

Stock return volatility based on intraday high frequency data: double-threshold ACD-GARCH model

Sunah Chung^a · S.Y. Hwang^{a,1}

^aDepartment of Statistics, Sookmyung Women's University

(Received December 28, 2015; Revised January 2, 2016; Accepted January 2, 2016)

Abstract

This paper investigates volatilities of stock returns based on high frequency data from stock market. Incorporating the price duration as one of the factors in volatility, we employ the autoregressive conditional duration (ACD) model for the price duration in addition to the GARCH model to analyze stock volatilities. A combined ACD-GARCH model is analyzed in which a double-threshold is introduced to accommodate asymmetric features on stock volatilities.

Keywords: ACD, high frequency GARCH, double-threshold ACD-GARCH

1. 서론

우리가 이용하고 분석할 수 있는 자료는 컴퓨터 처리 능력이 급속히 향상됨에 따라 실시간으로 고빈도(high frequency)로 축적되고 있다. 주식시장의 경우, 주식거래는 1 초 간격 사이에도 여러 건이 발생되고 있고 고빈도 자료는 이러한 모든 거래에 수반되는 특성들(시간, 가격, 거래량 등)을 포함하여 기록하고 있다. 고빈도 자료는 불규칙하게 발생하는 거래정보를 담고 있는데 Engle과 Russell (1997, 1998)은 고빈도 자료 분석기법으로 ACD(autoregressive conditional duration) 모형을 소개하였다. ACD 모형은 거래들 사이의 시간간격인 듀레이션에 대해 GARCH 모형과 비슷한 수리 구조를 통해 변동성을 설명하는 모형이다. 듀레이션 정보만으로 구성된 기본적인 ACD 모형은 가격 정보를 포함하고 있지 않는데, Engle (2000)은 듀레이션과 가격 수익률을 연결하여 ACD-GARCH 모형을 연구한 바 있다. Engle (2000)의 ACD-GARCH 모형은 UHF-GARCH 모형으로 언급되기도 하지만 (여기서 UHF는 초고빈도 ultra high frequency의 약자이다) 본 논문에서는 ACD-GARCH 모형으로 부르기로 한다. 먼저, 기대듀레이션을 ACD 모형으로 분석하고 모수 추정을 위해 최우추정(ML)과 추정방정식(estimated function; EF) 방법을 적용하였다. ACD 추정식을 수익률과 변동성 분석으로 연결하는 이중-분계점(double-threshold) ACD-GARCH 모형을 사용하여 듀레이션의 영향과 정보의 비대칭성(asymmetry)에 대한 연구를 진행하였다. 2장에서는 본 논문에서 사용할 ACD 모형들을 소개하고, 3장에서는 ACD-GARCH 모형을 설명하고 4장에서는 자료 분석을 통해 예시하고 있다.

¹Corresponding author: Department of Statistics, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-Gu, Seoul 04310, Korea. E-mail: shwang@sookmyung.ac.kr

2. ACD(autoregressive conditional duration) 모형

거래자료는 두 개의 확률변수로 구성된다. 하나는 거래시간이고 다른 하나는 그에 수반되는 거래량, 거래가격과 같은 마크(mark)들이다. 기호 t_i 는 i 번째 거래가 일어나는 시간이고, 거래들 사이의 시간 간격 $x_i = t_i - t_{i-1}$ 을 듀레이션(duration)이라고 한다. 듀레이션은 관심 사건에 따라 구분되는데, 거래 듀레이션은 연속되는 두 거래 사이의 시간 차이로 정의되고, 가격듀레이션은 누적적으로 주어진 크기의 가격 변화가 있을 때까지의 시간으로 정의되며 변동성의 척도로서 사용될 수 있다. 이와 유사하게, 볼륨 듀레이션은 누적적으로 정해놓은 크기의 주문량이 될 때 까지의 시간으로 정의된다 (Bauwens와 Hautsch, 2009).

Engle과 Russell (1997, 1998)이 소개한 ACD 모형에서는 조건부 기대 듀레이션 ψ_i 을 다음과 같이 정의한다.

$$\psi_i = E(x_i | x_{i-1}, x_{i-2}, \dots, x_1) = E(x_i | F_{i-1}),$$

여기서 F_{i-1} 은 $i-1$ 시점까지의 정보집합이다. 실제 듀레이션 x_i 는 조건부 기대 듀레이션 ψ_i 와 예측하지 못한 듀레이션 ϵ_i 로 구성된다고 가정하고 곱의 형태로 구성 하였다.

$$x_i = \psi_i \epsilon_i, \quad \epsilon_i \stackrel{i.i.d.}{\sim} (1, \sigma_\epsilon^2).$$

예측하지 못한 듀레이션 ϵ_i 에 대해서는 모수 추정을 위해 분포를 가정한다. 듀레이션은 양의 값을 가지므로, ϵ_i 도 양의 값을 가지며 평균은 1이고 분산이 σ_ϵ^2 인 확률분포를 가정하는데 보통 지수분포와 와이블분포를 이용한다. 그 외에 감마분포, Burr분포 등도 고려한다.

조건부 기대 듀레이션 ψ_i 모형은 변동성에 대한 정보를 담고 있다고 생각하고 GARCH(p, q)와 유사하게 설정한 모형이며, ACD(p, q) 모형은 다음과 같다 (Bauwens와 Giot, 2003).

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

여기서 p, q 는 시차를 나타내며, 듀레이션은 양의 값을 가져야 하므로 $\omega > 0, \alpha_j \geq 0, \beta_j \geq 0$ 을 가정한다. 예를 들어, 간단한 모형 ACD(1, 1)은 다음과 같다.

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

오차항으로 지수 분포를 가정한 EACD(1, 1)의 조건부 확률밀도함수는 아래와 같다.

$$f(\epsilon_i) = \exp(-\epsilon_i),$$

$$f(x_i | \psi_i) = \frac{1}{\psi_i} \exp\left(-\frac{x_i}{\psi_i}\right).$$

오차항으로 와이블 분포를 가정한 WACD(1, 1)의 조건부 확률밀도함수는 아래와 같다.

$$f(\epsilon_i | F_{i-1}; \psi_i) = \kappa \left[\Gamma\left(1 + \frac{1}{\kappa}\right) \right]^\kappa \epsilon_i^{\kappa-1} \exp\left(-\left[\Gamma\left(1 + \frac{1}{\kappa}\right) \epsilon_i \right]^\kappa\right),$$

$$f(x_i | F_{i-1}; \psi_i) = \frac{\kappa}{x_i} \left[\Gamma\left(1 + \frac{1}{\kappa}\right) \frac{x_i}{\psi_i} \right]^\kappa \exp\left(-\left[\Gamma\left(1 + \frac{1}{\kappa}\right) \frac{x_i}{\psi_i} \right]^\kappa\right),$$

여기서 κ 는 scale-모수이다. 본 논문에서 ACD 모형의 모수들은 ML(maximum likelihood)과 EF(estimated function) 두 가지 방법을 이용하여 추정하였다. 먼저, MLE는 EACD(1, 1)과 WACD(1, 1)를 가정하고 다음의 로그우도함수를 사용하여 추정하였다 (Engle과 Russell, 1997, 1998).

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log f(\epsilon_i | F_{i-1}; \psi_{i-1}), \quad \boldsymbol{\theta} = (\omega, \alpha, \beta),$$

$$\hat{\boldsymbol{\theta}} = \arg \max l(\boldsymbol{\theta}).$$

한편, 오차항에 대한 분포는 모르는 경우가 현실적이므로 분포에 대한 가정을 하지 않는 추정함수(EF) 모수추정 방법인 Godambe (1985)의 추정함수 $g(\mathbf{x}; \boldsymbol{\theta})$ 를 이용한다. 적절한 조건하에서 $E[g(\mathbf{x}; \boldsymbol{\theta})] = 0$ 이 되는 최적의 선형불편추정함수는 다음과 같이 주어진다.

$$g^*(\mathbf{x}; \boldsymbol{\theta}) = h_i a_{i-1}^*,$$

여기서 $a_{i-1}^* = E_{i-1}[\partial h_i / \partial \boldsymbol{\theta}] / E_{i-1}[h_i^2]$ 이며 h_i 는 \mathbf{x} 와 $\boldsymbol{\theta}$ 의 함수인 마팅계일 차(martingale difference)이다. 즉, ACD 모형의 경우에는 $h_i = x_i - \psi_i$ 를 사용하면 a_{i-1}^* 와 최적의 선형불편추정함수는 다음과 같이 된다 (Allen 등, 2013).

$$a_{i-1}^* = E_{i-1} \left[\frac{\partial h_i}{\partial \boldsymbol{\theta}} \right] / \psi_i^2 \sigma_\epsilon^2,$$

$$g^*(\mathbf{x}; \boldsymbol{\theta}) = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} E_{i-1} \left[\frac{\partial h_i}{\partial \boldsymbol{\theta}} \right] (x_i - \psi_i).$$

추정함수 추정치 $\hat{\boldsymbol{\theta}}$ 은 $g^*(\mathbf{x}; \boldsymbol{\theta})$ 을 풀어서 얻을 수 있다. $\hat{\boldsymbol{\theta}}$ 에 대한 분산-공분산 행렬 추정은 Allen 등 (2012)을 따라 다음과 같이 정의된 \mathbf{V} 행렬의 역행렬로 계산하였다.

$$\mathbf{V} = \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{bmatrix},$$

여기서 각 원소들은 다음과 같다.

$$V_{11} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \omega} \right)^2, \quad V_{12} = V_{21} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \omega} \right) \left(\frac{\partial \psi_i}{\partial \alpha} \right),$$

$$V_{22} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \alpha} \right)^2, \quad V_{23} = V_{32} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \alpha} \right) \left(\frac{\partial \psi_i}{\partial \beta} \right),$$

$$V_{33} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \beta} \right)^2, \quad V_{13} = V_{31} = \sum_{i=1}^n \frac{1}{\psi_i^2 \sigma_\epsilon^2} \left(\frac{\partial \psi_i}{\partial \omega} \right) \left(\frac{\partial \psi_i}{\partial \beta} \right).$$

3. ACD-GARCH 모형

ACD 모형의 확장은 GARCH 모형의 일반화 형태와 유사하다. 참고로, Bauwens와 Hautsch (2009)에 다양한 ACD 모형들이 정리되어 있다. 확장된 여러 모형 중에서 본 논문에서는 Engle (2000)의 ACD-GARCH 모형을 이용하여 주식수익률의 변동성을 분석하고자 한다. 기본 ACD 모형들은 거래 사이에 듀레이션만을 고려하고 가격 과정이 주는 정보는 포함하고 있지 않다. 각 거래에서 가격은 중요한 정보

이며 이를 모형에 수용하는 것이 유용할 것이다. 거래 사이의 듀레이션과 가격 과정의 정보를 연결하여 확장한 모형이 ACD-GARCH 모형이다. ACD모형이 듀레이션을 설명하고 GARCH 모형에서 듀레이션이 수익률의 변동성을 설명하는데 사용되는 것이다 (Bauwens와 Giot, 2003). 가격 듀레이션 모형은 가격 변화당 기대 시간을 측정하는 것인데 직관적으로 이것은 변동성 측정과 연결 지어 생각할 수 있고 변동성과 듀레이션의 역의 관계를 경험적으로 확인해 볼 수 있게 된다 (Engle과 Russell, 1997). 즉, 듀레이션이 짧을수록 변동성이 커지는 것을 예상할 수 있게 된다.

Engle (2000)의 ACD-GARCH 모형은 UHF(ultra-high-frequency)-GARCH 모형으로도 불리는데, 가격변동성을 측정하고 거래시간이 이 변동성에 어떻게 영향을 미치는지를 파악하고자 만든 모형이다. Jo (2010)은 이 모형을 국내주가분석에 응용하였다. r_i 가 $i-1$ 번째 거래와 i 번째 거래사이의 수익률이라고 하면, 거래당 조건부 분산은 다음과 같다.

$$V_{i-1}(r_i|x_i) = h_i,$$

여기서 분산은 과거의 수익률과 듀레이션 뿐만 아니라 현재의 듀레이션을 조건으로 하고 있다. 분산은 고정된 시간간격에 대해서 측정되므로 여기서 시간당 조건부 분산은 다음과 같이 추정한다.

$$V_{i-1}\left(\frac{r_i}{\sqrt{x_i}} \middle| x_i\right) = \sigma_i^2.$$

두개의 조건부 분산은 $h_i = x_i\sigma_i^2$ 으로 연결된다. 관측되는 이산적 로그가격 p_i 와 관측할 수 없는 연속적 로그가격 m_i 와의 관계는 다음과 같다.

$$p_i = m_i + \zeta_i,$$

여기서 오차 ζ_i 는 호가단위에 의한 절단의 결과이다. m_i 는 $\sqrt{x_i}$ 와 비례하는 이노베이션을 가지는 마팅계일로 가정하고 $\sqrt{x_i}$ 시간당 측정오차는 ARMA(1,1)으로 가정한다. 즉,

$$m_i = m_{i-1} + \sqrt{x_i}\nu_i, \quad \frac{p_i - m_i}{\sqrt{x_i}} \equiv \eta_i = \rho\eta_{i-1} + \xi_i + \chi\xi_{i-1}.$$

이 식은 $\sqrt{x_i}$ 시간당 수익률이 ARMA(1,1)인 것을 의미하게 된다. 즉,

$$\begin{aligned} \frac{r_i}{\sqrt{x_i}} &\equiv \frac{\Delta p_i}{\sqrt{x_i}} = \nu_i + \eta_i = \rho \frac{r_{i-1}}{\sqrt{x_{i-1}}} + \nu_i - \rho\nu_{i-1} + \xi_i + \chi\xi_{i-1}, \\ \frac{r_i}{\sqrt{x_i}} &= \rho \frac{r_{i-1}}{\sqrt{x_{i-1}}} + e_i + \phi e_{i-1}. \end{aligned}$$

시간당 수익률에 대한 간단한 변동성 모형은 다음과 같은 ACD-GARCH(1,1) 모형이다.

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2,$$

여기에는 듀레이션에 대한 정보를 반영하고 있지 않기 때문에 이를 고려하여 확장된 모형은 다음과 같다.

$$\sigma_i^2 = \omega + \alpha e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \xi_{i-1} + \gamma_4 \psi_i^{-1},$$

여기서 x_i 은 듀레이션, ψ_i 은 기대듀레이션, ξ_i 은 장기변동성 설명변수이다. ξ_i 계산은 r^2/x 을 지수평활하여 $\xi_i = \lambda\xi_{i-1} + (1-\lambda)(r_{i-1}^2/x_{i-1})$, $\lambda = 0.995$ 로 계산한다. x_i/ψ_i 은 기대하지 못한 듀레이션으로 단기 변동성의 영향을 알아보기 위한 설명변수라 할 수 있다. 듀레이션과 기대듀레이션이 짧을수록, 단기 변동성과 장기변동성이 클수록 수익률의 변동성이 커질 것을 예상하는 모형이라 할 수 있다.

4. 예시 및 자료 분석

ACD 모형 자료는 Engle과 Russell (1998)에서와 같이 일중패턴(diurnal pattern) 제거와 가격 간략화 과정(thinning price process)을 거친 후 사용한다. 일중패턴은 거래와 관련된 특징들이 규칙적으로 반복되는 경향을 나타낸다. 예를 들면, 주식시장 시작 시간과 끝나는 시간에 거래가 활발하게 일어나고 중간 시간에는 상대적으로 거래가 덜 일어나는 경향을 말한다. 가격 간략화 과정은 가격변화의 임계값을 설정하여 임계값 이하의 가격변화는 의미가 없다고 생각하고 임계값 이상의 가격변화를 갖는 자료들만을 사용하여 듀레이션을 측정하는 것을 말한다.

ACD 모형 분석에서는 다양한 방법으로 위의 제거과정을 거치고 있다. Engle (2000)은 듀레이션이 0인 자료를 제외한 모든 자료를 사용하였다. 일중패턴은 piecewise linear spline을 사용하여 시간에 따른 듀레이션의 전형적인 유형을 제거하였다. Jo (2010)에서는 동일시각 초 단위에 발생한 거래가 서로 구분되어 있을 경우 이들을 하나의 동일한 거래로 취급하고 거래량은 구분되어 있는 각 거래량을 합산하여 사용하였다. 9시 이전과 15시 이후의 거래는 제외하고 일중 효과를 삼차스플라인 회귀함수를 사용하여 제거하였다. Bauwens와 Giot (2003)은 의미 있는 가격변화 사이의 간격을 듀레이션으로 사용하였는데 기준을 bid-ask의 중간가격의 누적변화가 0.125가 일어나는 시간 간격을 자료로 사용하여 분석하였다. 일중효과 제거에는 커널스무딩을 사용하였다. Park과 Kim (2014)은 동일 시간에 거래된 여러 자료는 하나로 간략화 하였는데, 거래량은 합하고 가격은 마지막 자료의 것을 사용한 후, 임계값을 설정하여 두 틱을 연속해서 임계값을 넘어서는 경우의 자료만을 사용하였다. 자료조정을 거치게 되면, 자료분석에 사용되는 듀레이션 x_i 은 결정적 요소가 제거된 듀레이션이 된다. 즉, 듀레이션은 확률적 요소과 결정적 요소로 구성되는데, 확률적 요소는 ACD 모형으로 설명되고, 결정적 요소는 일중효과 같은 것으로 ACD 모형을 분석하기 전에 원자료로부터 제거되어야 할 요소로 본다. 즉, $X_i = x_i\phi(t_i)$ 에서 X_i 는 원자료, $\phi(t_i)$ 는 일중패턴함수, x_i 는 조정된 듀레이션이다 (Bauwens와 Giot, 2003).

본 논문에서 실증분석에 사용한 자료는 우리나라 대표종목인 삼성전자이다. 자료는 미래에셋증권 HTS에서 다운받은 자료를 사용하였는데, 자료에는 거래가 이루어진 시간이 초단위로 기록되어 있고 해당 거래의 가격과 거래량이 기록되어 있다. 기간은 2015년 8월 7일 부터 2015년 11월 10일까지 약 3개월간 자료이고 자료의 개수는 1,180,880개이다. 초 단위 이하에 거래된 자료는 해당 초단위로 합하여 거래량을 만들고 가격 자료는 해당 초에서 마지막 가격자료를 사용하였다. 거래시간 오전 9시부터 오후 3시까지(32400초~54000초) 이외의 자료와 매 거래일의 첫 번째 거래는 제외하였다. 가격에 대한 임계값은 따로 설정하지 않았고 0보다 큰 가격변화가 있을 때까지의 시간간격을 가격 듀레이션으로 계산하였다. 이상의 제거 과정 후의 자료의 개수는 250,569개이다. 일중 효과는 Jo (2010)에서와 같이 다음과 같은 삼차스플라인 회귀함수를 사용하여 추정하였다.

$$\phi(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \sum_{k=1}^5 \beta_{k+3} (t - t_k)^3 I(t \geq t_k),$$

여기서 $I(\cdot)$ 는 0과 1값을 갖는 지시함수이다. 스피라인 마디점(knots)은 1시간 간격인 $t_k = 32400 + 3600 \times k$, $k = 1, 2, 3, 4, 5$ 으로 하였고 $I(t \geq t_k)$ 은 마디점에 대한 지시함수이다. 추정된 일중효과로 조정된 듀레이션은 $x_i = X_i/\phi(t_i)$ 이 된다.

Figure 4.1에서 가로축은 일중 주식거래시간을 초단위로 나타낸 것이고 세로축은 듀레이션에 대한 삼차스플라인 모형 추정 결과이다. 삼성전자의 경우 주식 시장 시작 시간과 마감 시간 근처에서 주식거래가 활발하여 듀레이션이 짧고 10시경 근처에서 가장 긴 듀레이션이 길어지는 것을 볼 수 있다. 반대로 초당 주식수익률은 주식 시장 시작 시간과 마감 시간 근처에서 높고 12경 점심시간 근처에서 가장 낮은 것을 볼 수 있다. 주식가격 역시 듀레이션과 마찬가지로 전형적인 일중 효과를 가진다. 주식시장은 중간

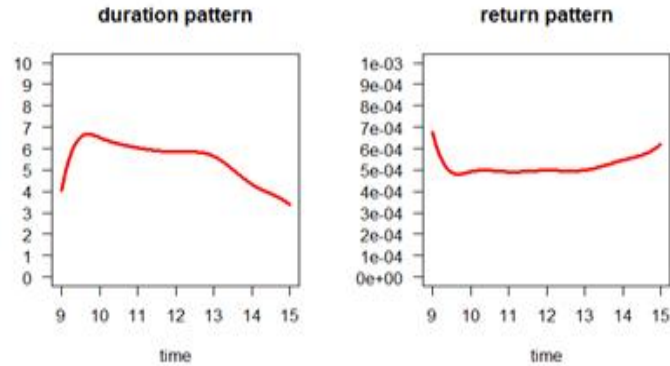


Figure 4.1. Diurnal patterns using cubic spline (Samsung Electronics).

Table 4.1. Summary statistics

	관측치수	1일 평균 관측치수	평균	표준편차	최솟값	중위수	최댓값	왜도	첨도
제거전	250569	3915	5.358	6.458	1	3	123	3.609	22.179
제거후	250569	3915	0.992	1.151	0.150	0.602	21.228	3.386	19.434

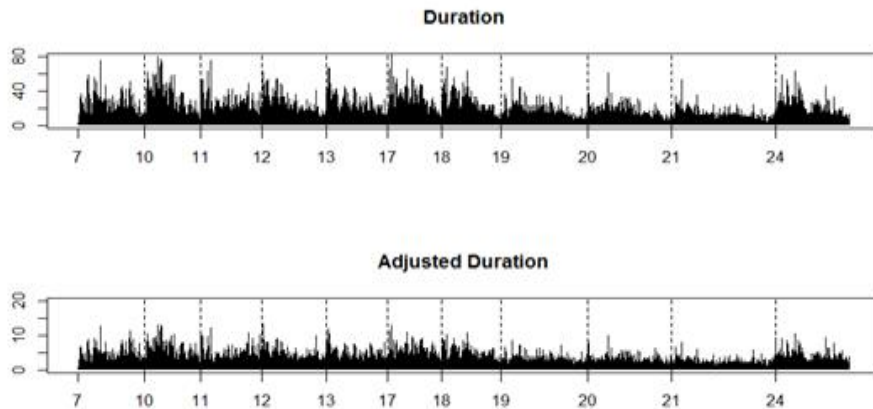


Figure 4.2. A plot of duration and adjusted duration.

시간 보다 시작과 끝 시간에 좀 더 활발하고 이것이 듀레이션과 변동성 모두에 나타난다. 수익률 역시 변동성이 일증패턴을 가지고 있기 때문에 조정되어야 하는데 여기서는 시간당 수익률의 단위가 너무 작아서 전반적으로는 조정의 효과가 미미하여 조정을 하지 않고 사용하였다.

Table 4.1과 같이 일증패턴 제거 전에는 듀레이션의 평균이 5.358초이고 표준편차는 6.458초, 범위는 1~123초였는데, 일증패턴 제거 후에는 평균이 0.992초이고 표준편차는 1.151초, 범위는 0.150~21.228초로 나타났다. 왜도는 3.609에서 3.386으로 줄어들었으나 비대칭 형태이고, 첨도도 22.179에서 19.434로 줄어들었으나 중심부분이 뾰족한 형태이다. Figure 4.2는 일증패턴 제거 전/후의 듀레이션의 모습을 일부기간(8월7일~8월24일)만 나타낸 것이다.

Table 4.2. Estimates for ACD(1, 1) model (standard errors in parenthesis)

	ML-EACD	ML-WACD	EF
$\hat{\omega}$	0.0012 (0.0001)	0.0011 (0.0001)	0.0012 (0.0001)
$\hat{\alpha}$	0.0126 (0.0002)	0.0131 (0.0002)	0.0127 (0.0004)
$\hat{\beta}$	0.9862 (0.0001)	0.9859 (0.0001)	0.9861 (0.0004)
$\hat{\kappa}$		1.0901 (0.0016)	
Log Likelihood	-240856	-239176	

Table 4.3. Estimates for ACD-GARCH model

Parameter	Model (4.1)		Model (4.2)		Model (4.3)	
	Estimate	Approx Std Err	Estimate	Approx Std Err	Estimate	Approx Std Err
$\hat{\mu}$	0.0163	0.0020	0.0172	0.0021	0.0136	0.0016
$\hat{\eta}$	-0.9627	0.0007	-0.9623	0.0007	-0.8573	0.0012
$\hat{\phi}$	0.5756	0.0026	0.5737	0.0026	0.4995	0.0026
$\hat{\omega}$	0.0801	0.0024	0.0790	0.0026	-0.1564	0.0016
$\hat{\alpha}_1$	0.1886	0.0037	0.1907	0.0044	0.2503	0.0046
$\hat{\alpha}_2$			-0.0080	0.0032	-0.1315	0.0047
$\hat{\beta}$	0.6547	0.0079	0.6586	0.0083	0.0950	0.0026
$\hat{\gamma}_1$					0.1912	0.0012
$\hat{\gamma}_2$					0.0625	0.0009
$\hat{\gamma}_3$					0.1285	0.0023
$\hat{\gamma}_4$					-0.1502	0.0018
Log Likelihood	-255280		-255277		-227745	

삼성전자의 조건부 기대 듀레이션 ACD(1, 1) 모형의 추정결과는 Table 4.2와 같다. 여기서 x_i 는 조정된 듀레이션이다. 기대 듀레이션 ψ_i 에 대한 모수추정치는 세 가지 경우 모두 비슷한 결과를 주었는데, 이중 오차항에 대한 분포를 지수분포로 가정한 경우와 오차항에 대한 분포를 가정하지 않은 EF 방법의 결과가 더 비슷하게 나왔다. 표준오차는 지수분포와 와이블분포가 비슷하고 EF 방법이 지수분포 보다는 미미한 수준으로 수치 차이를 보였다. $\hat{\alpha} + \hat{\beta}$ 는 약 0.998로 과거의 변동성이 오래 동안 지속되는 것으로 나타났다. 또한, EF 방법으로 오차항에 대한 분포를 가정하지 않고 추정한 결과는 지수분포와 거의 차이가 없는 결과를 보여주었다.

다음은 단위거래시간당 일중 수익률의 변동성을 ACD-GARCH 모형으로 분석한 결과이다. 모형 (4.1)은 Engle의 기본 모형을 적용한 결과이고, 모형 (4.2)와 (4.3)은 기본 모형의 변동성에 분계점(threshold)을 준 모형이다. 모형 (4.3)에 사용한 기대 듀레이션 ψ_i 는 EACD(1, 1)으로 추정한 결과이다. ACD-GARCH 모형의 추정엔 SAS/ETS의 MODEL 프로시저를 사용하였다.

$$\frac{r_i}{\sqrt{x_i}} = \mu + \eta \frac{r_{i-1}}{\sqrt{x_{i-1}}} + e_i + \phi e_{i-1},$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \beta \sigma_{i-1}^2, \tag{4.1}$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \alpha_2 I(e_{i-1} < 0) e_{i-1}^2 + \beta \sigma_{i-1}^2, \tag{4.2}$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \alpha_2 I(e_{i-1} < 0) e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \xi_{i-1} + \gamma_4 \psi_i^{-1}. \tag{4.3}$$

Table 4.3 결과를 보면, 공통적으로 $\hat{\eta}$ 가 음의 수치로 크게 나와 현재의 시간당 수익률은 이전 시점의 수익률과 반대방향으로 움직이는 것으로 나타난다. 기본 모형 (4.1)에서 변동성은 과거의 변동성이 지속됨을 보여주는데 식 (4.2)와 같이 잔차가 양일 경우와 음일 경우로 나누어 보면 변동성에 미치는 영향이

Table 4.4. Estimates for double-threshold ACD-GARCH model

Parameter	Model (4.4)		Model (4.5)		Model (4.6)	
	Estimate	Approx Std Err	Estimate	Approx Std Err	Estimate	Approx Std Err
$\hat{\mu}$	-0.0028	0.0009	-0.0031	0.0010	-0.0035	0.0009
$\hat{\eta}_1$	-0.9464	0.0020	-0.9459	0.0021	-0.8561	0.0022
$\hat{\eta}_2$	0.9787	0.0021	0.9793	0.0022	0.8824	0.0022
$\hat{\phi}$	0.5744	0.0026	0.5745	0.0028	0.5049	0.0026
$\hat{\omega}$	0.0789	0.0025	0.0791	0.0038	-0.1911	0.0020
$\hat{\alpha}_1$	0.1869	0.0038	0.1863	0.0071	0.1589	0.0029
$\hat{\alpha}_2$			0.0009	0.0050	0.0118	0.0038
$\hat{\beta}$	0.6587	0.0081	0.6585	0.0121	0.1118	0.0029
$\hat{\gamma}_1$					0.2051	0.0013
$\hat{\gamma}_2$					0.1053	0.0014
$\hat{\gamma}_3$					0.1098	0.0023
$\hat{\gamma}_4$					-0.1707	0.0022
Log Likelihood	-255274		-255272		-226067	

0.1907과 0.1827으로 양일 경우 영향력이 다소 크다. 듀레이션의 정보를 고려한 모형 (4.3)으로 확장하면, 잔차의 양과 음의 효과는 0.2503과 0.1188으로 차이가 더 커지는 방향으로 나타나고, 현재 가격변화 듀레이션 x_i 가 짧아질수록 변동성이 커지지만, 반대로 ψ_i 기대 듀레이션은 짧아질수록 변동성은 작아지는 것으로 나타났다. 실제 듀레이션과 기대 듀레이션의 비율 x_i/ψ_i 가 클수록 변동성이 커지는 것으로 나타났고, ξ_{i-1} 이전 시점의 장기변동성이 클수록 단위시간당 수익률에 양의 방향으로 영향을 미치고 있다.

이제, 이중-분계점(double-threshold) ACD-GARCH 모형을 도입하기로 하자. 이중-분계점 모형은 주가 수익률의 비대칭 패턴이 기대 조건부 수익률(평균함수)과 변동성(분산함수) 모두에 동시에 나타날 때 유용한 모형이다. 모형 (4.4)는 Engle의 기본 모형 평균방정식에도 분계점을 도입한 모형이다. 여기서 모형 (4.5)와 (4.6)은 기대 조건부 수익률에만 분계점이 있는 모형 (4.4)에 변동성에도 분계점을 도입해 이중-분계점(double-threshold)을 사용한 모형이다.

$$\frac{r_i}{\sqrt{x_i}} = \mu + \eta_1 \left(\frac{r_{i-1}}{\sqrt{x_{i-1}}} \right)^+ + \eta_2 \left(\frac{r_{i-1}}{\sqrt{x_{i-1}}} \right)^- + e_i + \phi e_{i-1},$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \beta \sigma_{i-1}^2, \quad (4.4)$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \alpha_2 I(e_{i-1} < 0) e_{i-1}^2 + \beta \sigma_{i-1}^2, \quad (4.5)$$

$$\sigma_i^2 = \omega + \alpha_1 e_{i-1}^2 + \alpha_2 I(e_{i-1} < 0) e_{i-1}^2 + \beta \sigma_{i-1}^2 + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \xi_{i-1} + \gamma_4 \psi_i^{-1}. \quad (4.6)$$

Table 4.4에서, 이전 시점의 시간당 수익률이 양일 경우와 음일 경우로 나누어 보면 $\hat{\eta}_1$ 보다 $\hat{\eta}_2$ 의 절대치가 다소 크므로 음의 수익률인 경우에 양의 방향으로 더 큰 영향을 보이는 것으로 나타났다. 모형 (4.5)과 같이 잔차가 양일 경우와 음일 경우로 나누어 보면 음일 경우 효과가 유의하지 않아 양일 경우와 차이를 보이지 않는다. 듀레이션을 고려한 모형 (4.6)으로 확장하면, 잔차의 양과 음의 효과는 0.1589와 0.1707으로 모두 유의미 하고 음일 경우에 변동성에 미치는 영향이 양일 때 보다 다소 크게 나타난다. 이는 모형 (4.3)과는 다른 모습인데 이것은 평균함수에 도입한 분계점의 영향으로 변동성의 분계점 효과가 기대수익률의 분계점 효과로 옮겨간 결과로 보인다. 그러나, 듀레이션과 관련한 효과들은 모형 (4.3)과 방향성이 같게 나왔는데, 현재 가격변화 듀레이션 x_i 는 짧아질수록 변동성이 커지지만, 반

대로 기대 듀레이션 ψ_i 는 짧아질수록 변동성은 작아지고, 실제 듀레이션과 기대 듀레이션의 차이(비율) x_i/ψ_i 이 클수록 변동성이 커지고, ξ_{i-1} 이전 시점의 장기변동성이 클수록 단위시간당 수익률에 양의 방향으로 영향을 미치고 있다.

주식시장에서는 거래에서 발생하는 방대한 규모의 고빈도 자료를 저장하고 있다. 본 논문에서는 고빈도 자료가 갖고 있는 정보를 활용하여 주식수익률에 영향을 미치는 요인들을 살펴보았다. 우선, 거래간의 시간간격적인 듀레이션을 ACD 모형으로 추정하였는데 오차항의 분포를 가정한 경우와 가정하지 않은 경우에 비슷한 결과를 얻었다. 다음으로 기대수익률과 변동성을 ACD-GARCH 모형으로 분석하였는데 비대칭적 정보를 수량화하기 위해 이전 시점의 양의 효과와 음의 효과를 분계점을 사용하여 구분해 보았다. 삼성전자 주식의 경우, 변동성 부분에만 분계점을 주었을 때 나타났던 양과 음의 효과를 기대수익률과 변동성 모두에 분계점을 사용한 경우와 비교하면, 변동성 부분에 나타났던 분계점 효과가 기대수익률 부분의 분계점 효과로 옮겨가면서 변동성 부분의 비대칭 효과는 다소 감소되는 것으로 나타났다.

References

- Allen, D., Ng, K. H. and Peiris, S. (2012). Estimating and simulating Weibull models of risk or price durations: an application to ACD models, *North American Journal of Economics and Finance*, **25**, 214–224.
- Allen, D., Ng, K. H. and Peiris, S. (2013). The efficient modelling of high frequency transaction data: a new application of estimating functions in financial economics, *Economics Letters*, **120**, 117–122.
- Bauwens, L. and Giot, P. (2003). Asymmetric ACD models: Introducing price information in ACD models, *Empirical Economics*, **28**, 709–731.
- Bauwens, L. and Hautsch, N. (2009). *Modelling Financial High Frequency Data Using Point Processes*, Handbook of Financial Time Series, Springer.
- Engle, R. F. (2000). The econometrics of ultra-high-frequency data, *Econometrica*, **68**, 1–22.
- Engle, R. F. and Russell, J. R. (1997). Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model, *Journal of Empirical Finance*, **4**, 187–212.
- Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data, *Econometrica*, **66**, 1127–1162.
- Godambe, V. P. (1985). The foundation of finite sample estimation in stochastic processes, *Biometrika*, **72**, 419–428.
- Jo, S. P. (2010). A study on determinants of volatility in intra-day stock return: an application of UHF-GARCH-Leverage model, M.A thesis, Hanyang University.
- Park, S. N. and Kim, Y. J. (2014). Bayesian forecasting with nonlinear autoregressive conditional duration models, *Journal of Industrial Economics and Business*, **27**, 1–33.
- Tsay, R. S. (2010). *Analysis of Financial Time Series*, Third Ed. Wiley, New York.

이중-분계점 ACD-GARCH 모형을 이용한 일중 고빈도 자료의 주식 수익률 변동성 분석

정선아^a · 황선영^{a,1}

^a숙명여자대학교 통계학과

(2015년 12월 28일 접수, 2016년 1월 2일 수정, 2016년 1월 2일 채택)

요약

주식시장 거래에서 기록되는 고빈도 자료를 사용하여 주식 수익률에 대한 변동성을 분석하였다. 변동성을 설명할 수 있는 한 요소로 주식거래에서 불규칙한 간격으로 발생하는 가격 듀레이션을 생각할 수 있는데, 실제 자료에 ACD 모형을 사용하여 듀레이션을 추정해 보았고, ACD-GARCH 모형을 사용하여 주식 수익률과 변동성에 미치는 듀레이션의 영향을 살펴보았다. 이 과정에서 ACD 모형 추정에는 ML과 EF 방법을 적용하였고, ACD-GARCH 모형에는 이중-분계점(double-threshold)을 추가하여 평균수익률의 비대칭성 및 변동성의 비대칭성을 동시에 분석해 보았다.

주요용어: ACD, 고빈도-GARCH, 이중-분계점 ACD-GARCH

¹교신저자: (04310) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 통계학과.
E-mail: shwang@sookmyung.ac.kr