Modified QMLEs for Location and Zero-Augmented Multiplicative Error Models: Order of (1,1)

Qian Li *

Abstract
We extend the linear Multiplicative Error Model (MEM) at order of (1,1) by adding the location parameter. The minimum of the sample is shown to be a consistent estimator for this parameter, and used to truncate the data set. If the truncated data set contains none or a trivial proportion of zeros, the remaining coefficients are estimated by the Gaussian Quasi Maximum Likelihood Estimator (QMLE). If a large proportion of zeros exist in the truncated data set, we adopt a Zero-Augmented (ZA) distribution for the random errors in MEM and propose a modified QMLE (ZA-QMLE) without specifying the continuous density to estimate the coefficients in this ZA model. Consistency and asymptotic normality are discussed for both estimators under mild assumptions at order of (1,1). We also conduct simulation studies at both (1,1) and higher orders and empirical analysis on IBM High Frequency trading data, to illustrate the asymptotic results and model improvement for both cases.

1. Introduction

A common model employed in characterizing financial time-varying volatility is the AutoRegressive Conditional Heteroscedasticiy (ARCH) and GARCH model, as proposed by Engle (1982) and Bollerslev (1986),respectively. A variety of structure specifications are further developed [1]. One of the most notable examples is the Autoregressive Conditional Duration model (ACD) of Engle and Russell (1998). It models irregularly-spaced transaction data, via parameterization of the conditional distribution by focusing on time intervals (or durations) between events. In fact, the ACD model is isomorphic to the GARCH model through the use of the square root of duration as dependent variable[2]. Engle (2002) generalizes ACD as the Multiplicative Error Model (MEM), which is focused on non-negativity of the dependent variable and allows the entry of predetermined variables into the model. The MEM can be applied to a wide range of non-negative variables in finance, such as volume of shares traded, the daily high-low range of price, and the ask-bid spread, as discussed in Engle and Gallo (2006).

In many empirical analyses concerning financial data, diurnal adjustment is a necessary procedure in data preparation for modeling financial variables, including duration and price volatility (see Engle and Russell (1998), Hautsch (2002) and Manganelli (2005)). With the assistance of cubic spline or piecewise regression smoothing, diurnal adjustment removes the intradaily seasonality in the data and provides a general daily pattern for these variables. Although it is reasonable to apply this technique to many variables, signs of a significant daily pattern have not been found in variables like volume per trade or cumulative trading volume.

*Department of Biostatistics, University of Kansas Medical Center, 3901 Rainbow Blvd, Kansas City, KS, 66160

[1]For example, Multiplicative GARCH by Geweke (1986), Exponential GARCH by Nelson (1991), Nonlinear Asymmetric GARCH by Engle and Ng (1993), Asymmetric Power ARCH (APARCH) model by Ding et al. (1993) and Threshold GARCH (TGARCH) by Zakoian (1994). See Tsay (2010)

[2]Pacurar (2008) provides a thorough survey on the ACD models and existing extensions.

To avoid the potential overfitting problem caused by smoothing, we investigate another method to directly model nonnegative time series data without diurnal adjustment. Our approach is also inspired by a feature in nonnegative series, which is not captured by existing models. Namely, there exists a non-zero lower bound, for example, in adjusted (or deseasonalized) volume per trade. The absence of such lower bounds in existing MEMs can lead to a poor empirical fit for some variables, and this motivates us to add a location parameter to the original MEM, which provides a 'shift' in the lower bound from zero to any positive number. For this reason, we name the extended model as Location MEM.

In this paper, we investigated the extension to linear MEM(1,1) and develop estimation for the conditional mean equation. There has been substantial research on GARCH type models that directly apply to MEM, such as local estimation for all parameters in stationary GARCH(1,1) in Lee and Hansen (1994), analysis on vector ARMA-GARCH with unit root restriction in Ling and McAleer (2003) and asymptotic properties of the estimator for (G)ARCH parameters in nonstationary ARCH(1) and GARCH(1,1) in Jensen and Rahbek (2004a) and Jensen and Rahbek (2004b). Apparently, this research is all based on Quasi Maximum Likelihood Estimation (QMLE), as well as requires certain assumptions about the stationarity of the process. Ling and McAleer (2003) extend asymptotic inference for all parameters in a more general model, vector ARMA-GARCH(p,q), but it requires assumptions of unit root and higher order moment of the unconditional error. Although Jensen and Rahbek (2004a) and Jensen and Rahbek (2004b) investigate asymptotic theories without requiring moment conditions and stationarity, the results in their work only apply to (G)ARCH parameters. Our work in this paper will follow the assumptions and methods adopted in Lee and Hansen (1994), which insures stationary distribution of the process, develops estimators for all parameters and allows the model to to be integrated or mildly-explosive. We do not consider any exogenous variable in the current case, because exogenous variables, especially those negatively correlated with the dependent variable, may cause more restrictions on parameters.

Because our goal is to estimate parameters in the extended MEM, which does not depend on the true density of disturbance, it is ideal to employ a quasi likelihood function without bringing in other parameters and assumptions on the correct density. For MEM, Engle (2002) has clearly shown that a log likelihood function based on unit exponential disturbance can be interpreted as a quasi likelihood function. In fact, this function coincides with the quasi likelihood for GARCH models in literature such as Lumsdaine (1996) and Jensen and Rahbek (2004a). Hence, we will follow the idea of deriving estimators by a likelihood function of exponential error. In addition, the minimum of the observed data will be demonstrated to be a consistent estimator for the location parameter, under an assumption that we set and a theory in Nelson (1990). This estimator will be substituted into the quasi likelihood function and the QMLE for other parameters is the maximizer of this modified quasi likelihood function. Theorems in Amemiya (1985) pp. 106-111 will be extended to cover the modified QMLE, although the estimator for the location parameter is not asymptotically normal. Eventually, consistency and asymptotic normality will both be developed for the univariate Location MEM(1,1).

The other possible case for a nonnegative financial variable is that the minimum is fixed at a constant value and occurs with a nontrivial proportion in a given sample. Examples include NYSE raw volume per trade and cumulative volume, for which the location parameter (or the minimun) is set at 100 and 0, respectively.

The error term remains unity expectation, but follows a special mixture distribution consisting of a probability mass at zero and a continuous density at positive values. Such a mixture distribution is called Zero Augmented (or Zero Inflated) distribution, abreviated as ZA, and is proposed for the error term of a MEM in Hautsch et al. (2014). They present simulation evidence to show that the Exponential QMLE ignoring the large fraction of zeros are not robust for ZA MEM. The continuous density of the ZA distribution proposed in Hautsch et al. (2014) is a Generalized F distribution. In this paper, we propose ZA Location MEM (1,1) to deal with the large proportion of zeros. The ZA Location MEM(1,1) that we propose can be treated as the Zero-Augmented MEM or ZA MEM(1,1), via subtracting each observation by the constant sample minimum. We further propose density-free ZA-QMLEs for the parameter estimation. Such ZA QMLE in fact coincides with the MLE of ZA-Exponential MEM, whose random errors follow ZA-Exponential(β), where $\beta \neq 1$[3].

The rest of this paper is organized as follows. Section 2 shows specification and assumptions for the extended model, as well as the connection between MEM(1,1) and the extended model. Section 3 provides a consistent estimator for all parameters in Location MEM(1,1) and develops the asymptotic normality for such estimator. Some lemmas in Lee and Hansen (1994) will be carried over to the current case to demonstrate consistency and asymptotic normality. Section 4 provides the specification of ZA MEM(1,1) to account for the case of constant lower bound in Location MEM(1,1). The ZA QMLE is proposed and its asymptotic properties are developed in this section. Section 5 provides simulation examples and empirical improvement using IBM trade volume.

2. The Model

Suppose that we observe a process $\{r_t\}$, $t = 1, ..., n$ and $\theta_0 = (\mu_0, \omega_0, \alpha_0, \beta_0)'$ are the true values of parameters describing this process. A MEM (1,1) process without exogenous variables in the mean equation is usually written as:

$$r_t = h_{0t} z_t, E(z_t) = 1, z_t \geq 0 \tag{1}$$

$$h_{0t} = \omega_0 + \alpha_0 r_{t-1} + \beta_0 h_{0t-1} \tag{2}$$

where $D(1, \phi^2)$ represents the distribution with mean 1 and variance ϕ^2. $\sqrt{h_{0t}}$ is also called the conditional scaling parameter, since the above process can be viewed as a GARCH(1,1) by taking the square root of both sides of equation (1).

2.1 Location MEM(1,1)

Now we consider an improvement on the above model by adding a constant to the right hand side of equation (1) and changing the ARCH term in equation (2). Then we have

$$r_t = \mu_0 + \epsilon_t, \ \epsilon_t = h_{0t} z_t, E(z_t) = 1, z_t \geq 0 \tag{3}$$

$$h_{0t} = \omega_0(1 - \beta_0) + \alpha_0(r_{t-1} - \mu_0) + \beta_0 h_{0t-1} \tag{4}$$

[3]This likelihood function does not contain any parameter other than the probability mass π and MEM coefficients, because the density parameter β can be represented by π as long as the unity mean holds for the process, see equation (13).

where $r_t - \mu_0 \geq 0$ for $t = 1, ..., n$. In other words, the upper bound for μ_0 is the minimum of r_t, denoted by $r_{n(1)}$, that is $\mu_0 \leq r_{n(1)}$. Since $\sqrt{h_{0t}}$ can be viewed as 'scale' , correspondingly, we name μ_0 as the location parameter. The model for the unknown parameters $\theta = (\mu, \omega, \alpha, \beta)'$ is

$$r_t = \mu + \ e_t, \mu \leq r_{n(1)} \tag{5}$$

$$h_t^* = \omega(1-\beta) + \alpha \ e_{t-1} + \beta \ h_{t-1}^*, \ \ h_1^* = \omega \tag{6}$$

For the observed sequence, we have

$$h_t^* = \omega + \alpha \sum_{k=0}^{t-2} \beta^k \ e_{t-1-k} \tag{7}$$

Analogous to the quasi-likelihood estimation of MEM in Engle (2002) and other literature, the observed log likelihood function takes the form:

$$L_n^*(\theta) = \frac{1}{n}\sum_{t=1}^{n} l_t^*(\theta) \qquad l_t^*(\theta) = -\left(\ln h_t^*(\theta) + \frac{e_t}{h_t^*(\theta)}\right) \tag{8}$$

Here we ignore the distribution of z_t, and use the above log likelihood function to derive estimation, because any assumption about the density function will bring in additional parameters. Usually, QMLE is the maximizer of $L_n^*(\theta)$. However, as mentioned the upper bound for μ depends on the observed data and its sample size, hence, the parameter space for the extended model varies with observations. It would be problematic to locate a maximizer of the score function in such a space. Since Location MEM(1,1) can be reduced to MEM(1,1) at $\mu = \mu_0$, it occurred to us that a plausible solution is to separate μ from the other parameters in estimation. The first question that arises is whether there exists a consistent estimator for μ_0 without involving the other parameters. If we can find such an estimator, how do we derive an estimator for the remaining parameters and verify its econometrics properties using the above quasi-likelihood function? These questions will be answered in the following section.

Analogous to the corollary about the exponential ACD model in Engle and Russell (1998), certain assumptions about true innovation terms are necessary.

Assumption 1. Suppose the following conditions are met.
(1). z_t is stationary and ergodic.
(2). z_t is nondegenerate.
(3). $E(z_t^2|\mathscr{F}_{t-1}) < \infty$ a.s.
(4). $\sup_t E(\ln(\beta_0 + \alpha_0 z_t)|\mathscr{F}_{t-1}) < 0$, a.s.
(5). η_0 is in the interior of Θ^*.

Note that condition (3) is actually stronger than necessary for consistency only. According to Lee and Hansen (1994), it is sufficient to establish local consistency for MEM(1,1), if there is some δ such that $E(z_t^{1+\delta}|\mathscr{F}_{t-1}) < \infty$. Existence of the second moment of z_t is just a prerequisite for asymptotic normality. For simplicity, we choose the stronger version of the moment condition in this paper. Condition (4), which is also required in proposition 1 below, not only assures the consistency of $r_{n(1)}$, but also serves as a sufficient condition for stationarity and ergodicity of h_{0t}. Although the analysis in this paper only focuses on stationary processes, we plan to investigate similar properties for a nonstationary case in the future, that is when $\sup_t E(\ln(\beta_0 + \alpha_0 z_t)|\mathscr{F}_{t-1}) \geq 0$, a.s. as in Jensen and Rahbek (2004a) and

Jensen and Rahbek (2004b). When $\alpha_0 + \beta_0 \leq 1$, condition (4) is automatically satisfied due to Jensen's inequality. But it is not necessary to require $\alpha_0 + \beta_0 \leq 1$ in current case. Asymptotic properties of the local estimator can be established under condition (4), which allows integrated and mildly explosive cases.

2.2 Connection

An obvious relation between Location MEM and MEM is that Location MEM reduces to MEM when μ is zero. In fact, there is another transition from Location MEM(1,1) to MEM(1,1) with certain assumptions.

Suppose r_t is a process described by equation (3). Define $x_t = \dfrac{\mu_0 + h_{0t} z_t}{\mu_0 + h_{0t}}$. It can be shown that $E(x_t|\mathscr{F}_{t-1}) = 1$ and x_t is stationary and ergodic. Let $\psi_{0t} = \mu_0 + h_{0t}$, then

$$\begin{aligned} r_t &= \psi_{0t} x_t \\ \psi_{0t} &= \omega_0' + \alpha_0 r_{t-1} + \beta_0 \psi_{0t-1} \end{aligned} \tag{9}$$

where $\omega_0' = \omega_0 + (1 - \alpha_0 - \beta_0)\mu_0$. Hence, Location MEM(1,1) can be transformed to MEM(1,1) with ARCH and GARCH parameters unchanged. In order to ensure positivity of ω_0' and distinguish it from ω, here we only consider $\alpha_0 + \beta_0 < 1$. Thus, Assumption 1 (1),(3),(4) and (5) are easily satisfied by x_t. However, the range of x_t is slightly different from that of z_t. When $\mu_0 \neq 0$, x_t is strictly positive and the exponential distribution is excluded from the alternative true densities. Lee and Hansen (1994) have demonstrated that no matter what density function the innovation term has, the log likelihood function in the form of unity exponential always produces a consistent and asymptotically normal estimation under certain conditions. Therefore, if all parts of Assumption 1 are satisfied by x_t, the model represented in (9) can still be estimated by QMLE regardless of the range of x_t. The only concern is whether condition (2) can be trivially met if μ_0 is too large. When $h_{0t}/\mu_0 \to 0$, $x_t \to 1$ a.s., that is x_t is degenerate. In that case, it is impossible to ensure consistency and asymptotic normality for the MEM(1,1). Therefore, Location MEM(1,1) can be correspondingly viewed as MEM(1,1) only if μ_0 is not too large compared to h_{0t}, which is difficult to evaluate in real data.

3. Asymptotic Properties for Modified QMLE

In order to estimate Location MEM(1,1) with QML as mentioned in subsection 2.1, the first task is to locate a consistent estimator of μ_0. The second question that arises is how to derive a consistent and asymptotically normal estimator for the remaining parameters. If consistency in estimation for μ_0 is established, is it valid to substitute the estimator of μ_0 for μ in likelihood function (8) and then maximize it over $\omega_0, \alpha_0, \beta_0$? Before we explore the asymptotic properties of the estimator for $\omega_0, \alpha_0, \beta_0$, the primary issue is to validate convergence results based on a score function composed from observed data, unknown parameters and a partial estimator.

3.1 A Consistent Estimator of the Location Parameter

If the observations are independently and identically distributed as Exponential, Weibull or Generalized Gamma, it is not hard to show that the minimum statistic of the sample converges to the location parameter in the density function. Inspired by this property, for a process described by equations (3) and (4), $r_{n(1)}$ might be a consistent estimator for μ_0. However, as mentioned we ignore the true distribution of z_t. The distribution of r_t cannot be specified due to the lack of an assumption on z_t's density. Therefore, it is challenging to draw the following conclusion.

PROPOSITION 1. Under Assumption 1, $r_{n(1)} \to_p \mu_0$.

Estimating μ_0 so far has not involved any of the other parameters. For Location MEM(p,q), the convergence of $r_{n(1)}$ can also be validated under certain conditions of stationarity. Note that normality is not applicable to $r_{n(1)}$, because it has a lower bound of μ_0. Now that the consistency of the estimator for μ is demonstrated, we will move on to the modified QMLE for ω, α and β.

3.2 Asymptotic Properties of The Modified QMLE

In order to utilize the above two theorems, we have to first specify and draw some conclusions on the limit of $L_n^*(r_{n(1)}, \eta)$ at first, which is an unobserved likelihood, defined as

$$L_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} l_t(\theta) \qquad l_t(\theta) = -\left(\ln h_t(\theta) + \frac{e_t}{h_t(\theta)}\right)$$

where

$$h_t = \omega + \alpha \sum_{k=0}^{\infty} \beta^k e_{t-1-k} \tag{10}$$

Equation (10) indicates that $h_t(\theta_0) = h_{0t}$, because when $\theta = \theta_0$, $e_t = r_t - \mu_0 = \epsilon_t$. Conclusions on $L_n(\theta)$ and $L_n^*(r_{n(1)}, \eta)$ are all presented as lemmas in Appendix A.

In order to establish asymptotic properties, the parameter space is restricted to be: $\Theta = \{\eta : 0 \le \omega_l \le \omega \le \omega_u, 0 \le \alpha_l \le \alpha \le \alpha_u, 0 \le \beta_l \le \beta \le \beta_u < 1\}$, where $\eta = (\omega, \alpha, \beta)$. There is no need to discuss the estimator for μ, as the consistency of $r_{n(1)}$ is demonstrated in proposition 1 and asymptotic normality is not applicable to it. We will directly apply results in Lee and Hansen (1994) to the current model with μ fixed at μ_0. Similar to the analysis about GARCH (1,1) in that paper, we need to split the parameter space in the same way to bound h_t.

Let $\mathscr{R}_l = \mathscr{R}(K_l^{-1}\alpha_l) < 1$ and pick positive constants η_l and η_u, which satisfy $\eta_l < \beta_0(1 - \mathscr{R}_l^{1/6})$ and $\eta_u < \beta_0(1 - \mathscr{R}_0^{1/6})$.

Define for $1 \le r \le 6$ the constants $\beta_{rl} = \beta_0 \mathscr{R}_l^{1/r} + \eta_l < \beta_0$, $\beta_{ru} = \dfrac{\beta_0 - \eta_u}{\mathscr{R}_0^{1/r}} > \beta_0$,

and the subspaces
$\Theta_r^l = \{\eta \in \Theta : \beta_{rl} \le \beta \le \beta_0\}$, $\Theta_r^u = \{\eta \in \Theta : \beta_0 \le \beta \le \beta_{ru}\}$,
$\Theta_r = \Theta_r^l \bigcup \Theta_r^u$. When $r = 2$, $\Theta_2 = \Theta_2^l \bigcup \Theta_2^u$.
Now we can define the modified QMLE and prove its local asymptotic properties.

Definition 1. Local modified QMLE: $\hat{\eta}_n(r_{n(1)}) = \arg\max_{\eta \in \Theta_2} L_n^*(r_{n(1)}, \eta)$.

Let ∇_η denote the first order gradient w.r.t $\eta' = (\omega, \alpha, \beta)$.
Set $\hat{G}_n(\theta) = -\frac{1}{n}\sum_{t=1}^{n} \nabla_\eta^2 l_t^*(\theta)$ and $\hat{C}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} \nabla_\eta l_t^*(\theta) \nabla_\eta l_t^*(\theta)'$.

Assumption 2. $E|r_{n(1)} - \mu_0|^s = o(\frac{1}{\sqrt{n}})$ for some $s < 1$.

THEOREM 3.1. *Under Assumptions 1 and 2,*
(1) $\hat{\theta}_n \to_p \theta_0$.
(2) $\sqrt{n}(\hat{\eta}_n(r_{n(1)}) - \eta_0) \to_D N(0, V_0)$, *where* $V_0 = G_0^{-1} C_0 G_0^{-1}$, $C_0 = E(\nabla_\eta l_t(\theta_0) \nabla_\eta l_t(\theta_0)')$ *and* $G_0 = -E\nabla_\eta^2 l_t(\theta_0)$.
(3) $\hat{V}_n = \hat{G}_n^{-1}(\hat{\theta}_n)\hat{C}_n(\hat{\theta}_n)\hat{G}_n^{-1} \to_p V_0 = G_0^{-1} C_0 G_0^{-1}$.

Now let's restrict attention to nonintegrated processes. We can establish consistency of the global modified QMLE with the same assumptions.

Definition 2. Global modified QMLE: $\tilde{\theta}_n = (r_{(1)}, \arg\max_{\eta \in \Theta} L_n^*(r_{(1)}, \eta))$.

THEOREM 3.2. *Under Assumption 1 and* $\alpha_0 + \beta_0 < 1, \tilde{\theta}_n \to_p \theta_0$.

PROPOSITION 2. If $\alpha_0 + \beta_0 < 1$, $L_n^*(\theta)$ is maximized at $\mu = r_{n(1)}$ for any given η.

Therefore, if we derive the global estimator for a nonintegrated Location MEM(1,1) by maximizing $L_n^*(\theta)$ over all parameters, the optimal solution coincides with the global modified QMLE. That is, $\tilde{\theta}_n$ is equivalent to $\arg\max_{\theta \in \Theta} L_n^*(\theta)$ when $\alpha_0 + \beta_0 < 1$.

4. Zero-Augmented MEM and ZA-QMLE

According to the results in the previous section, a positive process can be modeled by the Location MEM and estimated by applying Gaussian QMLE to the truncated process. There must be at least one zero in the truncated data set. If the minimum of the observations occur at a nontrivial proportion, there must be a nontrivial proportion of zeroes in the truncated series, see the plot in Figure 1 for truncated data of IBM volume per trade in NYSE. In order to capture such clustering at a zero, a Zero-Augmented distribution is appropriate for the random errors in the truncated process.

4.1 Zero-Augmented MEM

If a nonnegative continuous r.v. X has a non-trivial proportion of zero realizations, a discrete probability mass is assigned as following:

$$\pi = P(X > 0), \; 1 - \pi = P(X = 0) \tag{11}$$

The distribution of X has a semicontinuous density as

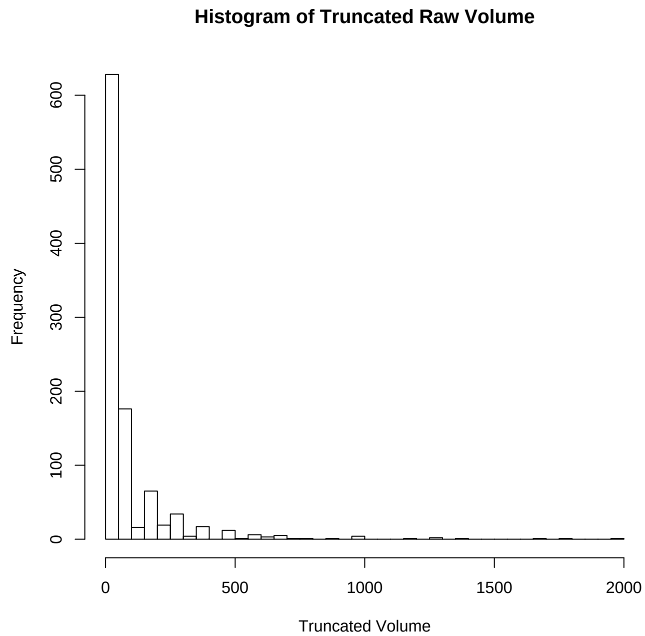
$$f_X(x) = (1 - \pi)^{1 - I_{(x>0)}} [\pi g(x|\boldsymbol{\xi})]^{I_{(x>0)}} \tag{12}$$

where $g(x|\boldsymbol{\xi})$ is the continuous density at nonzero values with parameters $\boldsymbol{\xi}$. X is called a Zero-Augmented (ZA) random variable. Let $e(\xi)$ be the unconditional expectation of the continuous distribution $g(x|\boldsymbol{\xi})$, then

$$E(X) = \pi e(\xi) \tag{13}$$

Examples of the continuous part of density function $g(x|\boldsymbol{\xi})$ include, but not limited to, density functions of Exponential, Weibull, Gamma, Generalized F and

Figure 1: Histogram of IBM Truncated Raw Volume in NYSE



Log-logistic. As proposed in Hautsch et al. (2014), the Generalized F (GF) distribution has a density function

$$g(x|\boldsymbol{\xi}) = \frac{dx^{dm-1}[\phi + (x/\lambda)^d]^{(-\phi-m)}\phi^{\phi}}{\lambda^{dm}\mathcal{B}(m,\phi)} \tag{14}$$

where $\boldsymbol{\xi} = (d, m, \phi, \lambda)$, $d > 0, m > 0, \phi > 0$ and $\lambda > 0$, $\mathcal{B}(m, \eta) = \dfrac{\Gamma(m)\Gamma(\eta)}{\Gamma(m+\eta)}$. Such density function nests several distributions. When $m = 1$ and $\phi = 1$, it is Log-Logistic. If $\phi \to \infty$, it is approaching Generalized Gamma. If $\phi \to \infty$ and $m = 1$, it reduces to Weibull.

A MEM(p,q) process incorporating Zero-Augmented random errors is the so-called ZA-MEM. Let $\{r_t\}_{t=1}^n$ be a series of nonnegative observations, modeled by the Location MEM with location parameter $\mu = \mu_0$, where μ_0 is known. Let $y_t = r_t - \mu_0$, then y_t can be described by a ZA-MEM with true parameters as

$$y_t = h_{0t} z_t \tag{15}$$

$$h_{0t} = \omega + \alpha_0 y_{t-1} + \gamma_0 I_{(y_{t-1}=0)} + \beta_0 h_{0,t-1} \tag{16}$$

where z_t's are i.i.d. and $E(z_t) = 1$. An indicator variable $I_{(y_{t-i}=0)}$ is included in current research, as specified in the autoregressive equation (16), to account for the potential asymmetric impact of zero observations on h_t, similar to the structure of Asymmetric GARCH (AGARCH) and Augmented ACD (AACD), (see Hentschel (1995) and Fernandes and Grammig (2006)). If z_t's follow ZA-GF with $g(x|\boldsymbol{\xi})$ defined by (14), then y_t is called a ZA-GF MEM. Since $E(z_t) = 1$, then $\lambda = (\pi\zeta)^{-1}$, where $\zeta = \phi^{1/d}[\Gamma(m + 1/d)\Gamma(\phi - 1/d)][\Gamma(m)\Gamma(\phi)]^{-1}$. However, the GF density function is not the optimal choice for $g(x|\boldsymbol{\xi})$, as it brings in additional density parameters and the estimate of ϕ becomes biased if the true density of z_t is Generalized Gamma, in which the true value of ϕ is ∞.

4.2 ZA-QMLE and Asymptotic Properties

The Gaussian QMLE may not be valid for Zero-Augmented MEM due to the large fraction of zeroes. Hence, we propose an estimator based on the following quasi likelihood.

Let $\boldsymbol{w}$ be the unknown parameters vector for $\boldsymbol{w}_0 = (\omega_0, \alpha_0, \gamma_0, \beta_0)$, π be unknown parameter for true probability mass π_0 and $\boldsymbol{v} = (\pi, \boldsymbol{w})$. The likelihood function $L_n^{**}(\boldsymbol{v})$ is defined as

$$\begin{aligned} L_n^{**}(\boldsymbol{v}) &= \frac{1}{n}\sum_{t=1}^{n} l_t^{**}(\boldsymbol{v}) \\ l_t^{**}(\boldsymbol{v}) &= I_{(y_t=0)}\ln(1-\pi) + I_{(y_t>0)}\left(\ln\frac{\pi^2}{h_t(\boldsymbol{w})} - \frac{\pi y_t}{h_t(\boldsymbol{w})}\right) \end{aligned} \tag{17}$$

where $h_t(\boldsymbol{w})$ is h_{0t} in equation (16) in terms of unknown parameters $\boldsymbol{w}$.

If the true density of z_t is ZA-Exponential(β), equation (13) and $E(z_t) = 1$ imply $\pi\beta = 1$ and $g(x|\boldsymbol{\xi}) = \pi e^{-\pi x}$, which leads to a true log likelihood function that coincides with equation (17). In other words, the quasi likelihood function that we propose in (17) is the derived by the true likelihood function of ZA-Exponential MEM, but does not require any specification on the continuous part of density function $g(x|\boldsymbol{\xi})$. The parameters to be estimated are still the propability mass π and those in (16). There is no additional parameters brought into the model and estimation does not depend on the true distribution of errors. Therefore, this is an ideal likelihood function for ZA-MEM.

Definition 3. $\hat{\boldsymbol{v}} = \arg\max_{\boldsymbol{v}\in\Theta_2} L_n^{**}(\boldsymbol{v})$, where Θ_2 is the compact parameter space for $\boldsymbol{v}$.

The following theorem states that $\hat{\boldsymbol{v}}$ serves as the Quasi MLE for ZA-MEM, named as ZA-QMLE.

THEOREM 4.1. *Under Assumptions 1 and 2,*
(1) $\hat{\boldsymbol{v}} \to_p \boldsymbol{v}_0$.
(2) $\sqrt{n}(\hat{\boldsymbol{v}} - \boldsymbol{v}_0) \to_D N(0, V_0)$, *where* $V_0 = G_0^{-1}C_0G_0^{-1}$, $C_0 = -E(\nabla_{\boldsymbol{v}} l_t(\boldsymbol{v}_0)\nabla_{v'} l_t(\boldsymbol{v}_0))$ *and* $G_0 = -E\nabla^2_{\boldsymbol{v}} l_t(\boldsymbol{v}_0)$.
(3)$\hat{V}_n = \hat{G}_n^{-1}(\hat{\boldsymbol{v}})\hat{C}_n(\hat{\boldsymbol{v}})\hat{G}_n^{-1}(\hat{\boldsymbol{v}}) \to_p V_0$, *where* $\hat{G}_n(\boldsymbol{v}) = -\frac{1}{n}\sum_{t=1}^{n}\nabla^2_{\boldsymbol{v}} l_t(\boldsymbol{v})$,
$\hat{C}_n(\boldsymbol{v}) = \frac{1}{n}\sum_{t=1}^{n}\nabla_{\boldsymbol{v}} l_t(\boldsymbol{v})\nabla_{\boldsymbol{v}'} l_t(\boldsymbol{v})$.

5. Simulation and Empirical Results

In this section , we will illustrate asymptotic results in the previous sections as well as the relation between Location MEM and MEM, by both simulation and using real data. GARCH software with QMLE can be used to fit a positive series into either Location MEM or MEM, by taking the square root of the input variable and setting the mean to be zero. For Location MEM, the input is the observed data from which the minimum of the sample is subtracted.

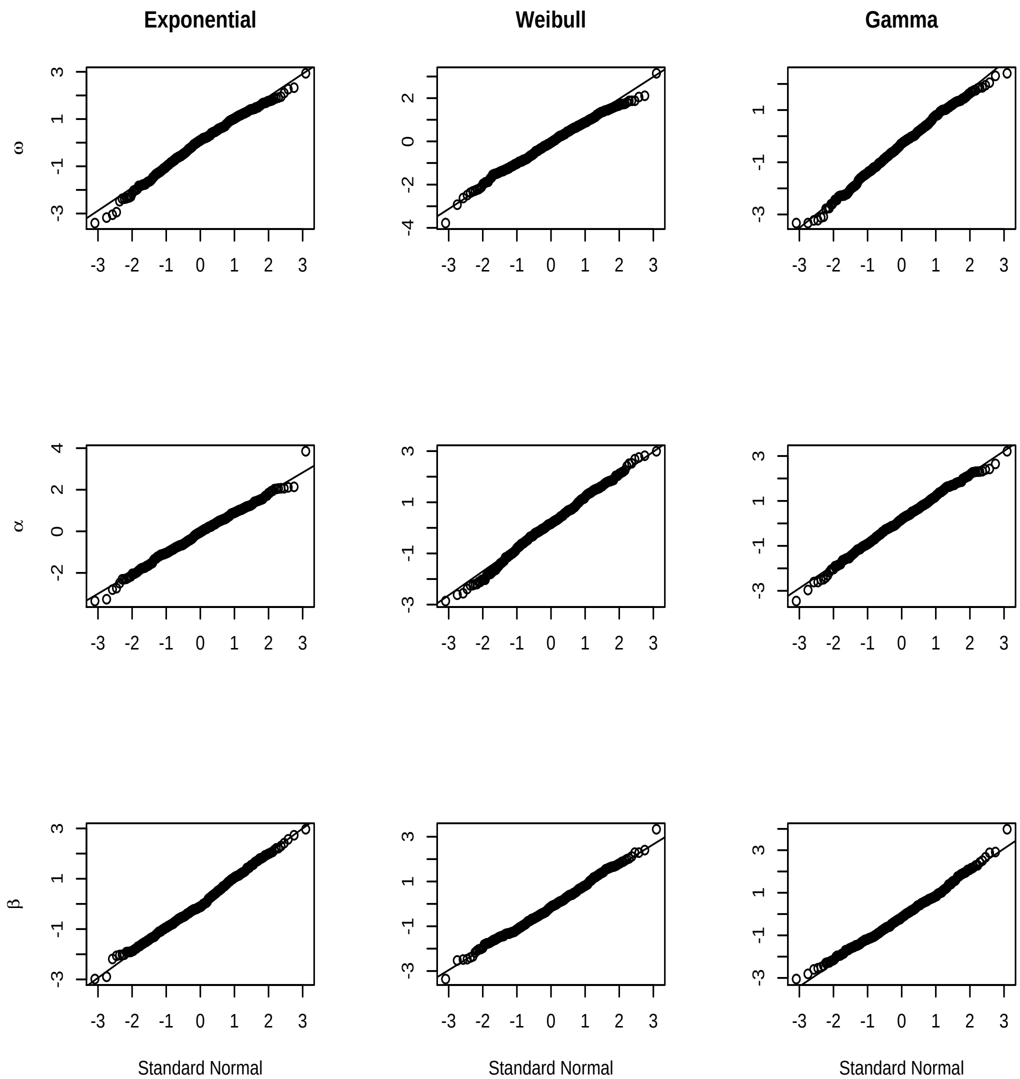


Figure 2: QQ Plot of standardized estimates for different distributions

5.1 Simulation

In the first step, we generate 500 Location MEM(1,1) data sets by a DGP (Data Generating Process) with sample size n=10000 and parameter values at $(\mu_0, \omega_0, \alpha_0, \beta_0) = (0.5, 0.03, 0.1, 0.85)$ for alternative innovation term distributions: Exponential, Weibull and Gamma with unity mean. Asymptotic properties of modified QMLE for Location MEM(1,1) with different innovation distributions are confirmed by the QQ plots of standardized estimates displayed in Figure 2. The standardized estimate in each QQ plot is $\sqrt{n}(\hat{\eta}(r_{n(1)}) - \eta_0)$ standardized by $\hat{V}_n$ (defined in Theorem 3.1 (3)).

In the second step, we illustrate asymptotic results in for Location MEM at a higher order (2,2), along with the contrast between MEM(2,2) and Location MEM(2,2) by 500 simulated data sets with sample size n=30000 from Location Gamma-MEM(2,2) with parameter values $(\omega_1, \alpha_1, \alpha_2, \beta_1, \beta_2) = (0.03, 0.05, 0.1, 0.5, 0.3)$.

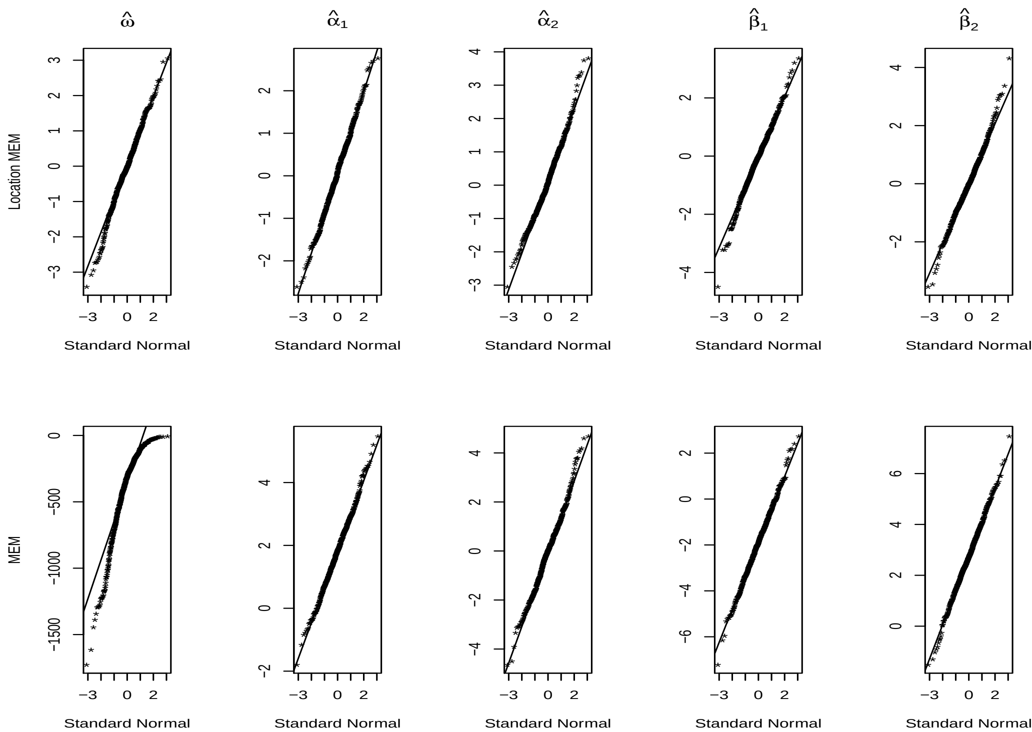


Figure 3: QQ Plot of standardized estimates for $\mu_0 = 8$ at order (2,2)

Consequently, parameters for the corresponding MEM(2,2) process are $(\omega_2, \alpha_1, \alpha_2, \beta_1, \beta_2) = (0.08, 0.05, 0.1, 0.5, 0.3)$.

The estimation result in Table 1 not only confirms consistency and asymptotic normality of the modified QMLE for a Location MEM(2,2) process, but also presents the bias in the estimation of MEM(2,2) when μ_0 is relatively large. Violation of asymptotic normality in MEM(2,2) is shown by the QQ-plot in Figure 3 as well. The tremendous contrast between two models at order of (2,2) implies that there must exist non-negligible improvement brought by location parameter at order of (1,1).

Table 1: Estimates of Simulated Data by Two Models

	Location MEM					MEM				
	$\hat{\omega}_1$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\omega}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
True	0.03	0.05	0.1	0.5	0.3	0.08	0.05	0.1	0.5	0.3
$\mu_0 = 1$										
Mean	0.029	0.05	0.1	0.503	0.297	0.081	0.05	0.1	0.5	0.297
S.D.	0.002	0.006	0.009	0.085	0.074	0.007	0.006	0.009	0.087	0.076
$\mu_0 = 8$										
Mean	0.03	0.05	0.1	0.503	0.296	0.0055	0.06	0.1	0.408	0.433
S.D.	0.003	0.006	0.009	0.084	0.074	0.0001	0.006	0.008	0.072	0.066

Next, we generate 1000 samples from DGP of ZA MEM(2,2) described by (15) and (16) with $\pi_0 = 0.5$ and sample size $n = 7500$. The random errors are from Log-normal$(0, 2\ln 2)$ and Gamma$(0.5, 4)$. The simulation results in table 2 not only illustrates the consistency and asymptotic normality of ZA QMLE, but also shows an obvious contrast between ZA QMLE and Gaussian QMLE. The QQ plots in Figure 4 imply that Gaussian QMLE may require a larger sample size in the convergence of standardized estimates to standard normal distribution, in comparison to ZA QMLE. Such contrast between two estimators at order (2,2) also implies better performance of ZA QMLE at order (1,1)

Table 2: Coefficients Estimates of DGP with 50% Zero Observations

	ZA QMLE				Gaussian QMLE			
	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}$
True	.03	.05	.05	.8	.03	.05	.05	.8
				Log Normal$(0, 2\ln 2)$				
Mean	.0326	.0515	.0518	.7887	.0327	.0517	.0526	.7870
Med.	.0294	.0501	.0511	.7991	.0288	.0503	.0528	.7990
RMSE	.0185	.0148	.0158	.0679	.0210	.0163	0.0170	0.0769
				Gamma (0.5, 4)				
Mean	.0323	.0503	.0505	.7926	.0327	.0505	.0514	.7901
Med.	.0295	.0501	.0505	.7963	.0293	.0499	.0512	.7979
RMSE	.0164	.0110	.0125	.0570	.0199	.0122	.0139	.0683

5.2 Empirical Application on IBM Data

The first application of MEM is to model trading duration in a high-frequency market, proposed as ACD by Engle and Russell (1998). Later, it was applied to other positive processes, such as volume, bid-ask spread and price return volatility. Hence, we first explored on the difference between Location MEM(1,1) and MEM(1,1) by fitting eaching model to IBM transaction duration data. However, we noticed only trivial improvement in log likelihood and the Ljung-Box test statistics brought by the location parameter. The trivial improvement is caused by the fact that the minimum of the observations is too small, since the ultimate transaction duration-one millisecond-is frequently reached and the average level of transaction duration is much larger than 0.001 second for IBM, as mentioned in section 3.

On the other hand, we found that trading volumes have a relatively large minimum compared to duration. Furthermore, according to Manganelli (2005), high volume may increase price volatility in the next trade. Hence, trade volume is a key economic element to be molded and forecasted. We use IBM trading volume in the NYSE with the time span April 8th-12th, 2013, taken from Trade and Quotes(TAQ) database, to investigate the contrast between Location MEM and MEM. Applying these two models to IBM trading volume raw data, we found significant improvement in Location MEM(1,1) as shown in Table 3. There still exists remained autocorrelation in the residuall of Location MEM(1,1), indicating that (1,1) is not the optimal order of lags. Based on BIC and significance of parameters, (1,2) is the optimal order of lags for both models on the raw data. The improvement on goodness-of-fit brought by the location parameter is more pronounced at lag of (1,2), compared to (1,1), as shown by the Q-statistics in Table 4. Standard Errors and likelihood function values in this table also illustrate that Location MEM(1,2) greatly outperforms the traditional MEM on raw data.

Next, we remove the intraday pattern by smoothing with a piecewise regression

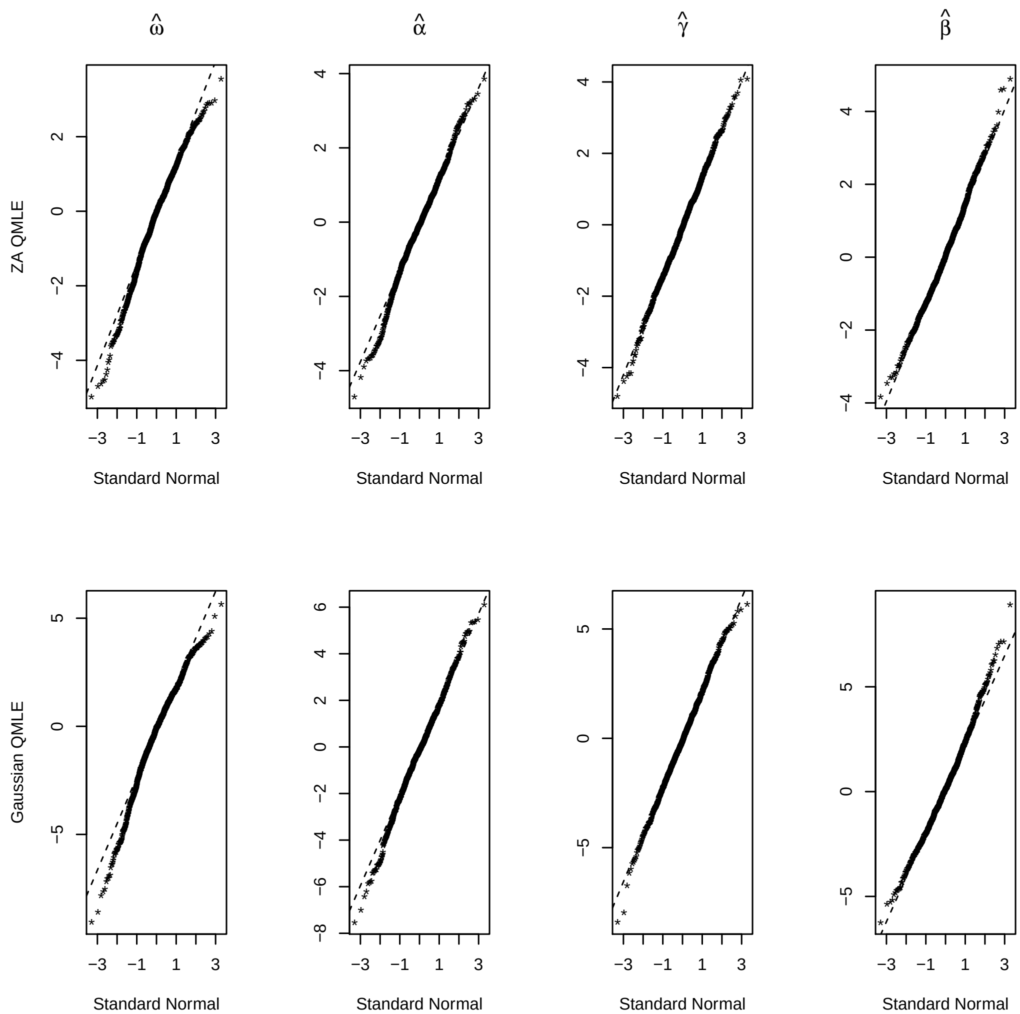


Figure 4: QQ Plots of standardized ZA QMLE and Gaussian QMLE with n=15000

Table 3: Model Improvement on Raw IBM Trading Data at Order (1,1)

	Location MEM			MEM		
	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\omega}'$	$\hat{\alpha}$	$\hat{\beta}$
Estimate	1.4057	0.019	0.966	3.331	0.022	0.961
Std. Error	0.4043	0.0029	0.0063	0.9216	0.0029	0.0063
Log Likelihood		-56778.3			-62600.27	
Ljung-Box Q(30)		38.611			49.0837	
p Value of Q(30)		0.04027			0.00276	

Table 4: Improvement on Raw IBM Trading Data at Order of (1,2)

	Location MEM				MEM			
	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\omega}'$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\beta}_2$
Estimate	2.0548	0.0285	0.4398	0.5097	4.8176	0.0328	0.4137	0.5286
Std. Error	0.6803	0.0047	0.1248	0.1251	1.6745	0.0052	0.1509	0.1531
Log Likelihood		-56772.35			-62598.45			
L-B Q(15)		23.6975			32.2416			
p Value of Q(15)		0.07041			0.005968			

function, as used by Engle and Russell (1998). We select knots with interval of 30 minutes from 10 AM to 4 PM to compute diurnal factors. The adjusted series is the original data divided by diurnal factors. Table 5 and 6 show the estimates for the deseasonalized data by two models at order (1,1) and (1,2). The improvement on goodness-of-fit can still be observed in likelihood function and Q-statistics.

The transformation between two models also presents a deficiency in MEM, based on the estimates in Table 4 and 6. The minimum is 100 for the raw data and 0.35243 for the adjusted data, which can be viewed as the estimates of the location parameter. Estimates of the parameters in h_{0t} are $(2.0548, 0.0285, 0.4398, 0.5097)$ for the raw data, and $(0.0429, 0.032, 0.4802, 0.4187)$ for the adjusted data. Hence, equation (9) indicates that estimated ω' in MEM(1,2) should be 4.2548 for raw data and 0.0673 for deseasonalized data. However, results in the above two tables illustrate that the estimates by MEM(1,2) for both data sets are biased. Therefore, combined with improvement in log likelihood and Q-statistics, Location MEM appears superior to MEM at order (1,2).

6. Conclusion

This paper This chapter illustrates that, under weak conditions, the estimator for the location parameter and the modified QMLEs for the other parameters in Location MEM and ZA MEM at order of (1,1) are both consistent. Asymptotic normality of the modified QMLEs are also well developed. Asymptotic properties and the connection between MEM(1,1) and the extended model as described by equation (9), are illustrated by both simulated data sets and real data of IBM trading.

Table 5: Model Improvement on Adjusted IBM Trading Data at Order(1,1)

	Location MEM			MEM		
	ω	α	β	ω'	α	β
Estimate	0.034	0.024	0.922	0.077	0.028	0.893
Std. Error	0.0189	0.0069	0.0361	0.0482	0.0096	0.0578
Log Likelihood	-18210.63			-21718.21		
Ljung-Box Q(30)	37.6396			39.4853		
p Value of Q(30)	0.05014			0.03291		

Table 6: Improvement on Adjusted IBM Trading Data at Order of (1,2)

	Location MEM				MEM			
	ω	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\omega}'$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\beta}_2$
Estimate	0.0429	0.032	0.4802	0.4187	0.0892	0.0353	0.4755	0.3975
Std. Error	0.0218	0.0081	0.1549	0.1601	0.0453	0.0097	0.1655	0.1741
Log Likeilhood	-18208.07				–21717.05			
L-B Q(12)	20.3357				21.1042			
p Value of Q(12)	0.061				0.04888			

The model and assumptions discussed in this paper focus on order (1,1). There exist open questions in other versions of Location MEM. The first and the most interesting one must be estimation of Location MEM and ZA MEM(p,q) under conditions of stationarity, ergodicity and moments for GARCH(p,q) as generalized in Bougerol and Picard (1992). Berkes et al. (2003) discuss the structure of GARCH(p,q) and establish its asymptotic properties under this condition. Berkes and Horvath (2004) propose a class of estimators under the same condition by introducing various density-based likelihood functions and investigates their efficiency. It would be ideal to extend current results to Location MEM(p,q) following the methodology of Berkes et al. (2003). Following the empirical improvement for order of (2,2) presented in section 5, a generalized model and asymptotic analysis for order of (p,q) is investigated in Li (2016)

Since current work is conducted under mild conditions for stationarity and ergodicity, it will be of great interest to consider a nonstationary process in future by the approach used in Jensen and Rahbek (2004a) and Jensen and Rahbek (2004b), which utilizes other convergence theorems and imposes different assumptions. Another promising direction on this model is allowing for explanatory variables (or covariates) in the dynamic component, as in GARCH-X in Han and Kristensen (2014). Similar to the investigation about GARCH-X in Han (2015), asymptotic results concerning sample autocorrelation, variance and kurtosis can be built for Location MEM as well. Analysis on forecasting can also be developed by incorporating exogenous variables in the conditional mean equation. It would be practical to develop an estimator for a linear multivariate location MEM, in which different variables of a mark are jointly modeled, analogous to the framework in Manganelli

(2005). Asymptotic properties of such multivariate model can be established under weak conditions of stationarity, because entry of predetermined variables in each equation excludes integrated and explosive processes.

Moreover, various GARCH-type models imply that considerable attention should be paid in the future to nonlinear structures of Location MEM in the sense of Log-ACD in Bauwens and Giot (2000), TGARCH in Zakoian (1994) and Asymmetric ACD (AACD) in Fernandes and Grammig (2006). Another methodology to capture nonlinearity in this model is the regime-switching approach or TACD proposed by Zhang et al. (2001). Obviously, the location parameter in a regime-switching model varies along with threshold values. One should be careful while selecting the thresholds for Location MEM, because there is a possible impact of threshold values on the estimate of the location parameter, and precision in this estimate largely depends on the number of observations within each regime. Although existing results provide evidence that nonlinear structures are usually more adequate to model duration, a linear system, on the other hand, is more ideal while taking into account the interaction between variables.

References

Amemiya, T. (1985). *Advanced Econometrics.* Cambridge: Harvard University Press.

Bauwens, Luc and Giot, Pierre (2000). "The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks". *Annales d'conomie et de Statistique* 60, pp. 117–149.

Berkes, Istvan and Horvath, Lajos (2004). "The Efficiency of the Estimators of the Parameters in GARCH Processes". *The Annals of Statistics* (32), pp. 633–655.

Berkes, Istvan, Horvath, Lajos, and Kokoszka, Piotr (2003). "GARCH Processes: Structure and Estimation". *Bernoulli* (9), pp. 201–227.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity". *Journal of Econometrics* 31, pp. 307–327.

Bougerol, Philippe and Picard, Nico (1992). "Strict Stationarity of Generalized Autoregressive Processes". *The Annals of Probability* 20, pp. 1714–1730.

Ding, Zhuanxin, Granger, Clive, and Engle, Robert (1993). "A long memory property of stock market returns and a new model". *Journal of Empirical Finance* 1.1, pp. 83–106.

Engle, Robert (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom inflation". *Econometrica* 50, pp. 987–1007.

— (2002). "New frontiers for ARCH models". *Journal of Applied Econometrics* 17.5, pp. 425–446.

Engle, Robert and Gallo, Giampiero M. (2006). "A Multiple Indicators Model for Volatility Using Intra-daily Data". *Journal of Econometrics* 131, pp. 3–27.

Engle, Robert and Ng, Victor K. (1993). "Measuring and Testing the Impact of News on Volatility". *The Journal of Finance* 48, pp. 1749–1778.

Engle, Robert and Russell, Jefferey (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data". *Econometrica* 66, pp. 1127–1162.

Fernandes, M and Grammig, J (2006). "A Family of Autoregressive Conditional Duration Models." *Journal of Econometrics* 130, pp. 1–23.

Geweke, John (1986). "Modeling the Persistence of Conditional Variance: A Comment". *Econometric Reviews* 5, pp. 57–61.

Han, Heejoon (2015). "Asymptotic Properties of GARCH-X Processes". *Journal of Financial Econometrics* 13, pp. 188–221.

Han, Heejoon and Kristensen, Dennis (2014). "Asymptotic Theory for the QMLE in GARCH-X Models with Stationary and Nonstationary Covariates". *Journal of Business & Economic Statistics* 32.3, pp. 416–429.

Hautsch, Nikolaus (2002). "Modelling Intraday Trading Activity Using Box-Cox ACD Models". *Discussion Paper, CoFE, University of Konstanz.*

Hautsch, Nikolaus, Malec, Peter, and Schienle, Melanie (2014). "Capturing the Zero: A New Class of Zero-Augmented Distributions and Multiplicative Error Processes". *Journal of Financial Econometrics* 12, pp. 89–121.

Hentschel, L. (1995). "All in the Family: Nesting Symmetric and Asymmetric GARCH Models". *Journal of Financial Economics* 39, pp. 71–104.

Jensen, S.T. and Rahbek, Anders (2004a). "Asymptotic Inference for Nonstationary GARCH". *Econometric Theory* 20.

— (2004b). "Asymptotic Normality of the QMLE Estimator for ARCH in the Non-Stationary Case". *Econometrica* 72, pp. 641–646.

Lee, S.W. and Hansen, B.E. (1994). "Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator". *Econometric Theory* 10, pp. 29–52.

Li, Qian (2016). "Modified QMLEs for Location and Zero-Augmented Multiplicative Error Models". *Social Science Research Network, Working Paper.*

Ling, Shiqing and McAleer, Michael (2003). "Asymptotic Theory for a Vector ARMA-GARCH Model". *Econometric Theory* 19, pp. 280–310.

Lumsdaine, R.L. (1996). "Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models". *Econometrica* 64, pp. 575–596.

Manganelli, Simone (2005). "Duration, Volume and Volatility Impact of Trades". *Journal of Financial Markets* 8, pp. 377–399.

Nelson, D.B. (1990). "Stationarity and Persistence in the GARCH(1,1) Model". *Econometric Theory* 6, pp. 318–334.

— (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". *Econometrica* 59, pp. 347–370.

Pacurar, Maria (2008). "Autoregressive Conditional Duration Models in Finance: A Survey of the Theoretical and Empirical literature". *Journal of Economic Surveys* 22, pp. 711–751.

Tsay, Ruey (2010). *Analysis of Financial Time Series.* John Willey and Sons.

Zakoian, Jean-Michel (1994). "Threshold Hereroskedastic Models". *Journal of Economic Dynanmic and Control* 18, pp. 931–955.

Zhang, Michael Yuanjie, Russell, Jeffrey R, and Tsay, Ruey S (2001). "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data". *Journal of Econometrics* 104.1, pp. 179–207.