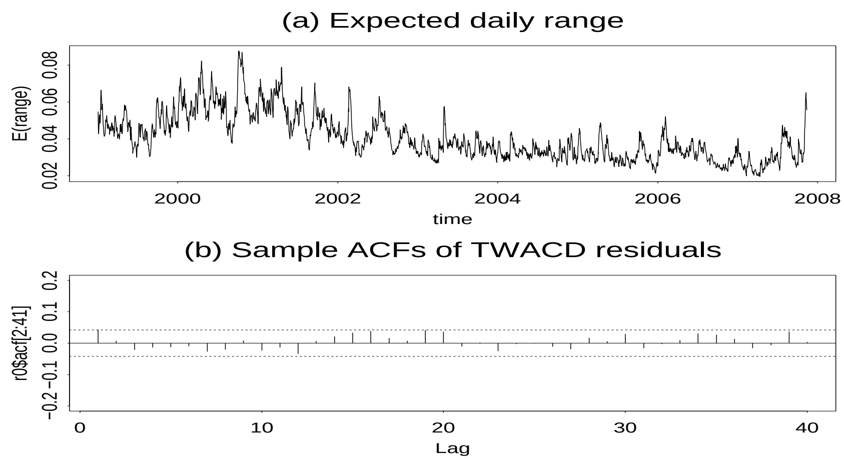

Figure 9: Model fitting for the daily range of the log price of Apple stock from January 4, 1999 to November 20, 2007: (a) The conditional expected durations of the fitted TWACD(2;1,1) model and (b) the sample ACF of the standardized residuals.

observations in the sample are after the intervention, we define the indicator variable

$$I_i^{(t_o)} = \begin{cases} 1 & \text{if } i \leq t_o, \\ 0 & \text{otherwise,} \end{cases}$$

to signify the absence of intervention. Since a larger tick size tends to increase the observed daily price range, it is reasonable to assume that the conditional expected range would be higher before the intervention. A simple intervention model for the daily range of Apple stock is then given by

$$x_i = \psi_i \begin{cases} \epsilon_{1i} & \text{if } x_{i-1} \leq 0.04753, \\ \\ \epsilon_{2i} & \text{otherwise,} \end{cases}$$

where ψ_i follows the model

$$\psi_i = \alpha_0 + \gamma I_i^{(t_o)} + \alpha_1 x_{i-1} + \beta_1 \psi_{i-1} \tag{16}$$

where γ denotes the decrease in expected duration due to the decimalization of stock prices. In other words, the expected durations before and after the intervention are

$$\frac{\alpha_0 + \gamma}{1 - \alpha_1 - \beta_1} \quad \text{and} \quad \frac{\alpha_0}{1 - \alpha_1 - \beta_1},$$

respectively. We expect $\gamma > 0$.
The fitted duration equation for the intervention model is

$$\psi_i = 0.0021 + 0.0011 I_i^{(522)} + 0.1595 x_{i-1} + 0.7828 \psi_{i-1},$$

where the standard errors of the estimates are 0.0004, 0.0003, 0.0177, and 0.0264, respectively. The estimate $\hat{\gamma}$ is significant at the 1% level. For the innovations, we have

$$\epsilon_i \sim \begin{cases} W(2.2835) & \text{if } x_{i-1} \leq 0.04753, \\ \\ W(2.7322) & \text{otherwise.} \end{cases}$$

The standard errors of the two estimates of the shape parameter are 0.0413 and 0.0780, respectively. Figure 10(a) shows the expected durations of the intervention model and Figure 10(b) shows the ACF of the standardized residuals. All residual ACFs are within the two-standard-error limits. Indeed, for the standardized residuals, we have $Q(1) = 2.37(0.12)$ and $Q(10) = 6.24(0.79)$. For the squared series of the standardized residuals, we have $Q^*(1) = 0.34(0.56)$ and $Q^*(10) = 6.79(0.75)$. As expected, $\hat{\gamma} > 0$ so that the decimalization indeed reduces the expected value of the daily range. This simple analysis shows that, as expected, adopting the decimal system reduces the volatility of Apple stock. Note that a general intervention model that allows for changes in the dynamic dependence of the expected duration can be used, even though our analysis only allows for a change in the expected duration. Of course, more flexible models are harder to estimate and understand.

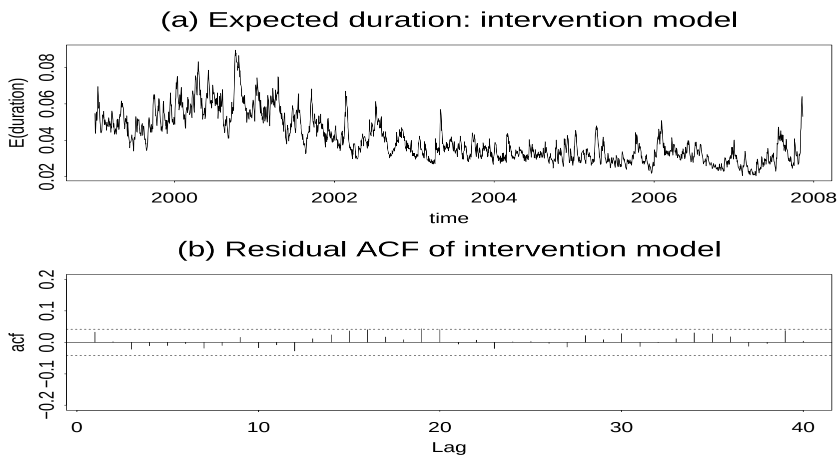

Figure 10: Model fitting for the daily range of the log price of Apple stock from January 4, 1999 to November 20, 2007: (a) The conditional expected durations of the fitted TWACD(2;1,1) model with intervention and (b) the sample ACF of the corresponding standardized residuals.

7 Conclusion

In this chapter, we introduced the autoregressive conditional duration models and discussed their properties and statistical inference. Among many applications, we used the model to study the daily volatility of stock price and found that, for the Apple sotck, adopting the decimal system on January 29, 2001 ideed significantly reduces the price volatility.

Note 1. The estimation of all ACD models in this chapter is carried out by the FMINCON function in Matlab.

REFERENCES

Box, G. E. P. and Tiao, G. C. (1975), "Intervention analysis with applications to economic and environmental problems," *Journal of the American Statistical Association*, 70, 70-79.

Chou, R. Y. (2005), "Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model," *Journal of Money, Credit and Banking*, 37, 561-582.

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

Newey, W. and West, K. (1987), "A simple positive semidefinite, heteroscedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 863-898.

Parkinson, M. (1980), "The extreme value method for estimating the variance of the rate of return," *Journal of Business*, 53, 61-65.

Tong, H. (1978). On a threshold model. In *Pattern Recognition and Signal Processing*, ed. C.H. Chen. The Netherlands: Sijthoff and Noordhoff.

Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford, U.K..

Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association*, 84, 231-240.

Tsay, R. S. (2005), *Analysis of Financial Time Series*, 2nd edition, John Wiley, Hoboken, New Jersey.

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