Article:

Koul, H.L., Perera, I. and Balakrishna, N. (2023) A class of minimum distance estimators in Markovian multiplicative error models. Sankhya B, 85 (Suppl 1). pp. 87-115. ISSN 0976-8386

https://doi.org/10.1007/s13571-021-00274-x

A class of Minimum Distance Estimators in Markovian Multiplicative Error Models

Hira L. Koul
Michigan State University

Indeewara Perera
The University of Sheffield

Narayana Balakrishna
Cochin University of Science and Technology

Abstract

This paper proposes a class of minimum distance estimators for the underlying parameters in a Markovian parametric multiplicative error time series model. This class of estimators is based on the integrals of the square of a certain marked residual process. The paper derives the asymptotic distributions of the proposed estimators. In a finite sample comparison, some members of the proposed class of estimators dominate a generalized method of moments estimator in terms of the finite sample bias at a variety of chosen error distributions while neither dominate each other in terms of the mean squared error at these error distributions. A real data example is considered to illustrate the proposed estimation procedures.

Keywords: Marked empirical process, GMM estimator.

JEL Classifications: C13, C51.

1 Introduction

The analysis of financial time series often requires modeling of sequences of non-negative random variables. Examples include financial durations, absolute values of returns, trade volumes, and realized volatility. Due to the nature of high frequency financial data one often encounters zeros or near zeros in such non-negative random variables with non-trivial probability. Therefore, the usual practice of analyzing the log-transformed data is often not applicable. Furthermore, there are several drawbacks to taking logs when observations very near zero are possible, see, e.g. Engle (2002). Multiplicative error models (MEMs) of Engle (2002) decompose the observable process into its conditional mean and a non-negative valued multiplicative error and hence avoid the need to rely on log-transforms. These models have been found to be useful for modeling a variety of non-negative time series. Numerous applications and properties of these models are discussed in Engle and Russell (1998), Bauwens and Giot (2000), Bauwens and Veredas (2004), Manganelli (2005), Chou (2005), Engle and Gallo (2006), Fernandes and Grammig (2006), Gao, Kim and Saart (2015) and Koul and Perera (2021), among others.

The problem of estimating MEMs with parametric conditional means has been addressed by several authors. The quasi maximum likelihood estimator (QMLE) based on the standard exponential distribution was investigated by Engle and Russell (1998), Engle (2002), and Drost and Werker (2004). Cipollini, Engle and Gallo (2013) developed a generalized method of moments (GMM) estimator when the error distribution is unknown. The QMLE is consistent and asymptotically normal, provided that the innovation has unit mean and a finite second moment, even if the true distribution is not exponential (see Engle, 2002). The GMM estimator is more suited to the vector specification of the MEM and allows practitioners to estimate the assumed model without specifying an innovation distribution. Some of the key features of QML and GMM estimators have been discussed in Pacurar (2008), Hautsch (2012), Brownlees, Cipollini and Gallo (2012), Cipollini, Engle and Gallo (2013) and Perera, Hidalgo and Silvapulle (2016), among others. Although these estimators have desirable asymptotic properties, their finite sample performance can at times be sensitive to the (unknown) conditional distribution of the observable process, and hence, in practice, when available, fully efficient maximum likelihood (ML) estimates are often preferred (see Grammig and Maurer 2000, Perera and Silvapulle 2021). For example, even though the asymptotic distribution of the QMLE is independent of the innovation distribution, when the data generating process is based on an innovation distribution that induces a non-monotonic hazard rate function (e.g. Burr distribution), the QMLE based on the standard exponential distribution may perform poorly, even with quite large sample sizes (see Grammig and Maurer 2000). It is thus desirable to develop new estimation methods for MEMs that improve upon the existing methods. The objective of this paper is to contribute to this line of research by developing a new class of minimum distance (m.d.) estimators in a parametric Markovian MEM, that are applicable when fully efficient ML estimates are not available.

More precisely, let p, q be known positive integers, $\mathbb{R}_+^p := [0, \infty)^p$, $\mathbb{R}^q := (-\infty, \infty)^q$ and let $\psi(z, \vartheta)$, $z \in \mathbb{R}_+^p, \vartheta \in \Theta \subset \mathbb{R}^q$ be a family of positive functions. In the MEM of interest here we observe positive stationary time series Y_i, $i \in \mathbb{Z} := \{0, \pm 1, \cdots\}$ and a known p-vector of nonnegative functions Z_{i-1} of $Y_{i-1}, \cdots, Y_{i-\ell}$, for a known $\ell \geq 1$ such that for some $\theta \in \Theta$,

$$Y_i = \psi(Z_{i-1}, \theta)\varepsilon_i, \quad i \in \mathbb{Z}, \tag{1.1}$$

where $\varepsilon_i, i \in \mathbb{Z}$ are independent and identically distributed (i.i.d.) non-negative random variables (r.v.'s), $E(\varepsilon_0) = 1$, $\mathrm{Var}(\varepsilon_0) = \sigma^2 > 0$ and for each $i \in \mathbb{Z}$, ε_i is independent of $Y_{i-1}, \cdots, Y_{i-\ell}$. It is implicitly assumed that the parameter space Θ is such that when $\theta \in \Theta$ is true, the above time series $Y_i, i \in \mathbb{Z}$ is stationary. An example of this model, with $p = 1$ and Z_{i-1}'s different from Y_{i-1}'s, is provided in Example 2 of Perera and Koul (2017), where the time series $\{Y_i\}$ consists of daily annualized realized volatility measures, constructed from some intraday spot price data for the S&P500 index. It was seen in a certain data set that the changes of Y_i seem to be influenced by the fluctuations of past realized volatilities, which suggested to take $Z_{i-1} = \sum_{t=1}^{\ell} w_t Y_{i-t}$, a weighted sum of past realized volatilities with known $\{w_i\}$ and ℓ. Several other applications of this family of models are discussed in Koul, Perera and Silvapulle (2012) and Guo and Li (2018).

In the current paper we propose a class of m.d. estimators of θ in the above semi-parametirc model (1.1) based on the data $(Z_{i-1}, Y_i), 1 \leq i \leq n$ where distributions of Z_0 and ε_0 are unknown. To proceed further, for any real numbers a, b, $a \vee b := \max(a, b)$. For any two vectors $x = (x_1, \cdots, x_p)', y = (y_1, \cdots, y_p)' \in \mathbb{R}_+^p$, $x'(\|x\|)$ denote transpose (Euclidean norm) of x, $\{x \leq y\} := \{x_j \leq y_j, 1 \leq j \leq p\}$ and $x \vee y := \big(x_1 \vee y_1, \cdots, x_p \vee y_p\big)'$.

The m.d. estimators of interest here are the ones that are based on the marked residual empirical process

$$U_n(z, \vartheta) := n^{-1/2} \sum_{i=1}^{n} \Big(\frac{Y_i}{\psi(Z_{i-1}, \vartheta)} - 1\Big) I(Z_{i-1} \leq z), \quad z \in \mathbb{R}_+^p, \ \vartheta \in \Theta.$$

Using the ideas from Stute (1997) and Koul and Stute (1999), Koul et al. (2012) provide some motivation for basing inference in the above MEM models on an analog of the process $U_n(z, \vartheta)$, $z \in \mathbb{R}_+^p$, $\vartheta \in \Theta$. In additive models several m.d. estimators based on integrated squared residual empirical processes are known to have robustness and efficiency properties, see, e.g., Donoho and Liu (1988a,b) and Koul (1985a,b, 1986, 1996). For these reasons it is desirable to develop the m.d. estimators based on similar statistics involving the process U_n.

Now, let L be a distribution function (d.f.) on $\mathbb{R}_+^p$, $G_n(z)$ denote the proportion of

$Z_{i-1}, 1 \le i \le n$ less than or equal to z, i.e., $G_n(z) := n^{-1} \sum_{i=1}^n I(Z_{i-1} \le z)$ and define

$$\begin{aligned} M_{n1}(\vartheta) &:= \int U_n^2(z,\vartheta) dL(z), \qquad \widehat{\theta}_{n1} := \text{argmin}_\vartheta M_{n1}(\vartheta), \\ M_{n2}(\vartheta) &:= \int U_n^2(z,\vartheta) dG_n(z), \qquad \widehat{\theta}_{n2} := \text{argmin}_\vartheta M_{n2}(\vartheta). \end{aligned} \tag{1.2}$$

A further motivation for the above estimators is as follows. Note that $E\big\{\big(Y_i/\psi(Z_{i-1},\theta) -1\big) I(Z_{i-1} \le z)\big\} = 0$, for all $i \in \mathbb{Z}, z \in \mathbb{R}_+^p$. Thus intuitively, one may estimate θ by that entity ϑ that brings the set of variables $Y_i I(Z_{i-1} \le z)/\psi(Z_{i-1},\vartheta), 1 \le i \le n, z \in \mathbb{R}_+^p$ close to the set of variables $I(Z_{i-1} \le z), 1 \le i \le n, z \in \mathbb{R}_+^p$, in some sense. The estimators $\widehat{\theta}_{nj}, j = 1,2$ bring these two data sets close to each other in terms of the integrated square differences $n^{-1}M_{nj}(\vartheta), j = 1,2$ between their averages. Note that $\widehat{\theta}_{n1}$ provides a class of estimators, one for each L.

In some applications Z_{i-1} is one dimensional. For example the lack-of-fit test proposed by Koul et al. (2012), indicates that a Markovian MEM of the form (1.1) provides a good fit for a time series of squared log returns of a stock, denoted $\{Y_i\}$, with $\psi(Z_{i-1};\theta) = \theta_1 + \theta_2 Z_{i-1}$, $Z_{i-1} = Y_{i-1}$, where $\theta_1 > 0,\ 0 \le \theta_2 < 1,\ \theta = (\theta_1, \theta_2)'$. In such cases we give a representation of $M_{nj}(\vartheta),\ j = 1,2$ that is useful for their computation and that of $\widehat{\theta}_{nj}, j = 1,2$. Since $Z_{i-1}, 1 \le i \le n$ are one-dimensional, one orders $Z_{i-1}, 1 \le i \le n$ as $Z_{(0)} \le Z_{(1)} \le \cdots \le Z_{(n-1)}$. Let Y_i^* denote the Y_i corresponding to $Z_{(i-1)}, 1 \le i \le n$. Direct computation shows that

$$\begin{aligned} M_{n1}(\vartheta) &= n^{-1} \sum_{i=1}^n \sum_{j=1}^n \Big(\frac{Y_i}{\psi(Z_{i-1},\vartheta)} - 1\Big)\Big(\frac{Y_j}{\psi(Z_{j-1},\vartheta)} - 1\Big) \int I(Z_{i-1} \le z, Z_{j-1} \le z) dL(z) \\ &= n^{-1} \sum_{i=1}^n \big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)^2 \big(1 - L\big(Z_{(i-1)} - \big)\big) \\ &\quad + 2n^{-1} \sum_{i=1}^n \sum_{1 \le i < j \le n} \big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)\big(\frac{Y_j^*}{\psi(Z_{(j-1)},\vartheta)} - 1\big) \big[1 - L\big(Z_{(j-1)} - \big)\big]. \end{aligned}$$

Similarly, using the fact $G_n\big(Z_{(i-1)} - \big) = (i-1)/n$, for all $1 \le i \le n$, we obtain

$$\begin{aligned} M_{n2}(\vartheta) &= n^{-1} \sum_{i=1}^n \big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)^2 [1 - G_n\big(Z_{(i-1)} - \big)] \\ &\quad + 2n^{-1} \sum_{i=1}^n \sum_{1 \le i < j \le n} \big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)\big(\frac{Y_j^*}{\psi(Z_{(j-1)},\vartheta)} - 1\big) \big[1 - G_n\big(Z_{(j-1)} - \big)\big] \\ &= n^{-2} \sum_{i=1}^n \big(n - i + 1\big)\big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)^2 \\ &\quad + 2n^{-2} \sum_{i=1}^n \sum_{j=i+1}^n \big(n - j + 1\big)\big(\frac{Y_i^*}{\psi(Z_{(i-1)},\vartheta)} - 1\big)\big(\frac{Y_j^*}{\psi(Z_{(j-1)},\vartheta)} - 1\big). \end{aligned}$$

In the computation of $M_{n1}(\vartheta)$, for the choice of L, one may consider any d.f. with positive support; for example, the standard exponential distribution or a member of the standard gamma family (with two parameters). We discuss some such choices of L in (2.2) below in the next section. Note that, once a suitable choice of L is made, one can easily compute $\widehat{\theta}_{n1}$ and $\widehat{\theta}_{n2}$ by using the computational formulas described above for $M_{n1}(\vartheta)$ and $M_{n2}(\vartheta)$.

Our goal here is to derive the asymptotic distributions of the estimators $\widehat{\theta}_{n1}$ and $\widehat{\theta}_{n2}$ and to study their finite sample bias and mean square error (MSE), in comparison with some competing estimators.

The next section reports the findings of a finite sample simulation study, where we find that the empirical bias of the several m.d. estimators is much less than that of the GMM estimator (GMME) at all chosen error distributions and some m.d. estimators compete very well in terms of the empirical MSE with the GMME. In Section 3 we illustrate the proposed estimation procedures in an empirical example. The asymptotic normality of these estimators along with the needed regularity conditions are presented in Section 4. Remark 4.1 below describes a limited comparison of the asymptotic variance of $\widehat{\theta}_{n1}$ and $\widehat{\theta}_{n2}$ with that of the GMME. Section 5 concludes the paper and some proofs are relegated to an Appendix A.

2 Simulation study

In this section we present the findings of a Monte Carlo simulation study that evaluates the finite sample performance, in terms of bias and MSE, of some members of the above proposed class of m.d. estimators $\widehat{\theta}_{nj}, j = 1, 2$, and GMME $\widehat{\theta}_{mm}$.

2.1 Design of the simulation study

In this simulation study, we use $p = 1$, $q = 2$, and

$$\psi(Z_{i-1};\theta) = \theta_1 + \theta_2 Z_{i-1}, \quad \theta_1 > 0, \quad 0 \le \theta_2 < 1, \qquad \theta = (\theta_1, \theta_2)'. \tag{2.1}$$

This process is stationary for $\theta \in \Theta := (0, \infty) \times [0, 1)$. Similar specifications have previously been considered for the conditional mean in Koul et al. (2012) and Guo and Li (2018) in the MEM setting. Also note that the $\psi(z, \vartheta) = \vartheta_1 + \vartheta_2 z$ satisfies the conditions (C.1), (C.2) and (4.8) in Section 4 below, and hence (C.3) and (C.4) in Section 4 also hold (see Lemma 4.1).

For Z_{i-1} we consider the following four cases:

$$(a):\ Z_{i-1} = Y_{i-1}, \quad (b):\ Z_{i-1} = \Big(\prod_{k=1}^{3} Y_{i-k}\Big)^{1/3}, \quad (c):\ Z_{i-1} = \Big(\sum_{k=1}^{3} Y_{i-k}\Big)\Big/3,$$
$$(d):\ Z_{i-1} = 3\Big\{\sum_{k=1}^{3}\Big(1/Y_{i-k}\Big)\Big\}^{-1}, \quad i = 1, \cdots, n.$$

The first case ($Z_{i-1} = Y_{i-1}$) corresponds to the Markovian model considered in Koul et al. (2012). A setup similar to (c) was considered in Perera and Koul (2017) for a closely related threshold model. In (b) and (d) we consider respectively the geometric and harmonic means of the three lagged values of Y_i.

For the error distribution, we consider the following families of densities on $(0, \infty)$.

1. **Exponential** [E]: $f(x) := e^{-x}, x > 0$.
2. **Gamma** [G]: $f(x) := \alpha^\alpha \Gamma(\alpha)^{-1} x^{\alpha-1} e^{-\alpha x}$, $x > 0, \alpha > 0$.
3. **Generalized gamma** [GG]:
   $f(x) = c\{\sigma\Gamma(a)\}^{-1}(x/\sigma)^{ac-1}\exp\{-(x/\sigma)^c\}$, $a, c > 0$, $\sigma = \{\Gamma(a + c^{-1})\}^{-1}\Gamma(a)$.
4. **Burr**: $f(x) = (a/\sigma)(x/\sigma)^{a-1}[1 + b(x/\sigma)^a]^{-(1+b^{-1})}$, $a > b > 0$, and
   $\sigma = \{\Gamma(1 + a^{-1})\Gamma(b^{-1} - a^{-1})\}^{-1} b^{(1+a^{-1})}\Gamma(1 + b^{-1})$.

For the exponential distribution, $E\varepsilon_0 = 1$, while for each of the other three distributions, the scale parameter is selected to have $E(\varepsilon_0) = 1$. For the sample sizes we consider $n = 100, 200$ and 500.

The first two distributions above have been identified as having important roles in multiplicative error models, see, Engle and Russell (1998), Drost and Werker (2004) and Engle and Gallo (2006). The next two have been suggested in various empirical studies, see, e.g., Lunde (1999), Grammig and Maurer (2000) and Grammig and Wellner (2002).

In the tables below we write $\widehat{\theta}_1 = (\widehat{\theta}_{11}, \widehat{\theta}_{12})$, $\widehat{\theta}_2 = (\widehat{\theta}_{21}, \widehat{\theta}_{22})$ for $\widehat{\theta}_{n1}, \widehat{\theta}_{n2}$, respectively, and call them MDEs, while $\widehat{\theta}_{mm,1}, \widehat{\theta}_{mm,2}$ denote the GMMEs of θ_1, θ_2, respectively. We evaluate the performance of $\widehat{\theta}_1$ corresponding to each of the following four choices of L.

$$
\begin{aligned}
L_1(y) &= 1 - \exp(-y/2), \quad L_2(y) = 1 - \exp(-y), \\
L_3(y) &= 1 - \exp(-2y), \quad L_4(y) = 1 - \exp(-4y), \quad y > 0.
\end{aligned} \tag{2.2}
$$

MEMs require the conditional mean to be always positive. We ensure this by imposing a positive intercept θ_1 and a non-negative coefficient θ_2 in the conditional mean equation (2.1). Further, the coefficient θ_2 is restricted to be less than 1 to ensure stationarity. Therefore, under the specification (2.1), the parameter space in the minimization problem (1.2) is inequality restricted. To implement the constrained minimization problem in (1.2), under the specification (2.1), we use the `fmincon` function in Matlab.

2.2 Summary of the results

The simulation findings about empirical bias are given in Tables 1, 3, 5 and 7, and those on MSE are given in Tables 2, 4, 6 and 8. The main findings of the simulations are as follows. All the chosen m.d. estimators perform considerably better than the GMME in terms of the empirical bias while being competitive in terms of the empirical MSE at all the chosen error distributions and the sample sizes. Recall that the GMME is equivalent to the MLE when the error term follows a Gamma distribution yet we find the empirical MSE of some of the chosen m.d. estimators at the Gamma error distribution to be very close to that of the GMME.

Table 1: Empirical bias of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.4, 0.6)'$, $Z_{i-1} \equiv Y_{i-1}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.015	.008	.006	.005	.005	.005	-.038	-.015	-.008	.002	.006	-.003
G	.020	.014	.013	.013	.014	.011	-.030	-.017	-.014	-.013	-.014	-.010
GG	.017	.007	.005	.004	.004	.004	-.039	-.006	.003	.011	.012	.009
Burr	.026	.017	.018	.020	.023	.017	-.042	-.016	-.017	-.021	-.028	-.010
			$n = 200$						$n = 200$			
E	.006	.001	.000	.000	.001	.000	-.016	-.002	.002	.004	.006	.004
G	.010	.008	.006	.005	.005	.005	-.014	-.009	-.005	-.001	.001	-.002
GG	.008	.003	.002	.002	.004	.002	-.012	.007	.009	.011	.011	.010
Burr	.013	.007	.006	.005	.006	.005	-.020	-.003	.003	.009	.007	.006
			$n = 500$						$n = 500$			
E	.003	.001	.000	.000	.000	.000	-.003	.003	.005	.006	.009	.005
G	.004	.002	.002	.003	.003	.002	-.005	-.002	-.002	-.002	-.004	-.001
GG	.005	.002	.002	.002	.002	.002	-.009	.002	-.001	.001	.003	.000
Burr	.003	.001	.000	.000	.000	.000	-.009	-.001	.001	.004	.006	.002

Table 2: Empirical MSE of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.4, 0.6)'$, $Z_{i-1} \equiv Y_{i-1}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.0084	.0084	.0086	.0092	.0102	.0084	.0261	.0285	.0338	.0450	.0680	.0319
G	.0076	.0078	.0081	.0090	.0109	.0082	.0155	.0167	.0191	.0255	.0437	.0197
GG	.0103	.0098	.0100	.0106	.0118	.0099	.0364	.0416	.0469	.0599	.0884	.0448
Burr	.0133	.0135	.0145	.0161	.0182	.0143	.0374	.0438	.0497	.0635	.0921	.0487
			$n = 200$						$n = 200$			
E	.0038	.0039	.0040	.0042	.0050	.0039	.0122	.0147	.0177	.0242	.0426	.0167
G	.0035	.0036	.0038	.0044	.0060	.0038	.0072	.0078	.0092	.0135	.0271	.0093
GG	.0054	.0055	.0057	.0062	.0071	.0056	.0188	.0229	.0277	.0388	.0629	.0261
Burr	.0067	.0073	.0076	.0083	.0096	.0075	.0172	.0223	.0272	.0394	.0635	.0261
			$n = 500$						$n = 500$			
E	.0014	.0015	.0016	.0017	.0021	.0015	.0048	.0058	.0069	.0098	.0195	.0065
G	.0013	.0013	.0014	.0016	.0022	.0015	.0026	.0029	.0035	.0053	.0104	.0037
GG	.0019	.0021	.0022	.0024	.0028	.0021	.0074	.0090	.0111	.0162	.0304	.0104
Burr	.0024	.0026	.0027	.0031	.0037	.0027	.0082	.0099	.0123	.0181	.0320	.0121

Table 3: Empirical bias of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv (\prod_{k=1}^{3} Y_{i-k})^{1/3}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.018	.013	.011	.009	.007	.007	-.076	-.054	-.046	-.035	-.021	-.016
G	.026	.021	.019	.018	.016	.014	-.067	-.052	-.047	-.041	-.035	-.027
GG	.019	.012	.010	.007	.005	.004	-.079	-.047	-.034	-.018	-.003	.009
Burr	.027	.020	.018	.016	.016	.014	-.084	-.057	-.047	-.038	-.033	-.021
			$n = 200$						$n = 200$			
E	.009	.005	.005	.004	.003	.002	-.039	-.024	-.020	-.014	-.009	-.005
G	.015	.012	.011	.010	.010	.009	-.036	-.028	-.026	-.023	-.021	-.016
GG	.012	.007	.006	.005	.006	.004	-.038	-.017	-.012	-.008	-.006	.002
Burr	.015	.011	.010	.009	.008	.007	-.046	-.029	-.024	-.018	-.012	-.006
			$n = 500$						$n = 500$			
E	.004	.002	.002	.001	.001	.001	-.010	-.003	-.002	.000	.001	.003
G	.006	.004	.004	.004	.004	.003	-.013	-.009	-.008	-.007	-.007	-.004
GG	.005	.003	.002	.002	.003	.002	-.019	-.008	-.007	-.006	-.006	-.003
Burr	.005	.002	.002	.002	.002	.001	-.019	-.009	-.007	-.006	-.005	-.002

Table 4: Empirical MSE of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv (\prod_{k=1}^{3} Y_{i-k})^{1/3}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.0061	.0059	.0059	.0060	.0064	.0063	.0602	.0582	.0593	.0636	.0783	.0726
G	.0069	.0067	.0067	.0068	.0074	.0072	.0395	.0387	.0392	.0414	.0496	.0470
GG	.0066	.0064	.0064	.0066	.0071	.0070	.0738	.0741	.0766	.0831	.0967	.0928
Burr	.0088	.0086	.0086	.0089	.0095	.0093	.0772	.0798	.0828	.0881	.0998	.0964
			$n = 200$						$n = 200$			
E	.0027	.0026	.0026	.0027	.0030	.0029	.0291	.0285	.0292	.0321	.0422	.0385
G	.0031	.0030	.0030	.0032	.0036	.0034	.0174	.0170	.0174	.0190	.0248	.0229
GG	.0038	.0037	.0038	.0040	.0044	.0043	.0404	.0417	.0439	.0490	.0613	.0578
Burr	.0048	.0049	.0050	.0052	.0056	.0054	.0393	.0418	.0437	.0488	.0605	.0573
			$n = 500$						$n = 500$			
E	.0010	.0010	.0011	.0011	.0013	.0013	.0102	.0107	.0114	.0131	.0176	.0160
G	.0011	.0011	.0011	.0012	.0014	.0014	.0060	.0061	.0063	.0071	.0095	.0088
GG	.0014	.0014	.0014	.0015	.0017	.0017	.0161	.0166	.0175	.0204	.0278	.0260
Burr	.0018	.0019	.0019	.0021	.0024	.0023	.0178	.0188	.0197	.0225	.0303	.0285

Table 5: Empirical bias of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv (\sum_{k=1}^{3} Y_{i-k})/3$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.033	.025	.023	.021	.021	.019	-.075	-.055	-.048	-.042	-.037	-.032
G	.035	.028	.026	.025	.025	.022	-.066	-.051	-.047	-.043	-.043	-.034
GG	.031	.021	.018	.016	.016	.014	-.070	-.042	-.032	-.024	-.015	-.012
Burr	.038	.027	.024	.021	.022	.020	-.073	-.045	-.035	-.026	-.023	-.019
			$n = 200$						$n = 200$			
E	.016	.011	.010	.009	.010	.008	-.038	-.026	-.022	-.019	-.019	-.014
G	.020	.016	.016	.015	.015	.013	-.037	-.029	-.027	-.025	-.026	-.020
GG	.016	.010	.009	.009	.009	.008	-.030	-.014	-.009	-.006	-.004	-.001
Burr	.020	.014	.012	.011	.011	.010	-.037	-.021	-.015	-.010	-.007	-.005
			$n = 500$						$n = 500$			
E	.006	.004	.004	.004	.005	.003	-.011	-.005	-.004	-.004	-.005	-.001
G	.007	.006	.005	.005	.006	.005	-.013	-.009	-.009	-.009	-.010	-.007
GG	.008	.005	.005	.005	.006	.005	-.016	-.009	-.008	-.008	-.009	-.006
Burr	.006	.003	.003	.002	.003	.002	-.016	-.007	-.005	-.004	-.004	-.002

Table 6: Empirical MSE of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv (\sum_{k=1}^{3} Y_{i-k})/3$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.0109	.0103	.0104	.0111	.0129	.0111	.0421	.0408	.0424	.0484	.0655	.0482
G	.0096	.0092	.0092	.0095	.0110	.0100	.0328	.0317	.0326	.0359	.0470	.0395
GG	.0108	.0105	.0107	.0116	.0134	.0113	.0467	.0506	.0535	.0616	.0809	.0598
Burr	.0133	.0129	.0129	.0139	.0158	.0138	.0495	.0544	.0593	.0694	.0874	.0683
			$n = 200$						$n = 200$			
E	.0045	.0044	.0044	.0046	.0055	.0047	.0191	.0188	.0197	.0235	.0345	.0237
G	.0043	.0042	.0042	.0045	.0055	.0049	.0143	.0142	.0146	.0165	.0229	.0187
GG	.0054	.0054	.0057	.0065	.0079	.0063	.0229	.0246	.0279	.0359	.0524	.0337
Burr	.0065	.0069	.0071	.0078	.0093	.0078	.0243	.0281	.0313	.0388	.0549	.0379
			$n = 500$						$n = 500$			
E	.0016	.0016	.0017	.0019	.0024	.0019	.0067	.0071	.0076	.0092	.0138	.0093
G	.0015	.0015	.0016	.0017	.0022	.0019	.0049	.0051	.0054	.0062	.0090	.0072
GG	.0019	.0020	.0021	.0023	.0028	.0023	.0091	.0098	.0110	.0139	.0215	.0137
Burr	.0024	.0026	.0028	.0031	.0038	.0031	.0105	.0119	.0132	.0160	.0256	.0168

Table 7: Empirical bias of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv 3\{\sum_{k=1}^{3}(1/Y_{i-k})\}^{-1}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.009	.006	.004	.003	.002	.003	-.067	-.042	-.034	-.021	-.005	-.004
G	.018	.014	.013	.011	.009	.008	-.059	-.045	-.040	-.032	-.022	-.016
GG	.010	.005	.004	.002	.000	.001	-.067	-.033	-.022	-.006	.012	.022
Burr	.017	.012	.011	.010	.009	.010	-.075	-.048	-.040	-.030	-.023	-.016
			$n = 200$						$n = 200$			
E	.004	.002	.001	.001	.000	.000	-.033	-.018	-.014	-.007	.001	.002
G	.010	.008	.007	.006	.005	.005	-.031	-.024	-.021	-.018	-.013	-.011
GG	.006	.003	.002	.002	.002	.008	-.032	-.008	-.003	.003	.007	.011
Burr	.010	.006	.006	.005	.005	.005	-.043	-.025	-.021	-.016	-.011	-.005
			$n = 500$						$n = 500$			
E	.002	.001	.001	.000	.000	.000	-.008	.000	.002	.004	.007	.008
G	.004	.003	.003	.002	.002	.002	-.011	-.008	-.007	-.006	-.004	-.002
GG	.003	.001	.001	.001	.001	.001	-.018	-.005	-.004	-.002	-.001	.001
Burr	.002	.001	.001	.000	.000	.000	-.017	-.007	-.006	-.004	-.002	.000

Table 8: Empirical MSE of the GMME and the MDEs of θ for the DGP (2.1) with $\theta = (0.3, 0.5)'$, $Z_{i-1} \equiv 3\{\sum_{k=1}^{3}(1/Y_{i-k})\}^{-1}$ and the error d.f. F_0.

F_0	$\widehat{\theta}_{mm,1}$	$\widehat{\theta}_{11}$				$\widehat{\theta}_{21}$	$\widehat{\theta}_{mm,2}$	$\widehat{\theta}_{12}$				$\widehat{\theta}_{22}$
		L_1	L_2	L_3	L_4			L_1	L_2	L_3	L_4	
			$n = 100$						$n = 100$			
E	.0037	.0036	.0036	.0036	.0038	.0038	.0695	.0680	.0692	.0732	.0850	.0867
G	.0046	.0045	.0045	.0045	.0048	.0048	.0414	.0407	.0410	.0427	.0499	.0502
GG	.0043	.0041	.0042	.0042	.0044	.0046	.0923	.0925	.0952	.1004	.1115	.1157
Burr	.0057	.0055	.0055	.0056	.0059	.0060	.0911	.0912	.0924	.0966	.1071	.1102
			$n = 200$						$n = 200$			
E	.0017	.0017	.0017	.0017	.0018	.0018	.0035	.0035	.0036	.0038	.0472	.0486
G	.0021	.0021	.0021	.0021	.0023	.0023	.0185	.0183	.0185	.0197	.0244	.0246
GG	.0024	.0024	.0024	.0025	.0027	.0027	.0539	.0552	.0572	.0620	.0727	.0771
Burr	.0032	.0032	.0033	.0034	.0036	.0037	.0488	.0509	.0523	.0564	.0674	.0709
			$n = 500$						$n = 500$			
E	.0007	.0007	.0007	.0007	.0008	.0008	.0125	.0129	.0135	.0152	.0198	.0208
G	.0008	.0008	.0008	.0008	.0009	.0009	.0064	.0066	.0068	.0074	.0095	.0097
GG	.0010	.0010	.0010	.0010	.0011	.0011	.0228	.0233	.0241	.0267	.0338	.0377
Burr	.0013	.0013	.0013	.0013	.0015	.0015	.0232	.0241	.0247	.0269	.0329	.0359

3 Empirical example

In this section, we illustrate the proposed m.d. estimation procedures by considering a real data example. The time series of interest $\{Y_i\}$ is the monthly squared log returns of Intel stocks. The data for this example were downloaded from the home page of Professor Ruey S. Tsay. The data spans the period January 1973 to December 2009. Koul et al. (2012) studied a shorter version of this data set, and concluded that a Markov MEM of the form (1.1) with $\psi(Z_{i-1};\theta) = \theta_1 + \theta_2 Y_{i-1}$, where $\theta_1 > 0,\ 0 \le \theta_2 < 1,\ \theta = (\theta_1, \theta_2)'$, provides a good fit. Therefore, in this empirical illustration we consider the following Markov MEM:

$$Y_i = \psi(Z_{i-1};\theta)\varepsilon_i, \quad \psi(Z_{i-1};\theta) = \theta_1 + \theta_2 Y_{i-1}, \quad \theta_1 > 0,\ 0 \le \theta_2 < 1,\ \theta = (\theta_1, \theta_2)'. \tag{3.1}$$

Model diagnostics

First, we apply several model diagnostic tests to check the adequacy of model (3.1) for the data. More specifically, we apply (a) the lack-of-fit test developed by Koul et al. (2012), denoted KPS, (b) the LjungBox Q test (Ljung and Box, 1978), (c) a LM test (Meitz and Teräsvirta, 2006), and (d) a generalized moment test (Chen and Hsieh, 2010).

The large sample level-α critical value c_α of the KPS test is equal to the $100(1-\alpha)\%$ quantile of $\sup_{0\le t\le 1}|W_0(t)|$ where W_0 is the standard Brownian motion (Koul et al., 2012).

For $\alpha = 0.01, 0.05$ and 0.10, these critical values are 2.807034, 2.241403 and 1.959964, respectively. For the LM test we use the form in Theorem 1 of Meitz and Teräsvirta (2006), with ACD$(2, 0)$ being the model under the maintained hypothesis. Hence, the asymptotic null distribution of the resulting LM statistic is χ^2_1 (Meitz and Teräsvirta, 2006). We compute the Ljung-Box Q statistic as $\text{LBQ}(\ell_0) = T(T+2)\sum_{k=1}^{\ell_0}(T-k)^{-1}\rho_k^2$, where ℓ_0 is the number of autocorrelation lags included in the statistic, and ρ_k^2 is the squared sample autocorrelation at lag k of the estimated residuals. The critical values to implement the LBQ(ℓ_0) are obtained from $\chi^2_{\ell_0}$ (see Pacurar, 2008). The generalized moment test is the M-test of Chen and Hsieh (2010). To compute the M-test statistic, as in Koul et al. (2012), we use $(\varepsilon_{i-1} - 1)$ as the "misspecification indicator" g_i (see page 354 in Chen and Hsieh, 2010). The asymptotic null distribution of the resulting M test statistic is χ^2_1 (see page 353 in Chen and Hsieh, 2010).

The observed value of the KPS test statistic turns out to be 1.2610. Since this value is smaller than the 10% level critical value 1.959964, the KPS test does not reject the Markov MEM in (3.1) at 10% level. The observed values of LBQ(ℓ_0), for $\ell_0 = 1, 5$ and 15, LM and M test statistics, together with their p-values (reported within parentheses) are as follows:

$$\begin{aligned} &\text{LBQ}(1) = 0.7963\ (0.3722), \quad \text{LBQ}(5) = 7.0291\ (0.2185), \quad \text{LBQ}(10) = 13.6782\ (0.1882), \\ &\text{LBQ}(15) = 31.9087\ (0.0067), \quad \text{LM} = 0.8946\ (0.3442), \quad M = 3.4992\ (0.0614). \end{aligned}$$

Therefore, apart from LBQ(15), all the other tests fail to reject the null Markov MEM in (3.1) at 5% level. Hence, we fit the model (3.1) to the observed data set.

Estimation

We estimate the parameter θ of the Markov MEM in (3.1) by using the proposed m.d. estimators $\widehat{\theta}_{nj}, j = 1, 2$, in (1.2). For comparison, we also include the GMME of Cipollini et al. (2013). In the model (1.1), the GMM estimator is obtained by solving the following sample moment conditions for ϑ:

$$n^{-1}\sum_{i=1}^{n}\dot{\psi}(Z_{i-1},\vartheta)\Big\{\frac{Y_i - \psi(Z_{i-1},\vartheta)}{\psi^2(Z_{i-1},\vartheta)}\Big\} = 0.$$

These conditions coincide with the first-order conditions of a log-likelihood function maximization in the model (1.1) when ε_i are i.i.d. Gamma distributed r.v.'s, see Engle and Gallo (2006). The above GMME, denoted by $\widehat{\theta}_{mm}$ herein, provides a consistent estimator of θ under the correct specification of the conditional mean, without relying on any parametric assumptions on the error distribution. Moreover, under some regularity conditions, $n^{1/2}(\widehat{\theta}_{mm} - \theta) \to_D N(0, \sigma^2\Sigma^{-1})$, $\Sigma = E[\varphi(Z_0)\varphi(Z_0)']$.

To compute the m.d. estimator $\widehat{\theta}_1$ we use each of the four choices of L in (2.2). The estimated Markov multiplicative error models for the selected methods are as follows:

$$\begin{aligned}
\text{For } \widehat{\theta}_{n1} \text{ with } L_1(y): &\quad \psi(Z_{i-1}; \widehat{\theta}_{n1}) = 0.01142 + 0.35452 Y_{i-1}, \\
\text{For } \widehat{\theta}_{n1} \text{ with } L_2(y): &\quad \psi(Z_{i-1}; \widehat{\theta}_{n1}) = 0.01139 + 0.35940 Y_{i-1}, \\
\text{For } \widehat{\theta}_{n1} \text{ with } L_3(y): &\quad \psi(Z_{i-1}; \widehat{\theta}_{n1}) = 0.01133 + 0.36803 Y_{i-1}, \\
\text{For } \widehat{\theta}_{n1} \text{ with } L_4(y): &\quad \psi(Z_{i-1}; \widehat{\theta}_{n1}) = 0.01125 + 0.38153 Y_{i-1}, \\
\text{For } \widehat{\theta}_{n2}: &\quad \psi(Z_{i-1}; \widehat{\theta}_{n2}) = 0.01062 + 0.51316 Y_{i-1}, \\
\text{For } \widehat{\theta}_{mm}: &\quad \psi(Z_{i-1}; \widehat{\theta}_{mm}) = 0.01137 + 0.35927 Y_{i-1}.
\end{aligned}$$

In view of these results, the observed values of $\widehat{\theta}_{n1}$ are very similar to those of the GMME $\widehat{\theta}_{mm}$, for all four choices of L, whereas the m.d. estimator $\widehat{\theta}_{n2}$, obtained by using the empirical d.f. G_n for L, produces slightly different estimated values for both θ_1 and θ_2. Figure 1 shows that the conditional mean processes for $\widehat{\theta}_{mm}$, $\widehat{\theta}_{n1}$ with $L_2(y)$, and $\widehat{\theta}_{n2}$ are all very similar to each other and closely match the observed monthly Intel squared log return series. The corresponding conditional mean processes produced by $\widehat{\theta}_{n1}$ with $L_1(y)$, $L_3(y)$, and $L_4(y)$ (not plotted here) are also very similar to those for $\widehat{\theta}_{mm}$ and $\widehat{\theta}_{n1}$ with $L_2(y)$.

4 Asymptotic distributions of $\widehat{\theta}_{nj}, j = 1, 2$

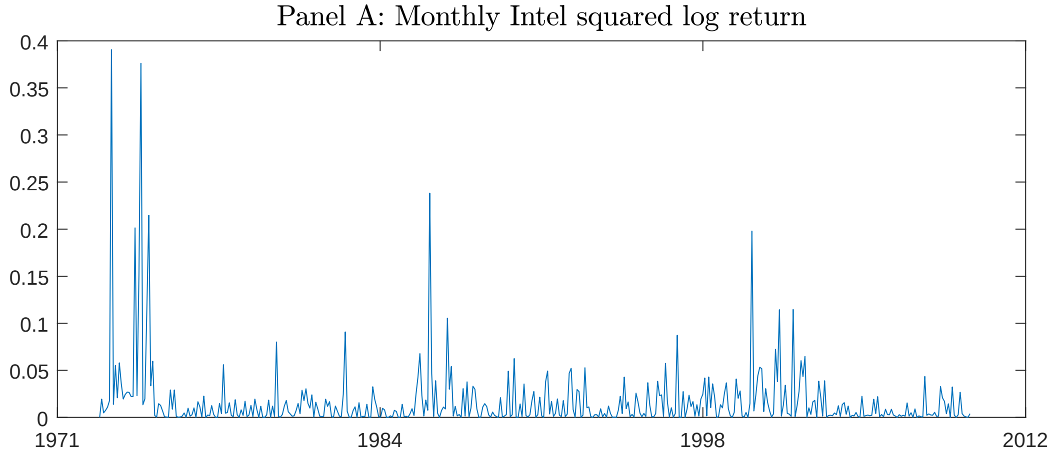
This section derives the asymptotic distributions of $\widehat{\theta}_{n1}$ and $\widehat{\theta}_{n2}$ under fairly general assumptions on the underlying entities. This is facilitated by applying the general method of Section 5.4 of Koul (2002) to the current setup. This method requires two steps. The first step requires to show that the $M_{nj}(\vartheta), j = 1, 2$ are AULQ (asymptotically uniformly locally quadratic) in $n^{1/2}(\vartheta - \theta)$ for $\vartheta \in \mathcal{N}_n(b) := \{\vartheta \in \Theta, n^{1/2}\|\vartheta - \theta\| \le b\}$, for every $0 < b < \infty$. The second step requires to show that $n^{1/2}\|\widehat{\theta}_{nj} - \theta\|, j = 1, 2$ are asymptotically bounded in probability. To begin with we focus on the class of estimator $\widehat{\theta}_{n1}$.

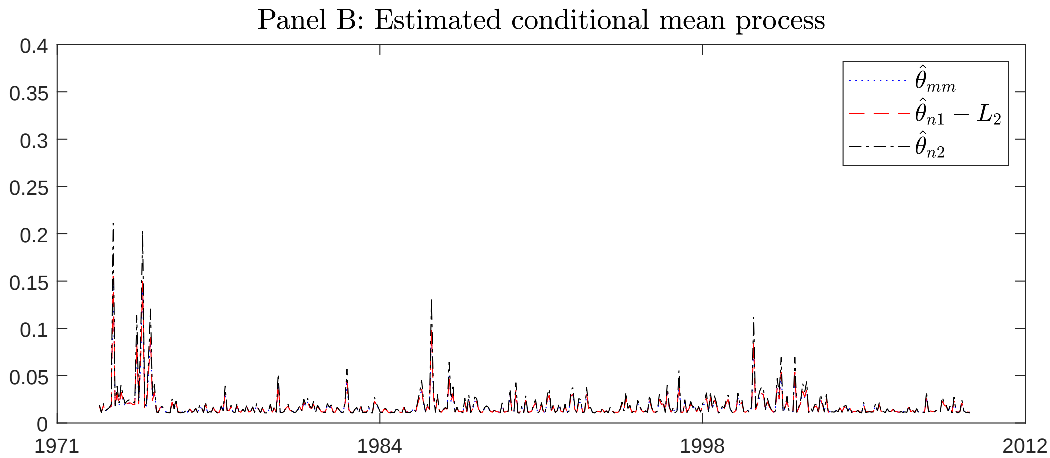
4.1 Asymptotic distribution of $\widehat{\theta}_{n1}$

Let $Y_{-\ell}, Y_{1-\ell}, \cdots, Y_n$ denote a given stretch of realizations of an observable stationary process $\{Y_i\}$ obeying (1.1). Let $T_{n1}(t) := M_{n1}(\theta + n^{-1/2}t)$, $t \in \mathbb{R}^q$. Proving that $M_{n1}(\vartheta), \vartheta \in \mathcal{N}_n(b)$ is AULQ is equivalent to proving that for every $0 < b < \infty$, $T_{n1}(t)$ is approximated by a quadratic form in t, uniformly in $\|t\| \le b$, in probability.

Let G denote the d.f. of Z_0. We shall write $U_n(z)$ for $U_n(z, \theta)$. Because ε_i are independent

Figure 1: Monthly squared log returns of Intel stocks (Panel A), and the corresponding conditional mean processes for the estimators $\widehat{\theta}_{mm}$, $\widehat{\theta}_{n1}$ with $L = L_2$, and $\widehat{\theta}_{n2}$ for fitting a Markov MEM (Panel B).





of Z_{i-1} and because $E\varepsilon_0 = 1$, we obtain

$$U_n(z) = n^{-1/2}\sum_{i=1}^{n}(\varepsilon_i - 1)I(Z_{i-1} \le z), \qquad T_{n1}(0) = \int U_n^2(z)dL(z),$$
$$EU_n(z) \equiv 0, \qquad EU_n^2(z) \equiv \sigma^2 G(z), \qquad ET_{n1}(0) = \sigma^2 \int G(z)dL(z) \le \sigma^2 < \infty.$$

Hence, by the Markov inequality,

$$M_{n1}(\theta) = T_{n1}(0) = O_p(1). \tag{4.1}$$

This fact is useful in showing that $\big\|n^{1/2}(\widehat{\theta}_{n1} - \theta)\big\| = O_p(1)$.

Before proceeding further, we shall state the additional needed assumptions as follows. In the sequel, the variables z, ϑ vary over $\mathbb{R}_+^p, \Theta$, respectively, unless specified otherwise.

(C.1) $\inf_{z,\vartheta} \psi(z,\vartheta) \ge C > 0$, for some $C < \infty$.

(C.2) There exists a q-vector $\dot{\psi}(z,\vartheta)$ such that $E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty$ and for every $0 < b < \infty$, $\sup_{1\le i\le n,\, \|t\|\le b} \sqrt{n}|\psi(Z_{i-1}, \theta + n^{-1/2}t) - \psi(Z_{i-1},\theta) - n^{-1/2}t'\dot{\psi}(Z_{i-1},\theta)| = o_p(1)$.

(C.3) $\forall\, \epsilon > 0$, $0 < \eta < \infty$, $\exists\, N_{\epsilon,\eta}$, $0 < b \equiv b_{\epsilon,\eta} < \infty$, such that $\forall\, n > N_{\epsilon,\eta}$,

$$P\Big(\inf_{\|t\|>b} M_{n1}(\theta + n^{-1/2}t) \ge \eta\Big) \ge 1 - \epsilon.$$

The assumptions (C.1) and (C.2) are used to derive the AULQ property of M_{nj}, $j = 1, 2$, while (4.1) and assumption (C.3) are used to show that $\|n^{1/2}(\widehat{\theta}_{n1} - \theta)\| = O_p(1)$.

Let $\varphi(z) := \dot{\psi}(z,\theta)/\psi(z,\theta)$. Assumptions (C.1) and (C.2) imply that

$$E\Big(\frac{\|\dot{\psi}(Z_0,\theta)\|^2}{\psi^j(Z_0,\vartheta)}\Big) \le C^{-j}E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty, \quad \forall\, j \ge 1, \tag{4.2}$$
$$E\|\varphi(Z_0)\|^2 < C^{-2}E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty.$$

To proceed further we need some more notation. Let

$$\Psi_n(z) := n^{-1}\sum_{i=1}^{n} \varphi(Z_{i-1})\, I(Z_{i-1} \le z), \qquad \Psi(z) := E\big(\varphi(Z_0)I(Z_0 \le z)\big),$$
$$V_{n1} := \int U_n(z)\Psi(z)dL(z), \qquad \mathcal{G}_1 := \int \Psi(z)\Psi(z)'dL(z),$$
$$Q_{n1}(t) := \int \big(U_n(z) - t'\Psi(z)\big)^2 dL(z) = T_{n1}(0) - 2t'V_{n1} + t'\mathcal{G}_1 t, \quad \widetilde{t}_{n1} := \mathrm{argmin}_t Q_{n1}(t).$$

Next, we give a representation of V_{n1} which is useful for determining its limiting distribution. For any d.f. L on $\mathbb{R}_+^p$, let $\overline{L}(y) := \int I(x \ge y)dL(x)$ and

$$\xi_1(x) := \int_{z\ge x} \Psi(z)dL(z), \quad x \ge 0.$$

The Fubini Theorem, L being a d.f. and (4.2) imply

$$\begin{aligned} \xi_1(x) &= \int_{z\geq x} E\big(\varphi(Z_0)I(Z_0\leq z)\big)dL(z) = E\Big(\varphi(Z_0)\overline{L}\big(Z_0\vee x\big)\Big), \\ \sup_{x\geq 0}\big\|\xi_1(x)\big\|^2 &\leq E\big(\big\|\varphi(Z_0)\big\|^2\big) < \infty. \end{aligned}$$

Let Z_{01}, Z_{02} be independent copies of Z_0. Then $\Sigma_1 := E\big(\xi_1(Z_{02})\xi_1(Z_{02})'\big)$ is well defined and

$$\Sigma_1 = E\Big\{E\Big(\varphi(Z_0)\varphi(Z_{01})'\overline{L}\big(Z_0\vee Z_{02}\big)\overline{L}\big(Z_{01}\vee Z_{02}\big)\Big|Z_{02}\Big)\Big\}.$$

Moreover,

$$\begin{aligned} V_{n1} &:= n^{-1/2}\sum_{i=1}^n(\varepsilon_i-1)\int I(Z_{i-1}\leq z)\Psi(z)dL(z) = n^{-1/2}\sum_{i=1}^n(\varepsilon_i-1)\xi_1(Z_{i-1}), \quad (4.3)\\ EV_{n1} &\equiv 0, \qquad E(V_{n1}V_{n1}') \equiv \sigma^2 E(\xi_1(Z_{02})\xi_1(Z_{02})') = \sigma^2\Sigma_1. \end{aligned}$$

We are now ready to state the following main result about $\widehat{\theta}_{n1}$.

Theorem 4.1. *Suppose (1.1), (C.1) and (C.2) hold. Then the following AULQ result holds.*

$$\sup_{\|t\|\leq b}\big|M_{n1}(\theta+n^{-1/2}t)-Q_{n1}(t)\big| = o_p(1), \quad \textit{for every } 0<b<\infty. \qquad (4.4)$$

If, in addition, (C.3) holds, then

$$\|n^{1/2}(\widehat{\theta}_{n1}-\theta)\| = O_p(1). \qquad (4.5)$$

If, further $\mathcal{G}_1$ is positive definite, then $\widetilde{t}_{n1} = \mathcal{G}_1^{-1}V_{n1}$ and

$$\text{(a)}\quad \|n^{1/2}(\widehat{\theta}_{n1}-\theta)-\widetilde{t}_{n1}\| = o_p(1), \qquad \text{(b)}\quad n^{1/2}(\widehat{\theta}_{n1}-\theta)\to_D N\big(0,\sigma^2\mathcal{G}_1^{-1}\Sigma_1\mathcal{G}_1^{-1}\big). \quad (4.6)$$

The proof of Theorem 4.1 appears in the Appendix A below.

4.2 Asymptotic distribution of $\widehat{\theta}_{n2}$

Here, we shall derive the asymptotic distribution of $\widehat{\theta}_{n2}$. This asymptotic distribution is equivalent to that of the $\widehat{\theta}_{n1}$ corresponding to $L=G$. Let

$$\begin{aligned} &\tilde{M}_{n2}(\theta) := \int U_n^2 dG, \quad V_{n2} := \int U_n(z)\Psi(z)dG(z), \quad \mathcal{G}_2 := \int \Psi(z)\Psi(z)'dG(z), \\ &Q_{n2}(t) := \int \big(U_n(z)-t'\Psi(z)\big)^2 dG(z) = \tilde{M}_{n2}(\theta) - 2t'V_{n2} + t'\mathcal{G}_{n2}t, \\ &\widetilde{t}_{n2} := \text{argmin}_t Q_{n2}(t), \qquad \xi_2(x) := \int_{z\geq x}\Psi(z)dG(z), \quad x\geq 0. \end{aligned}$$

By the Fubini Theorem,

$$\mathcal{G}_2 = \int E\big(\varphi(Z_0)I(Z_0 \le z)\big)E\big(\varphi(Z_{01})'I(Z_{01} \le z)\big)dG(z) = E\Big(\varphi(Z_0)\varphi(Z_{01})'\overline{G}\big(Z_0 \vee Z_{01}\big)\Big).$$

The Fubini Theorem and G being a d.f. imply

$$\begin{aligned}\xi_2(x) &= \int_{z\ge x} E\big(\varphi(Z_0)I(Z_0 \le z)\big)dG(z) = E\Big(\varphi(Z_0)\overline{G}\big(Z_0 \vee x\big)\Big),\\ \sup_{x\ge 0}\|\xi_2(x)\|^2 &\le E\big(\|\varphi(Z_0)\|^2\big) < \infty.\end{aligned}$$

Hence $\Sigma_2 := E(\xi_2(Z_{02})\xi_2(Z_{02})')$ is well defined and

$$\Sigma_2 = E\Big\{E\Big(\varphi(Z_0)\varphi(Z_{01})'\overline{G}\big(Z_0 \vee Z_{02}\big)\overline{G}\big(Z_0 \vee Z_{02}\big)\Big|Z_{02}\Big)\Big\}.$$

Moreover,

$$\begin{aligned}V_{n2} &:= n^{-1/2}\sum_{i=1}^{n}(\varepsilon_i - 1)\int I(Z_{i-1} \le z)\Psi(z)dG(z) = n^{-1/2}\sum_{i=1}^{n}(\varepsilon_i - 1)\xi_2(Z_{i-1}),\\ EV_{n2} &\equiv 0, \qquad E(V_{n2}V_{n2}') \equiv \sigma^2 E(\xi_2(Z_0)\xi_2(Z_0)') = \sigma^2\Sigma_2.\end{aligned}$$

Next, assume the following condition, which is the analog of (C.3) for M_{n2}.

(C.4) $\forall\, \epsilon > 0$, $0 < \eta < \infty$, $\exists\, N_{\epsilon,\eta}$, $0 < b \equiv b_{\epsilon,\eta} < \infty$, such that $\forall\, n > N_{\epsilon,\eta}$,

$$P\Big(\inf_{\|t\|>b} M_{n2}(\theta + n^{-1/2}t) \ge \eta\Big) \ge 1 - \epsilon.$$

The following theorem describes the AULQ property of $M_{n2}(\theta + n^{-1/2}t)$ in $\|t\| \le b$ and the asymptotic distribution of $n^{1/2}(\widehat{\theta}_{n2} - \theta)$. Its proof appears in the Appendix A below.

Theorem 4.2. *Under the above set up and (C.1) and (C.2), the following AULQ result holds.*

$$\sup_{\|t\|\le b}\big|M_{n2}(\theta + n^{-1/2}t) - Q_{n2}(t)\big| = o_p(1), \quad \textit{for every } 0 < b < \infty. \tag{4.7}$$

If, in addition, the condition (C.4) holds, then

$$\|n^{1/2}(\widehat{\theta}_{n2} - \theta)\| = O_p(1).$$

If, further $\mathcal{G}_2$ is positive definite, then $\widetilde{t}_{n2} = \mathcal{G}_2^{-1}V_{n2}$ and

(a) $\|n^{1/2}(\widehat{\theta}_{n2} - \theta) - \widetilde{t}_{n2}\| = o_p(1)$, (b) $n^{1/2}(\widehat{\theta}_{n2} - \theta) \to_D N\big(0, \sigma^2\mathcal{G}_2^{-1}\Sigma_2\mathcal{G}_2^{-1}\big)$.

Conditions (C.3) and (C.4). Here we shall discuss a sufficient condition for (C.3) and its analog (C.4) for M_{n2}. Any $\|\vartheta\| > b$ can be written as $\vartheta = re$, for some unit vector $e \in \mathbb{R}^q$, $\|e\| = 1$ and a real number r such that $|r| > b$. Assume the following.

$$\psi(z, \theta + n^{-1/2}re) \text{ is monotonic in } r,\ \forall\, z \in \mathbb{R}_+^p \text{ and } \forall\, e \in \mathbb{R}^q,\ \|e\| = 1. \tag{4.8}$$

The following lemma proves the said sufficiency of (4.8). Its proof appears in Appendix A.

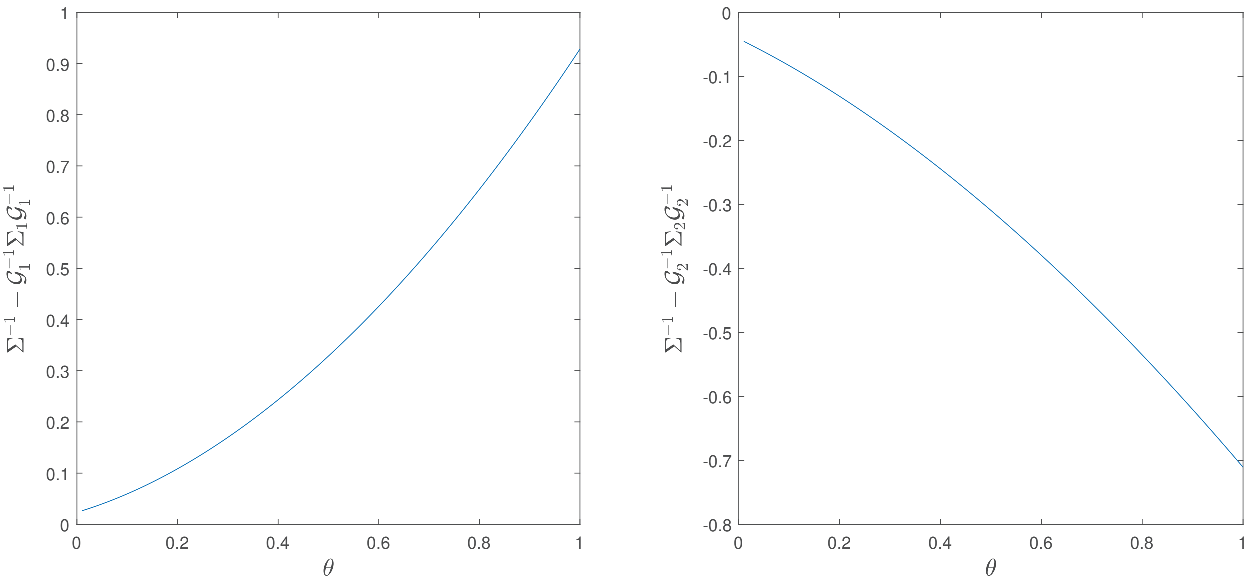
Lemma 4.1. *Assumptions (C.1), (C.2) and (4.8) imply (C.3) and (C.4).*

Remark 4.1. Here we shall make a limited comparison of the asymptotic variance of $\widehat{\theta}_{n1}, \widehat{\theta}_{n2}$ with that of $\widehat{\theta}_{mm}$. The estimator $\widehat{\theta}_{nj}$ is asymptotically more efficient than $\widehat{\theta}_{mm}$ whenever $\Sigma^{-1} - \mathcal{G}_j^{-1}\Sigma_j\mathcal{G}_j^{-1}$ is positive definite, $j = 1, 2$.

The expressions $\mathcal{G}_1^{-1}\Sigma_1\mathcal{G}_1^{-1}$ and $\mathcal{G}_2^{-1}\Sigma_2\mathcal{G}_2^{-1}$ are structurally different compared to Σ^{-1}, and hence it is difficult to compare them with Σ^{-1} analytically in general. However, here we provide a comparison of Σ^{-1} with $\mathcal{G}_j^{-1}\Sigma_j\mathcal{G}_j^{-1}, j = 1, 2$ in a limited restrictive setup.

To this end, let G and L be the uniform distributions on $[0, 1]$ and $[0, 2]$, respectively. Let $\psi(Z_{i-1}, \theta) = 0.1 + \theta Z_{i-1}$, $0 < \theta \le 1$ and $Z_{i-1} = Y_{i-1}, i \in \mathbb{Z}$. Then it is possible to derive algebraic expressions for the differences $\Sigma^{-1} - \mathcal{G}_j^{-1}\Sigma_j\mathcal{G}_j^{-1}$ and observe that $\Sigma^{-1} - \mathcal{G}_1^{-1}\Sigma_1\mathcal{G}_1^{-1} > 0$, for every $0 < \theta \le 1$ and $\Sigma^{-1} - \mathcal{G}_2^{-1}\Sigma_2\mathcal{G}_2^{-1} < 0$, for every $0 < \theta \le 1$. Figure 2 provides plots of $\Sigma^{-1} - \mathcal{G}_1^{-1}\Sigma_1\mathcal{G}_1^{-1}$ and $\Sigma^{-1} - \mathcal{G}_2^{-1}\Sigma_2\mathcal{G}_2^{-1}$ against θ for $0 < \theta \le 1$. Hence, $\widehat{\theta}_{n1}$ is asymptotically

Figure 2: Plots of $\Sigma^{-1} - \mathcal{G}_1^{-1}\Sigma_1\mathcal{G}_1^{-1}$ and $\Sigma^{-1} - \mathcal{G}_2^{-1}\Sigma_2\mathcal{G}_2^{-1}$ against θ, for $0 < \theta \le 1$, when G and L are the uniform distributions on $[0, 1]$ and $[0, 2]$, respectively, $\psi(Z_{i-1}, \theta) = 0.1 + \theta Z_{i-1}$ and $Z_{i-1} = Y_{i-1}, i \in \mathbb{Z}$.



more efficient than $\widehat{\theta}_{mm}$ while $\widehat{\theta}_{mm}$ is asymptotically more efficient than $\widehat{\theta}_{n2}$ in this example.

5 Conclusion

This paper advances the current state of econometric methodology in multiplicative error models for nonnegative valued time series. In particular, we propose a class of m.d. estimators for the underlying parameters in a Markovian parametric multiplicative error model.

We prove the asymptotic normality of the proposed m.d. estimators under fairly general and easily verifiable conditions. The simulation findings about empirical bias and mean squared error demonstrate that the proposed class of estimators can potentially complement, or serve as an alternative to, the GMM estimation approach of Cipollini et al. (2013) in estimating parametric Markovian multiplicative error models.

Acknowledgements

Authors are grateful to the referee and the editor for their useful comments and suggestions on the earlier draft of the paper.

A APPENDIX: Main proofs

A.1 Proofs of Theorems 4.1 and 4.2

Proof of Theorem 4.1. The proof of (4.4) is given later. Arguing as in the proof of Theorem 5.4.1 of Koul (2002), one verifies that (4.1), (C.3) and (4.4) imply (4.5) and (4.6)(a).

To prove (4.6)(b), by the positive definiteness of $\mathcal{G}_1$ we clearly have $\widetilde{t}_{n1} = \mathcal{G}_1^{-1} V_{n1}$. Moreover, V_{n1} is a vector of the sums of martingale difference arrays satisfying (4.3). By the martingale central limit theorem (CLT), see Hall and Heyde (1980), $V_{n1} \to_D N(0, \sigma^2 \Sigma_1)$ and $\widetilde{t}_{n1} = \mathcal{G}_1^{-1} V_{n1} \to_D N(0, \sigma^2 \mathcal{G}_1^{-1} \Sigma_1 \mathcal{G}_1^{-1})$. This fact together with the Slutsky Theorem, (4.5) and (4.6)(a) imply (4.6)(b).

Proof of (4.4). Let $\theta_{nt} := \theta + n^{-1/2} t$, $t \in \mathbb{R}^q$. Define, for $z \in \mathbb{R}_+^p$, $t \in \mathbb{R}^q$,

$$W_n(z,t) := n^{-1/2} \sum_{i=1}^{n} \Big[\frac{\psi(Z_{i-1},\theta)}{\psi(Z_{i-1},\theta_{nt})} - 1\Big]\, \varepsilon_i\, I(Z_{i-1} \le z),$$

$$S_n(z) := n^{-1} \sum_{i=1}^{n} \varphi(Z_{i-1})\,(\varepsilon_i - 1)\, I(Z_{i-1} \le z), \quad T_{n11}(t) := \int \big(W_n(z,t) + t'\Psi(z)\big)^2 dL(z),$$

$$T_{n12}(t) := \int (W_n(z,t) + t'\Psi(z))(U_n(z) - t'\Psi(z)) dL(z).$$

Use the model assumption $\varepsilon_i = Y_i/\psi(Z_{i-1}, \theta)$ to obtain

$$
\begin{aligned}
U_n(z, \theta_{nt}) &= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta_{nt})} - 1\Big] I(Z_{i-1} \le z) && \text{(A.1)}\\
&= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta_{nt})} - \frac{Y_i}{\psi(Z_{i-1}, \theta)} + \frac{Y_i}{\psi(Z_{i-1}, \theta)} - 1\Big] I(Z_{i-1} \le z)\\
&= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{\psi(Z_{i-1}, \theta)}{\psi(Z_{i-1}, \theta_{nt})} - 1\Big] \varepsilon_i \, I(Z_{i-1} \le z) + n^{-1/2} \sum_{i=1}^{n} (\varepsilon_i - 1) I(Z_{i-1} \le z)\\
&= W_n(z, t) + U_n(z), \qquad \forall \, z \in \mathbb{R}_+^p, \; t \in \mathbb{R}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
T_{n1}(t) &= \int U_n^2(z, \theta_{nt}) dL(z) = \int \Big(W_n(z, t) + t'\Psi(z) + U_n(z) - t'\Psi(z)\Big)^2 dL(z) && \text{(A.2)}\\
&= T_{n11}(t) + 2T_{n12}(t) + Q_{n1}(t).
\end{aligned}
$$

We shall shortly prove the following facts. For every $0 < b < \infty$,

$$
\text{(a)} \quad \sup_{\|t\| \le b} T_{n11}(t) = o_p(1), \qquad \text{(b)} \quad E\Big(\sup_{\|t\| \le b} Q_{n1}(t)\Big) = O(1). \qquad \text{(A.3)}
$$

Then by the Cauchy-Schwarz inequality, $\sup_{\|t\| \le b} \big|T_{n12}(t)\big|^2 \le \sup_{\|t\| \le b} T_{n11}(t) \sup_{\|t\| \le b} Q_{n1}(t) = o_p(1)$. Hence the claim (4.4).

Proof of (A.3)(a). Rewrite

$$
\begin{aligned}
W_n(z, t) &:= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{\psi(Z_{i-1}, \theta)}{\psi(Z_{i-1}, \theta_{nt})} - 1\Big] \varepsilon_i \, I(Z_{i-1} \le z)\\
&= -n^{-1/2} \sum_{i=1}^{n} \frac{1}{\psi(Z_{i-1}, \theta_{nt})} \Big[\psi(Z_{i-1}, \theta_{nt}) - \psi(Z_{i-1}, \theta) - n^{-1/2} t'\dot{\psi}(Z_{i-1}, \theta)\Big] \varepsilon_i \, I(Z_{i-1} \le z)\\
&\quad - t'n^{-1} \sum_{i=1}^{n} \Big[\frac{1}{\psi(Z_{i-1}, \theta_{nt})} - \frac{1}{\psi(Z_{i-1}, \theta)}\Big] \dot{\psi}(Z_{i-1}, \theta) \, \varepsilon_i \, I(Z_{i-1} \le z)\\
&\quad - t'n^{-1} \sum_{i=1}^{n} \varphi(Z_{i-1})(\varepsilon_i - 1) \, I(Z_{i-1} \le z)\\
&\quad - t'n^{-1} \sum_{i=1}^{n} \big[\varphi(Z_{i-1}) \, I(Z_{i-1} \le z) - E\big(\varphi(Z_0) I(Z_0 \le z)\big)\big] - t'\Psi(z).
\end{aligned}
$$

Let $d_{it} := \psi(Z_{i-1}, \theta_{nt}) - \psi(Z_{i-1}, \theta)$ and $\delta_{it} := d_{it} - n^{-1/2} t'\dot{\psi}(Z_{i-1}, \theta)$. Then the above identity is equivalent to

$$
\begin{aligned}
&W_n(z,t) + t'\Psi(z) && \text{(A.4)}\\
&= -n^{-1/2}\sum_{i=1}^n \frac{\delta_{it}}{\psi(Z_{i-1},\theta_{nt})}\,\varepsilon_i\, I(Z_{i-1}\le z) + t'n^{-1}\sum_{i=1}^n \frac{d_{it}}{\psi(Z_{i-1},\theta_{nt})}\varphi(Z_{i-1})\,\varepsilon_i\, I(Z_{i-1}\le z)\\
&\qquad - t'S_n(z) - t'\big(\Psi_n(z)-\Psi(z)\big)\\
&= A_1(z,t) + A_2(z,t) - t'S_n(z) - t'\big(\Psi_n(z)-\Psi(z)\big), \qquad \text{say.}
\end{aligned}
$$

In the sequel, the range of the vectors z,ϑ in $\inf_{z,\vartheta}$ is over $\mathbb{R}_+^p \times \Theta$, unless mentioned otherwise. By (C.1), $\inf_{z,\vartheta}\psi(z,\vartheta) \ge C > 0$ and for every $b<\infty$, $\sup_{1\le i\le n,\|t\|\le b} n^{1/2}\big|\delta_{it}\big| = o_p(1)$. Moreover, because $E(\varepsilon_0)=1$, we have $n^{-1}\sum_{i=1}^n \varepsilon_i = O_p(1)$. Hence

$$
\sup_{z\in\mathbb{R}_+^p,\|t\|\le b} |A_1(z,t)| \le C^{-1}\sup_{1\le i\le n,\|t\|\le b} n^{1/2}\big|\delta_{it}\big| n^{-1}\sum_{i=1}^n \varepsilon_i = o_p(1). \tag{A.5}
$$

Next, recall that the stationarity of the process $Y_i, i\in\mathbb{Z}$ and $E\|\dot\psi(Z_0,\theta)\|^2<\infty$ imply that $n^{-1/2}\max_{1\le i\le n}\|\dot\psi(Z_{i-1},\theta)\| = o_p(1)$, $E\big(n^{-1}\sum_{i=1}^n \|\varphi(Z_{i-1})\|\,\varepsilon_i\big) = E\|\varphi(Z_0)\|<\infty$, and by the Markov inequality, $n^{-1}\sum_{i=1}^n \|\varphi(Z_{i-1})\|\,\varepsilon_i = O_p(1)$. Hence, (C.1) and (C.2) imply that

$$
\begin{aligned}
D_n &:= \sup_{1\le i\le n,\|t\|\le b}|d_{it}| \le \sup_{1\le i\le n,\|t\|\le b}\big|\delta_{it}\big| + bn^{-1/2}\max_{1\le i\le n}\|\dot\psi(Z_{i-1},\theta)\| = o_p(1), && \text{(A.6)}\\
&\sup_{z\in\mathbb{R}_+^p,\|t\|\le b}|A_2(z,t)| \le b\,C^{-1}D_n n^{-1}\sum_{i=1}^n \|\varphi(Z_{i-1})\|\,\varepsilon_i = o_p(1).
\end{aligned}
$$

Next, consider $S_n(z)$. Observe that $S_n(z)$ is a vector of weighted empirical processes with the summands of each component being stationary and ergodic and $ES_n(z)\equiv 0$. Using a Glivenko-Cantelli Lemma type argument one obtains that $\sup_{z\in\mathbb{R}_+^p}\|S_n(z)\| = o_p(1)$. For details see Stute (1976) and Koul (2019). Similarly, $\sup_{z\in\mathbb{R}_+^p}\big\|\Psi_n(z)-\Psi(z)\big\|^2 = o_p(1)$. Upon combining these two facts with (A.4), (A.5) and (A.6) we obtain that for every $0<b<\infty$,

$$
\sup_{z\in\mathbb{R}_+^p,\|t\|\le b}\big|W_n(z,t)+t'\Psi(z)\big| = o_p(1). \tag{A.7}
$$

This fact combined with the definition of T_{n11} and L being a d.f. readily yields (A.3)(a).

Next, to prove (A.3)(b), note that

$$
\begin{aligned}
Q_{n1}(t) &:= \int \Big(U_n(z,\theta) - t'\Psi(z)\Big)^2 dL(z) \le 2T_{n1}(0) + 2t'\mathcal{G}_1 t,\\
E\Big(\sup_{\|t\|\le b} Q_{n1}(t)\Big) &\le 2ET_{n1}(0) + 2b^2\int \|\Psi(z)\|^2 dL(z)\\
&= 2\int G(z)dL(z)\big[1+b^2E\|\varphi(Z_0)\|^2\big] < \infty.
\end{aligned}
$$

This also completes the proof of Theorem 4.1. □

Proof of Theorem 4.2. The proof is similar to that of Theorem 4.1, with some differences. Let

$$\tilde V_{n2} := \int U_n(z)\Psi(z)dG_n(z), \qquad \mathcal{G}_{n2} := \int \Psi(z)\Psi(z)'dG_n(z),$$
$$\tilde Q_{n2}(t) := \int \big(U_n(z) - t'\Psi(z)\big)^2 dG_n(z) = M_{n2}(\theta) - 2t'\tilde V_{n2} + t'\mathcal{G}_{n2}t,$$
$$\tilde T_{n2}(t) := M_{n2}(\theta + n^{-1/2}t), \qquad \tilde T_{n21}(t) := \int \big(W_n(z,t) + t'\Psi(z)\big)^2 dG_n(z),$$
$$\tilde T_{n12}(t) := \int (W_n(z,t) + t'\Psi(z))(U_n(z) - t'\Psi(z))dG_n(z).$$

Then akin to (A.2),

$$\tilde T_{n2}(t) = \int U_n^2(z,\theta_{nt})dG_n(z) = \tilde T_{n21}(t) + 2\tilde T_{n22}(t) + \tilde Q_{n2}(t).$$

Now, $\sup_{\|t\|\le b}\tilde T_{n21}(t) \le \sup_{z\in\mathbb{R}_+^p,\|t\|\le b}\big(W_n(z,t) + t'\Psi(z)\big)^2 = o_p(1)$, by (A.7). We shall shortly prove that for every $0 < b < \infty$,

$$\sup_{\|t\|\le b}\big|\tilde Q_{n2}(t) - Q_{n2}(t)\big| = o_p(1). \tag{A.8}$$

Argue as for (A.3)(b) to conclude that $E\big(\sup_{\|t\|\le b} Q_{n2}(t)\big) = O(1)$. Hence $\sup_{\|t\|\le b}\tilde Q_{n2}(t) = O_p(1)$ and $\sup_{\|t\|\le b}\big|\tilde T_{n22}(t)\big|^2 \le \sup_{\|t\|\le b}\tilde T_{n21}(t)\sup_{\|t\|\le b}\tilde Q_{n2}(t) = o_p(1)$. These facts together yield the claim (4.7).

Next, to prove (A.8), note that the left hand side of (A.8) is bounded from the above by

$$\Big|\int U_n^2\big[dG_n - dG\big]\Big| + b\Big\|\int U_n\Psi\big[dG_n - dG\big]\Big\| + b^2\Big\|\int \Psi\Psi'\big[dG_n - dG\big]\Big\|.$$

By the Ergodic Theorem, $\big\|\int \Psi\Psi'\big[dG_n - dG\big]\big\| \to 0$, a.s. while by Lemmas A.1 and A.2 below, the first two terms tend to zero in probability. This completes the proof of (4.7). The proofs of the other two claims of this theorem are similar to those of (4.5) and (4.6)(a), (b) of Theorem 4.1. □

A.2 Proof of Lemma 4.1

Proof of the claim about (C.3). Let h be a positive function on $\mathbb{R}_+^p$ with $0 < \int h^2 dL < \infty$ and let $\boldsymbol{\infty}$ ($\mathbf{0}$) denote the vector of p infinities (zeros). Let $\varphi(z) := \int_{x\le z} h(x)dL(x)$. Note that φ is nondecreasing in each coordinate and $\varphi(\boldsymbol{\infty}) < \infty$. Moreover, $\gamma(z) := \int I(z \le x)h(x)dL(x) = \int I(z \le x)d\varphi(x) \ge 0$, for all $z \in \mathbb{R}_+^p$. Define

$$\mathcal{V}_{n1}(t) := \int U_n(z,\theta_{nt})h(z)dL(z) = \int U_n(z,\theta_{nt})d\varphi(z)$$

$$= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta + n^{-1/2}t)} - 1\Big] \int I(Z_{i-1} \le z) d\varphi(z)$$
$$= n^{-1/2} \sum_{i=1}^{n} \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta + n^{-1/2}t)} - 1\Big] \gamma(Z_{i-1}).$$

By the Cauchy-Schwarz inequality,

$$M_{n1}(\theta + n^{-1/2}t) \ge \Big(\int U_n(z, \theta_{nt}) h(z) dL(z)\Big)^2 \Big/ \int h^2 dL, \quad \forall\, t \in \mathbb{R}^q.$$

Because $\gamma(z) \ge 0$ for all z, by (4.8), $\forall\, e \in R^q, \|e\| = 1, \mathcal{V}_{n1}(re)$ is monotonic in r. Hence

$$\inf_{\|t\|>b} M_{n1}(\theta + n^{1/2}t) = \inf_{e \in \mathbb{R}^q, \|e\|=1, |r|>b} M_n(\theta + n^{-1/2}re) \ge \inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \mathcal{V}_{n1}^2(re) \Big/ \int h^2 dL.$$

Let

$$\widehat{\mathcal{V}}_{n1}(t) := \int \big(U_n(z) - t'\Psi(z)\big) h(z) dL(z) = \int U_n(z) d\varphi(z) - t' \int \Psi(z) d\varphi(z)$$
$$= \mathcal{V}_{n1}(0) - t' \int \Psi(z) d\varphi(z).$$

In view of (A.1),

$$\mathcal{V}_{n1}(t) - \widehat{\mathcal{V}}_{n1}(t) = \int \big\{U_n(z, \theta_{nt}) - U_n(z) + t'\Psi(z)\big\} d\varphi(z) = \int \big\{W_n(z, \theta_{nt}) + t'\Psi(z)\big\} d\varphi(z).$$

Because $\varphi(\infty) = \int h dL < \infty$, by (A.7),

$$\sup_{\|t\| \le b} \Big|\mathcal{V}_{n1}(t) - \widehat{\mathcal{V}}_{n1}(t)\Big| \le \sup_{\|t\| \le b} \big|W_n(z, \theta_{nt}) + t'\Psi(z)\big| \varphi(\infty) = o_p(1).$$

Therefore,

$$\Big|\inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \mathcal{V}_{n1}(re) - \inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \widehat{\mathcal{V}}_{n1}(re)\Big| = o_p(1). \tag{A.9}$$

Let $\tau_2 := \int h^2 dL$. Fix an $\epsilon > 0$ and $0 < \eta < \infty$. By (A.9), there exists $N_{\epsilon,\eta}$ such that

$$P\Big(\inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \Big|\mathcal{V}_{n1}(re)\Big| \ge (\tau_2 \eta)^{1/2}\Big)$$
$$\ge P\Big(\inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \Big|\widehat{\mathcal{V}}_{n1}(re)\Big| \ge (\tau_2 \eta)^{1/2}\Big) - \frac{\epsilon}{2}, \qquad \forall\, n > N_{\epsilon,\eta}.$$

Moreover, by the Cauchy-Schwarz inequality, $E\mathcal{V}_{n1}^2(0) = \sigma^2 E\gamma^2(Z_0) \le \sigma^2 \int h^2 dL < \infty$. Hence by the Markov inequality, for every $\epsilon > 0$ there is a b_ϵ such that

$$P\big(|\mathcal{V}_{n1}(0)| \le b_\epsilon\big) \ge 1 - (\epsilon/2), \quad \forall\, n \ge 1. \tag{A.10}$$

Recall the elementary fact that for any real numbers a, b, $\big||a| - |b|\big| \le |a \pm b|$. Also let $K := \int \|\Gamma\| d\varphi$. Let b_ϵ be as in (A.10). Choose $b \ge (b_\epsilon + (v_2\eta)^{1/2})K^{-1}$. Then, $\forall\ n \ge N_{\epsilon,\eta}$,

$$
\begin{aligned}
P\Big(\inf_{|t|>b} M_{n1}(\theta + n^{-1/2}t) \ge \eta\Big) &\ge P\Big(\inf_{e\in\mathbb{R}^q, \|e\|=1, |r|=b} \Big|\mathcal{V}_{n1}(re)\Big| \ge (\tau_2\eta)^{1/2}\Big) \qquad \text{(A.11)}\\
&= P\Big(\Big|\mathcal{V}_{n1}(\pm be)\Big| \ge (\tau_2\eta)^{1/2},\ \forall\, \|e\| = 1\Big)\\
&\ge P\Big(\Big|\widehat{\mathcal{V}}_{n1}(\pm be)\Big| \ge (\tau_2\eta)^{1/2},\ \forall\, \|e\| = 1\Big) - \epsilon/2\\
&\ge P\Big(\ \Big||\mathcal{V}_{n1}(0)| - bK\Big|\ \ge (\tau_2\eta)^{1/2}\Big) - \epsilon/2\\
&\ge P\Big(\ |\mathcal{V}_{n1}(0)| \le bK - (\tau_2\eta)^{1/2}\Big) - \epsilon/2\\
&\ge P\Big(\ |\mathcal{V}_{n1}(0)| \le b_\epsilon\Big) - \epsilon/2 \ge 1 - \epsilon,
\end{aligned}
$$

thereby proving the claim about (C.3).

Proof of the claim about (C.4). The proof of this claim is similar to that of the previous claim, but with some differences. In particular we need the following two preliminary lemmas.

Lemma A.1. *Under the above set up, U_n converges weakly to a continuous Gaussian process $\mathcal{Z}(x), x \in \mathbb{R}^p_+$ in Skorokhod space $D([0,\infty]^p)$ and uniform metric.*

The proof of this lemma is similar to that of the main theorem in Stute (1976).

Lemma A.2. *Let $\mathcal{U}$ be a relatively compact subset of $D[0,\infty]^p$. Let μ_n, μ be a sequence of possibly random multivariate distribution functions on $[0,\infty)^p$ such that $\sup_{x\in\mathbb{R}^p_+} \big|\mu_n(x) - \mu(x)\big| \to 0$, a.s. Then $\sup_{y\in\mathbb{R}^p_+, \alpha\in\mathcal{U}} \Big|\int_{x\le y} \alpha(x)[d\mu_n(x) - d\mu(x)]\Big| \to_p 0$.*

The proof of this lemma is similar to that of Lemma 3.1 of Chang (1990) and Lemma 4.2 of Koul and Stute (1999). Details are left out for the sake of brevity.

Now, let h be a positive function on $\mathbb{R}^p_+$ with $0 < \int h^2 dG < \infty$ and $\int h dG = 1$. Let

$$
\begin{aligned}
\beta_{nk}(z) &:= \int_{x\le z} h^k(x) dG_n(x) = n^{-1} \sum_{j=1}^n h^k(Z_{j-1}) I(Z_{j-1} \le z),\\
\beta_k(z) &:= \int_{x\le z} h^k(x) dG(x), \quad k = 1, 2,\ z \in \mathbb{R}^p_+, \qquad B_{n2} := \beta_{n2}(\boldsymbol{\infty}), \quad B_2 := \beta_2(\boldsymbol{\infty}).
\end{aligned}
$$

Note that $\beta_{nk}(\mathbf{0}) = 0 = \beta_k(\mathbf{0})$, $0 < \beta_k(z) \le \beta_k(\boldsymbol{\infty}) = \int h^k dG < \infty$, for all $z \in \mathbb{R}^p_+$ and $E\beta_{nk}(z) \equiv \beta_k(z)$, $k = 1, 2$. By the Ergodic Theorem and a Glivenko-Cantelli type argument, see Stute (1976) and Koul (2019),

$$
\sup_{z\in[0,\infty]^p} \big|\beta_{nk}(z) - \beta_k(z)\big| \to 0, \quad \text{a.s.}, k = 1, 2; \quad |B_{n2} - B_2| \to 0, \text{ a.s.} \qquad \text{(A.12)}
$$

Thus for all sufficiently large n, $\beta_{nk}(z) > 0$, for all $z \in (0,\infty]^p$, $k = 1, 2$, $B_{n2} > 0$ (a.s.). The arguments below are carried out on the event $B_{n2} = \int h^2 dG_n > 0$.

Recall $\theta_{nt} = \theta + n^{-1/2}t$. By the Cauchy-Schwarz inequality,

$$M_{n2}(\theta + n^{-1/2}t) \geq \Big(\int U_n(z, \theta_{nt})h(z)dG_n(z)\Big)^2 \Big/ B_{n2}, \quad \forall\, t \in \mathbb{R}^q.$$

Let $\alpha(z) := \int_{x \geq z} d\beta_1(x)$, $\alpha_n(z) := \int_{x \geq z} d\beta_{n1}(x)$, $z \in \mathbb{R}_+^p$ and

$$\begin{aligned}
\mathcal{V}_{n2}(t) &:= \int U_n(z, \theta_{nt})h(z)dG_n(z) = n^{-1/2}\sum_{i=1}^n \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta_{nt})} - 1\Big] \int I(Z_{i-1} \leq z)d\beta_{n1}(z) \\
&= n^{-1/2}\sum_{i=1}^n \Big[\frac{Y_i}{\psi(Z_{i-1}, \theta + n^{-1/2}t)} - 1\Big]\alpha_n(Z_{i-1}).
\end{aligned}$$

Because $\alpha_n(z) \geq 0$ for all $z \in \mathbb{R}_+^p, n \geq 1$, w.p.1, by (4.8), $\mathcal{V}_{n2}(re)$ is monotonic in r for all $e \in R^q, \|e\| = 1$. Hence

$$\inf_{\|t\|>b} M_{n2}(\theta + n^{1/2}t) = \inf_{e \in \mathbb{R}^q, \|e\|=1, |r|>b} M_{n2}(\theta + n^{-1/2}re) \geq \inf_{e \in \mathbb{R}^q, \|e\|=1, |r|=b} \mathcal{V}_{n2}^2(re) \Big/ B_{n2}.$$

Let

$$\begin{aligned}
\widehat{\mathcal{V}}_{n2}(t) &:= \int \big(U_n(z) - t'\Psi(z)\big)h(z)dG_n(z) = \int U_n(z)d\beta_{n1}(z) - t'\int \Psi(z)d\beta_{n1}(z) \\
&= \mathcal{V}_{n2}(0) - t'\int \Psi(z)d\beta_{n1}(z), \\
\mathcal{V}_n^* &:= \int U_n(z)d\beta_1(z) = n^{-1/2}\sum_{i=1}^n (\varepsilon_i - 1)\alpha(Z_{i-1}).
\end{aligned}$$

By Lemma A.1, the process $U_n(z), z \in [0, \infty]^p$ is relatively compact with respect to the uniform metric. Hence (A.12) and Lemma A.2 applied with $\mu_n = \beta_{n1}/\beta_{n1}(\infty)$, $\mu = \beta_1$ yield

$$\big|\mathcal{V}_{n2}(0) - \mathcal{V}_n^*\big| = \Big|\int U_n(z)[d\beta_{n1}(z) - d\beta_1(z)]\Big| = o_p(1).$$

Here we have used the fact $\beta_1(\infty) = \int h dG = 1$ and $\beta_{n1}(\infty) \to_p \beta_1(\infty) = 1$.

Let $\bar{\mathcal{V}}_{n2}(t) := \mathcal{V}_n^* - t'\int \Psi d\beta_1$. Because by the Ergodic Theorem, $\int \Psi d\beta_{n1} \to_p \int \Psi d\beta_1$ and

$$\sup_{\|t\| \leq b} \big|\widehat{\mathcal{V}}_{n2}(t) - \bar{\mathcal{V}}_{n2}(t)\big| = o_p(1), \tag{A.13}$$

by (A.1), we obtain

$$\mathcal{V}_{n2}(t) - \widehat{\mathcal{V}}_{n2}(t) = \int \big\{U_n(z, \theta_{nt}) - U_n(z) + t'\Psi(z)\big\}d\beta_{n1}(z) = \int \big[W_n(z, \theta_{nt}) + t'\Psi(z)\big]d\beta_{n1}(z).$$

Therefore, by (A.7), (A.12) and (A.13)

$$\begin{aligned}
\sup_{\|t\| \leq b}\Big|\mathcal{V}_{n2}(t) - \bar{\mathcal{V}}_{n2}(t)\Big| &\leq \sup_{\|t\| \leq b}\Big|\mathcal{V}_{n2}(t) - \widehat{\mathcal{V}}_{n2}(t)\Big| + \sup_{\|t\| \leq b}\Big|\widehat{\mathcal{V}}_{n2}(t) - \bar{\mathcal{V}}_{n2}(t)\Big| \\
&\leq \sup_{\|t\| \leq b}\big|W_n(z, \theta_{nt}) + t'\Psi(z)\big|\beta_{n1}(\infty) + o_p(1) = o_p(1).
\end{aligned}$$

Also note that $\sup_{\|t\|\le b}|\bar{\mathcal{V}}_{n2}(t)| = O_p(1) = \sup_{\|t\|\le b}|\mathcal{V}_{n2}(t)|$ and by (A.12),

$$\sup_{\|t\|\le b}\Big|\frac{\mathcal{V}_{n2}(t)}{B_{n2}^{1/2}} - \frac{\bar{\mathcal{V}}_{n2}(t)}{B_2^{1/2}}\Big| \le \sup_{\|t\|\le b}\big|\mathcal{V}_{n2}(t)\big|\,\Big|\frac{1}{B_{n2}^{1/2}} - \frac{1}{B_2^{1/2}}\Big| + \frac{1}{B_2^{1/2}}\sup_{\|t\|\le b}\big|\mathcal{V}_{n2}(t) - \bar{\mathcal{V}}_{n2}(t)\big| = o_p(1).$$

This fact in turn implies that

$$\Big|\inf_{e\in\mathbb{R}^q,\|e\|=1,|r|=b}\frac{\mathcal{V}_{n2}(re)}{B_{n2}^{1/2}} - \inf_{e\in\mathbb{R}^q,\|e\|=1,|r|=b}\frac{\bar{\mathcal{V}}_{n2}(re)}{B_2^{1/2}}\Big| = o_p(1). \tag{A.14}$$

Fix an $\epsilon > 0$ and $\eta > 0$. By (A.14), there exists $N_{\epsilon,\eta}$ such that

$$\begin{aligned} &P\Big(\inf_{e\in\mathbb{R}^q,\|e\|=1,|r|=b}\frac{\big|\mathcal{V}_{n2}(re)\big|}{B_{n2}^{1/2}} \ge \eta^{1/2}\Big) \\ &\qquad \ge P\Big(\inf_{e\in\mathbb{R}^q,\|e\|=1,|r|=b}\big|\bar{\mathcal{V}}_{n2}(re)\big| \ge (B_2\eta)^{1/2}\Big) - \frac{\epsilon}{2}, \qquad \forall\, n > N_{\epsilon,\eta}. \end{aligned}$$

Moreover, $E\big(\mathcal{V}_n^*\big)^2 = \sigma^2 E\alpha^2(Z_0) \le \sigma^2 B_2 < \infty$. Hence, by the Markov inequality, for any $\epsilon > 0$ there exists b_ϵ such that

$$P\big(|\mathcal{V}_n^*| \le b_\epsilon\big) \ge 1 - (\epsilon/2), \qquad \forall\, n \ge 1. \tag{A.15}$$

Let $K := \int\|\Gamma\|d\beta$ and b_ϵ be as in (A.15). Choose $b \ge (b_\epsilon + (B_2\eta)^{1/2})K^{-1}$ and argue as for (A.11) to obtain that $\forall\ n \ge N_{\epsilon,\eta}$,

$$\begin{aligned} P\Big(\inf_{|t|>b} M_{n2}(\theta_0 + n^{-1/2}t) \ge \eta\Big) &\ge P\Big(\ |\mathcal{V}_n^*| \le bK - (B_2\eta)^{1/2}\Big) - \epsilon/2 \\ &\ge P\Big(\ |\mathcal{V}_{n1}^*| \le b_\epsilon\Big) - \epsilon/2 \ge 1 - \epsilon, \end{aligned}$$

thereby proving the claim about (C.4). This also completes the proof of Lemma 4.1. □

Funding and Conflicts of interest

N. Balakrishna received partial financial suppport from SERB of India under the MATRICS scheme MTR/2018/000195. The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Bauwens, L. and Giot, P. (2000). The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Ann. Econ. Stat.*, **60**, 117–149.

Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *J. Econometrics*, **119(2)**, 381–412.

Brownlees, C.T., Cipollini, F. and Gallo, G.M. (2012). Multiplicative Error Models. In: Bauwens, L., Hafner, C., Laurent, S. (Eds.), Handbook of Volatility Models and Their Applications. John Wiley and Sons, Inc., Hoboken, New Jersey, pp. 225–247.

Chang, N. M. (1990). Weak convergence of a self-consistent estimator of a survival function with doubly censored data. *Ann. Statist.*, **18**, 391–404.

Chen, Y.-T. and Hsieh, C.-S. (2010). Generalized moment tests for autoregressive conditional duration models. *J. Financ. Econom.*, **8**, 345–391.

Chou, R.Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. *J. Money, Credit Banking*, **37(3)**, 561–582.

Cipollini, F., Engle, R.F. and Gallo, G.M. (2013). Semiparametric Vector MEM. *J. Appl. Econ.*, **28**, 1067–86.

Donoho, D.L. and Liu, R.C. (1988a). The "automatic" robustness of minimum distance functionals. *Ann. Statist.*, **16**, 552–586.

Donoho, D.L. and Liu, R.C. (1988b). Pathologies of some minimum distance estimators. *Ann. Statist.*, **16**, 587–608.

Drost, F. C. and Werker, B. J. M. (2004). Semiparametric duration models. *J. Bus. Econ. Stat.*, **22**(1), 40–50.

Engle, R. (2002) New frontiers for ARCH models. *J. Appl. Econ.*, **17**, 425–446.

Engle, R.F. and Gallo, G.M. (2006). A multiple indicators model for volatility using intradaily data. *J. Econometrics*, **131**, 3–27.

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

Fernandes, M. and Grammig, J. (2006). A family of autoregressive conditional duration models. *J. Econometrics*, **130(1)**, 1–23.

Grammig, J. and Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal*, **3**(1), 16–38.

Grammig, J. and Wellner, M. (2002). Modeling the interdependence of volatility and inter-transaction duration processes. *J. Econometrics*, **106**(2), 369–400.

Gao, J., Kim, N.H. and Saart, P.W. (2015). A misspecification test for multiplicative error models of non-negative time series processes. *J. Econometrics*, **189(2)**, 346–359.

Guo, B. and Li, S. (2018). Diagnostic checking of markov multiplicative error models. *Economics Letters*, **170**, 139 – 142.

Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its applications.* Academic Press, New York.

Hautsch, N. (2012). *Econometrics of Financial High-Frequency Data.* Springer, Heidelberg.

Koul, Hira L. (1985a). Minimum distance estimation in multiple linear regression. *Sankhya, Ser. A.*, **47, Part 1**, 57–74.

Koul, Hira L. (1985b). Minimum distance estimation in linear regression with unknown errors. *Statist. & Prob. Letters*, **3**, 1–8.

Koul, Hira L. (1986). Minimum distance estimation and goodness-of-fit tests in first-order autoregression. *Ann. Statist.*, **14**, 1194–1213.

Koul, Hira L. (1996). Asymptotics of some estimators and sequential residual empiricals in non-linear time series. *Ann. Statist.*, **24**, 380–404.

Koul, Hira L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models. 2nd Edition.* **Lecture Notes Series in Statistics, 166,** Springer, New York, N.Y., USA.

Koul, Hira L. (2019). A uniform convergence result for weighted residual empirical processes. RM#719, Department of Statistics and Probability, MSU.

Koul, H.L. and Stute, W. (1999). Nonparametric model checks for time series. *Ann. Statist.*, **27**, 204–236.

Koul, H.L. and Perera, I. (2021). A minimum distance lack-of-fit test in a markovian multiplicative error model. *J. Stat. Theory Pract.*, **15**, 34.

Koul, H.L., Perera, I. and Silvapulle, M.J. (2012). Lack-of-fit testing of the conditional mean function in a class of markov multiplicative error models. *Econometric Theory*, **28**, 1283–1312.

Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. *Discussion paper, Aalborg University.*

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *J. Financ. Mark.*, **8(4)**, 377–399.

Meitz, M. and Teräsvirta, T. (2006). Evaluating models of autoregressive conditional duration. *J. Bus. Econ. Stat.*, **24**, 104–124.

Pacurar, M. (2008) Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *J. of Economic Surveys*, **22**, 711–751.

Perera, I., Hidalgo, J., and Silvapulle, M. J. (2016). A goodness-of-fit test for a class of autoregressive conditional duration models. *Econometric Reviews*, **35**(6), 1111–1141.

Perera, I. and Koul, H.L. (2017). Fitting a two phase threshold multiplicative error model. *J. Econometrics*, **197**, 348–367

Perera, I. and Silvapulle, M. J. (2021). Bootstrap based probability forecasting in multiplicative error models. *J. Econometrics*, **221**(1), 1–24.

Stute, W. (1976). On a generalization of the Glivenko-Cantelli theorem. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **35(2)**, 167–175.

Stute, W. (1997). Nonparametric model checks for regression. *Ann. Statist.*, **25(2)**, 613–641.