Bootstrapping a linear estimator of the ARCH parameters

ARUP BOSE
Stat-Math Unit, Indian Statistical Institute, Kolkata,
Kolkata - 700108, India
and
Department of Economics, University of Cincinnati
Cincinnati, Ohio 45221
Email address: abose@isical.ac.in
and
KANCHAN MUKHERJEE
Dept. of Mathematical Sciences, The University of Liverpool,
Liverpool L69 7ZL, UK
Email address: k.mukherjee@liverpool.ac.uk

SUMMARY

A standard assumption while deriving the asymptotic distribution of the quasi likelihood estimator in ARCH models is that all ARCH parameters must be strictly positive. This assumption is also crucial in deriving the limit distribution of appropriate linear estimators.

We propose a weighted linear estimator (WLE) of the ARCH parameters in the classical ARCH model and show that its limit distribution is multivariate normal even when some of the ARCH coefficients are zero.

The asymptotic dispersion matrix involves unknown quantities. We consider appropriate bootstrapped version of the above WLE and prove that it is asymptotically valid in the sense that the bootstrapped distribution (given the data) is a consistent estimate (in probability) of the distribution of the WLE.

Though we do not show theoretically that the bootstrap outperforms the normal approximation, our simulations demonstrate that it yields better approximations than the limiting normal.

Running Head: Bootstrapping ARCH parameters

Some key words: ARCH model: QMLE: Bootstrapping.

1 Introduction

Volatility or the instantaneous variability of a financial time series is an important concept in many econometric models. In his seminal work, Engle (1982) modeled volatility of a series $\{X_t; 1 \leq t \leq n\}$ as a linear function of the squares of past observations and called it the autoregressive conditional heteroscedastic model or the ARCH model. ARCH models have been applied to numerous economic and financial data. For a survey of the literature, see for example, the book by Gouriéroux (1997).

Consider the following ARCH model where one observes $\{X_i; 1-p \leq i \leq n\}$ satisfying

$$X_i = \sigma_{i-1}(\boldsymbol{\beta})\epsilon_i;\ 1 \leq i \leq n, \tag{1.1}$$

with $\boldsymbol{\beta}' = [\beta_0, \beta_1, \cdots \beta_p]$ is the unknown parameter to be estimated; $\beta_0 > 0$ and $\beta_j \geq 0$, for $0 < j \leq p$; $\sigma_{i-1}(\boldsymbol{\beta}) = \{\beta_0 + \beta_1 X_{i-1}^2 + \cdots + \beta_p X_{i-p}^2\}^{1/2}$; $\{\epsilon_i;\ \ 1 \leq i \leq n\}$ are independent and identically distributed (iid) with zero mean and unit variance; and $\{\epsilon_i; 1 \leq i \leq n\}$ are independent of $\{X_i; 1-p \leq i \leq 0\}$. It is tacitly assumed that $\beta_p > 0$ so that p, which is assumed to be known, is the order of the model.

In this article, we are concerned with the problem of estimating $\boldsymbol{\beta}$. One of the most commonly used estimation procedures for ARCH models is the Gaussian likelihood approach. In this approach, the estimator is obtained as a maximizer of the logarithm of a Gaussian likelihood function. The resulting estimator is called the quasi-maximum likelihood estimator (QMLE). This yields a consistent estimator even when the conditional error density is non normal. The consistency and asymptotic normality of the QMLE was established by Weiss (1986).

However, the QMLE does not admit a closed form expression and is not easy to compute. The likelihood tends to be flat unless n is very large. A discussion of this problem can be found in Shephard (1996). As is the case with any estimator defined through an optimization problem, one also needs to specify a definite rule to choose the estimator in case of multiple solutions.

The proposal of Bose and Mukherjee (2003) circumvents these problems. They proposed an estimator whose computation involves solving two sets of linear equations. Nevertheless, it compares favourably with the QMLE in terms of accuracy. This linear estimator (termed as LE hereafter) turned out to have the same limiting dispersion as the QMLE and hence is also fully efficient when the error distribution is Gaussian. Moreover, the simulation results of Bose and Mukherjee (2003) showed that the LE performs better than the QMLE even for samples of size as small as $n = 30$ for a variety of error distributions.

However, one of the crucial assumptions for deriving the limiting distribution of the LE is that all ARCH parameters β_j must be strictly positive. This restricts the applications of the result. Also, in order to test any hypothesis regarding the parameters equal to zero, one needs to know the limiting distribution of the QMLE or LE under the null hypothesis. This is not available in the literature.

In order to tackle these problems, in this paper we propose a weighted linear estimator (WLE) and derive its limiting distribution. The limit distribution turns out to be multivariate normal even when some of the parameters are zero.

We next turn to the finite sample approximation of the distribution of the WLE. The bootstrap is an established method to obtain finite sample approximations of distributions of estimators. In a variety of cases, it has been shown that in many senses, it is better than the limiting normal approximation.

Here we consider appropriate bootstrapped version of the WLE, called BWLE. We prove that the bootstrap approximation is asymptotically valid and via simulations, we show that it yields better approximations than the normal.

2 Weighted LE and the Bootstrapped Weighted LE.

We first use a preliminary estimator of $\boldsymbol{\beta}$ to estimate the conditional variance in (1.1). We do this by using an appropriate weighted estimator. As shown in Lemma 3.1, this approach is useful to keep the moment restrictions to a minimum.

To motivate the preliminary estimator, let $Y_i = X_i^2$, $1-p \le i \le n$. For $1 \le i \le n$, let $\boldsymbol{Z}_{i-1} = [1, Y_{i-1}, \cdots, Y_{i-p}]' = [1, X_{i-1}^2, \cdots, X_{i-p}^2]'$ and $\eta_i = \epsilon_i^2 - 1$. Then

$$\sigma_{i-1}^2(\boldsymbol{\beta}) = \boldsymbol{Z}_{i-1}'\boldsymbol{\beta}. \tag{2.1}$$

Now squaring both sides of (1.1) and using the form of $\sigma_{i-1}^2(\boldsymbol{\beta})$ in (2.1), we get

$$Y_i = \boldsymbol{Z}_{i-1}'\boldsymbol{\beta} + \sigma_{i-1}^2(\boldsymbol{\beta})\eta_i, \qquad 1 \le i \le n, \tag{2.2}$$

where $E\{\sigma_{i-1}^2(\boldsymbol{\beta})\eta_i\} = E\{\sigma_{i-1}^2(\boldsymbol{\beta})\}E(\eta_i) = 0,\ 1 \le i \le n$.

Equation (2.2) reminds us of a standard linear autoregressive model with centered errors except that there is a multiplicative random scaling involved in the errors. Let $\boldsymbol{Z}$ be the matrix of order $n \times (1+p)$ with i-th row equal to $\boldsymbol{Z}_{i-1}'$ and $\boldsymbol{Y}$ be the vector with i-th entry Y_i, $1 \le i \le n$. Let $\{(u_{i-1}, v_{i-1}); i \ge 1\}$ be a sequence of nonnegative random variables (r.v.'s). They are typically called weights and let $\boldsymbol{U}$ and $\boldsymbol{V}$ be diagonal matrices of size n with the i-th diagonals given by u_{i-1} and v_{i-1} respectively, $1 \le i \le n$. In (2.2), ignoring the randomness of $\sigma_{i-1}^2(\boldsymbol{\beta})$ and also the presence of $\boldsymbol{\beta}$ in it, we obtain a preliminary weighted least squares estimator $\hat{\boldsymbol{\beta}}_{pr}$ of $\boldsymbol{\beta}$ as the solution of

$$\sum_{i=1}^{n} \left[u_{i-1}\boldsymbol{Z}_{i-1}\{Y_i - \boldsymbol{Z}_{i-1}'\boldsymbol{b}\}\right] = \boldsymbol{0}, \tag{2.3}$$

yielding the estimator

$$\hat{\boldsymbol{\beta}}_{pr} = (\boldsymbol{Z}'\boldsymbol{U}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{U}\boldsymbol{Y}.$$

In Section 3 we state the asymptotic normality of this estimator.

Based on $\hat{\boldsymbol{\beta}}_{pr}$, we may construct other estimators of $\boldsymbol{\beta}$ as follows. Notice that dividing (2.2) by $\sigma_{i-1}^2(\boldsymbol{\beta})$ and using (2.1), we obtain

$$Y_i/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}) = \{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})\}'\boldsymbol{\beta} + \eta_i.$$

Replacing $\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}$ by $\boldsymbol{Z}_{i-1}'\hat{\boldsymbol{\beta}}_{pr}$ in this expression, we obtain

$$Y_i/(\boldsymbol{Z}_{i-1}'\hat{\boldsymbol{\beta}}_{pr}) \approx \{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}'\hat{\boldsymbol{\beta}}_{pr})\}'\boldsymbol{\beta} + \eta_i. \tag{2.4}$$

So we may define a linear estimator (LE) of $\boldsymbol{\beta}$ as the solution of the equation

$$\sum_{i=1}^{n}\left[\left\{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}\hat{\boldsymbol{\beta}}_{pr})\right\}\left\{Y_i/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr}) - \{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr})\}'\boldsymbol{b}\right\}\right] = \boldsymbol{0}.$$

Likewise, a weighted linear estimate (WLE) $\hat{\boldsymbol{\beta}}_n$ may be defined as the solution of

$$\sum_{i=1}^{n} v_{i-1}\left[\left\{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}\hat{\boldsymbol{\beta}}_{pr})\right\}\left\{Y_i/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr}) - \{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr})\}'\boldsymbol{b}\right\}\right] = \boldsymbol{0}, \tag{2.5}$$

yielding the estimator

$$\hat{\boldsymbol{\beta}}_n = \left[\sum_{i=1}^{n} v_{i-1}\{\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr})^2\}\right]^{-1}\left[\sum_{i=1}^{n} v_{i-1}\{\boldsymbol{Z}_{i-1}Y_i/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})^2\}\right]. \tag{2.6}$$

Chatterjee and Bose (2005) developed the idea of weighted bootstrap versions of estimators that have been obtained as solutions of minimising problems or as solutions of equations in general dependent models. Though our situation does not quite fit in their set up, we are able to use their ideas and develop suitable bootstrap versions of the above estimators.

To describe this, let $\{w_{ni}; 1 \leq i \leq n, n \geq 1\}$ be a triangular array of r.v.'s such that for each $n \geq 1$, $\{w_{ni}; 1 \leq i \leq n\}$ are exchangeable and independent of $\{X_i; i \geq 1-p\}$ and $\{\epsilon_i, u_{i-1}, v_{i-1}; i \geq 1\}$. They are called the bootstrap weights. Let $\boldsymbol{W}$ be a diagonal matrix with i-th diagonal w_{ni}, $1 \leq i \leq n$.

From (2.3), the bootstrapped preliminary weighted least squares estimator $\hat{\boldsymbol{\beta}}_{Bpr}$ of $\boldsymbol{\beta}$ is defined as the solution of

$$\sum_{i=1}^{n} w_{ni}\left[u_{i-1}\boldsymbol{Z}_{i-1}\{Y_i - \boldsymbol{Z}'_{i-1}\boldsymbol{b}\}\right] = \boldsymbol{0}. \tag{2.7}$$

Hence,

$$\hat{\boldsymbol{\beta}}_{Bpr} = (\boldsymbol{Z}'\boldsymbol{W}\boldsymbol{U}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{W}\boldsymbol{U}\boldsymbol{Y}.$$

Finally, as in (2.4), the bootstrapped weighted linear estimator BWLE may be defined as the solution of

$$\sum_{i=1}^{n} w_{ni}v_{i-1}\left[\left\{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})\right\}\left\{Y_i/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr}) - \{\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})\}'\boldsymbol{b}\right\}\right] = \boldsymbol{0}, \tag{2.8}$$

which gives

$$\hat{\boldsymbol{\beta}}_B = \Big[\sum_{i=1}^{n} w_{ni} v_{i-1}\{\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})^2\}\Big]^{-1}\Big[\sum_{i=1}^{n} w_{ni} v_{i-1}\{\boldsymbol{Z}_{i-1}Y_i/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})^2\}\Big]. \tag{2.9}$$

Observe that while defining the bootstrap estimates, the structures of the estimators have been kept close to the orginal estimators, by following the same two step procedure and with the same weights as the original estimate. When $(w_{n1}, \ldots, w_{nn})$ follows a multinomial $(n, 1/n, \ldots, 1/n)$ distribution, we obtain the classical paired-bootstrap (see (2.7) and (2.8)). A host of other bootstrap methods in the literature like the Bayesian bootstrap, the deleted d-jackknives, the bootstrap clone are also special cases of the bootstrap formulation of (2.8). There are many advantages related to such general formulation of weighted bootstrap since it gives a unified way of studying several resampling schemes simultaneously. See, for example, Chatterjee and Bose (2005) and Praestgaard and Wellner (1993) for these details in other contexts.

3 Asymptotic distribution.

Throughout, we assume that the process $\{X_i; i \geq 1-p\}$ is stationary and ergodic. For necessary and sufficient conditions of stationarity, see Nelson (1990), Bougerol and Picard (1992) and Giraitis, Kokoszka and Leipus (2000). We also assume the following conditions on the errors and the weights.

(E1) $E(\epsilon^4) < \infty$.

Consider the increasing sequence of sigma-fields $\{F_{i-1} = \sigma < \boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_i >;\ i \geq 1\}$. Assume that

(W1) The weights $\{(u_{i-1}, v_{i-1}); i \geq 1\}$, are stationary and for each $i \geq 1$, (u_{i-1}, v_{i-1}) is F_{i-1}-measurable.

Model (1.1) together with (E1) and (W1) will be referred to as the **model assumptions.**

In the following $o_p(1)$ and $O_p(1)$ will denote the convergence in probability to zero and the boundedness in probability respectively with respect to the underlying stationary probability distribution.

Likewise, P_B, E_B, $o_B(1)$ and $O_B(1)$ will denote the probability, expectation, convergence in probability to zero and the boundedness in probability, respectively, with respect to the bootstrap distribution conditional on the data $\{X_i; 1-p \leq i \leq n\}$. Moreover, $\rightarrow$ will denote the convergence in distribution.

For the asymptotic normality of the preliminary estimator, we additionally assume the following for the preliminary weights $\{u_{i-1}\}$.

$$E(u_0^2) < \infty,\ E\{u_0 Y_{-j} Y_{-k}\} < \infty \text{ and } E\{u_0^2 Y_{-j} Y_{-k} Y_{-l} Y_{-m}\} < \infty,\ \forall 1 \leq j,\, k,\, l,\, m \leq p. \tag{3.10}$$

Condition (3.10) ensures that $E\{u_0(\boldsymbol{Z}_0\boldsymbol{Z}_0')\}$ and $E\{u_0^2(\boldsymbol{\beta}'\boldsymbol{Z}_0)^2\boldsymbol{Z}_0\boldsymbol{Z}_0'\}$ are all finite.

Lemma 3.1. *Suppose that the model assumptions and (3.10) hold. Then*

$$n^{1/2}(\hat{\boldsymbol{\beta}}_{pr} - \boldsymbol{\beta}) \rightarrow \boldsymbol{N}\left[\boldsymbol{0}, Var(\epsilon_1^2)\{E(u_0\boldsymbol{Z}_0\boldsymbol{Z}_0')\}^{-1}E\{u_0^2(\boldsymbol{\beta}'\boldsymbol{Z}_0)^2\boldsymbol{Z}_0\boldsymbol{Z}_0'\}\{E(u_0\boldsymbol{Z}_0\boldsymbol{Z}_0')\}^{-1}\right].$$

Let $\mathcal{P}_0 = \{0 < j < p;\ \beta_j = 0\}$. The next theorem concerns the asymptotic distribution of the WLE $\hat{\boldsymbol{\beta}}_n$. We assume the following conditions on the weights:

$$E(v_0^2) < \infty,\ E\{v_0 Y_{-j} Y_{-k}\} < \infty,\ E\{v_0^2 Y_{-j} Y_{-k}\} < \infty \text{ and } E\{v_0 Y_{-j} Y_{-k} Y_{-l}\} < \infty,\ \forall j, k, l \in \mathcal{P}_0. \tag{3.11}$$

Since

$$(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2} \leq \beta_0^{-2} \text{ and for } \beta_j > 0,\ y/(\beta_0 + \beta_j y) \tag{3.12}$$

is a bounded function of y, (3.11) ensures that

$$E\{v_0\boldsymbol{Z}_0\boldsymbol{Z}_0'(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2}\} < \infty \text{ and } E\{v_0^2\boldsymbol{Z}_0\boldsymbol{Z}_0'(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2}\} < \infty.$$

An example of a weight sequence satisfying (3.10) and (3.11) is given by

$$u_{i-1} = v_{i-1} = [(1 + Y_{i-1}) \ldots (1 + Y_{i-p})]^{-1}. \tag{3.13}$$

We will need to impose only the following property of $\hat{\boldsymbol{\beta}}_{pr}$:

$$n^{1/2}(\hat{\boldsymbol{\beta}}_{pr} - \boldsymbol{\beta}) = O_p(1). \tag{3.14}$$

Theorem 3.1. *Suppose that the model assumptions, (3.11) and (3.14) hold. Then,*

$$\begin{aligned} & n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \\ = \ & \Big[E\{v_0 \boldsymbol{Z}_0 \boldsymbol{Z}_0'/(\boldsymbol{Z}_0'\boldsymbol{\beta})^2\}\Big]^{-1} \Big[n^{-1/2}\sum_{i=1}^{n} \eta_i v_{i-1} \boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})\Big] + o_p(1), \end{aligned} \tag{3.15}$$

which converges in distribution to

$$\boldsymbol{N}\Bigg[0, Var(\epsilon_1^2)\Big[E\{v_0 \boldsymbol{Z}_0 \boldsymbol{Z}_0'(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2}\}\Big]^{-1}\Big[E\{v_0^2 \boldsymbol{Z}_0 \boldsymbol{Z}_0'(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2}\}\Big]\Big[E\{v_0 \boldsymbol{Z}_0 \boldsymbol{Z}_0'(\boldsymbol{\beta}'\boldsymbol{Z}_0)^{-2}\}\Big]^{-1}\Bigg].$$

Remark 3.1. Weiss (1986) and Bose and Mukherjee (2003) assumed that all β_j's are strictly positive in (1.1). Bose and Mukherjee (2003) obtained the limiting distribution of $\hat{\boldsymbol{\beta}}_n$ under the condition that for all $1 \le j,\ k,\ l \le p$,

$$E\{Y_{-j}Y_{-k}Y_{-l}/(\boldsymbol{\beta}'\boldsymbol{Z}_0)^3\} < \infty,$$

whereas Weiss (1986) assumed that for all $1 \le j,\ k \le p$, $E\{Y_{-j}Y_{-k}/(\boldsymbol{\beta}'\boldsymbol{Z}_0)^2\} < \infty$.

These conditions are automatically satisfied when $\beta_j > 0\ \forall\ 1 \le j \le p$. However, in our case, $\mathcal{P}_0$ could be nonempty. One approach to tackle this situation is to assume existence of higher moments for $\{X_i\}$. The introduction of the $\{v_{i-1}\}$'s avoids this since they down weight $\{X_i\}$'s in (3.11). See Ling (2005) for a related idea for the estimation in an infinite variance autoregressive model.

Finally, we discuss the bootstrapped estimators. Recall from Section 2 that $\forall n \ge 1$,

$$\{w_{ni}; 1 \le i \le n\} \text{ are exchangeable and independent of } \{\boldsymbol{Z}_{i-1}, \epsilon_i, u_{i-1}, v_{i-1}, 1 \le i \le n\}. \tag{3.16}$$

In addition, they are assumed to satisfy the following basic conditions (Conditions BW of Chatterjee and Bose (2005)) where $\sigma_n^2 = V_B(w_{ni})$ and $k_1 > 0$ is a constant.

$$E_B(w_{ni}) = 1,\ 0 < k_1 < \sigma_n^2 = o(n),\ \text{and}\ \forall\ 1 \le i \ne j \le n,\ \text{corr}_B\,(w_{ni}, w_{nj}) = O(1/n). \tag{3.17}$$

The next lemma establishes the boundedness in probability of the appropriately scaled bootstrapped preliminary estimator.

Lemma 3.2. *Suppose that the model assumptions, (3.16) and (3.17) hold. If for some* $\delta > 2$,

$$E\{(u_0 Y_{-j} Y_{-k})^{\delta}\} < \infty, \forall j, k,$$

then

$$\sigma_n^{-1} n^{1/2} (\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) = O_B(1).$$

To derive the bootstrap consistency of $\hat{\boldsymbol{\beta}}_B$, define for $\boldsymbol{x} \in \mathcal{R}^{1+p}$

$$F_n(\boldsymbol{x}) = P[n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \leq \boldsymbol{x}] \text{ and } F_B(\boldsymbol{x}) = P_B[\sigma_n^{-1} n^{1/2}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n) \leq \boldsymbol{x}].$$

We assume the following conditions (Conditions CLTW of Chatterjee and Bose (2005)) on the standardised exchangeable bootstrap weights $\{W_{ni} := (w_{ni} - 1)/\sigma_n\}$.

$$E_B(W_{ni}^4) < \infty, \text{ and } \forall\ 1 \leq i \neq j \leq n, \lim_{n \to \infty} E_B(W_{ni}^2 W_{nj}^2) = 1. \tag{3.18}$$

We also assume the following conditions on the weights of the original estimator $\hat{\boldsymbol{\beta}}_n$. For some $\delta > 2$,

$$\begin{aligned} & E(v_0^2) < \infty,\ E\{(v_0 Y_{-j})^{\delta}\} < \infty,\ E\{(v_0 Y_{-j} Y_{-k})^{\delta}\} < \infty,\ E\{v_0^2 Y_{-j} Y_{-k}\} < \infty \\ & \text{and} \qquad E\{v_0 Y_{-j} Y_{-k} Y_{-l}\} < \infty,\ \forall j, k, l \in \mathcal{P}_0. \end{aligned} \tag{3.19}$$

Finally, we assume only the following for $\hat{\boldsymbol{\beta}}_{Bpr}$ which is a consequence of Lemma 3.2.

$$\sigma_n^{-1} n^{1/2} (\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) = O_B(1). \tag{3.20}$$

Theorem 3.2. *Suppose that the model assumptions, (3.16)-(3.19) and (3.20) hold. Then,*

$$\begin{aligned} & \sigma_n^{-1} n^{1/2} (\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n) \\ & \quad = \Big[E\{v_0 \boldsymbol{Z}_0 \boldsymbol{Z}_0' / (\boldsymbol{Z}_0' \boldsymbol{\beta})^2\} \Big]^{-1} \times n^{-1/2} \sum_{i=1}^{n} \eta_i W_i v_{i-1} \boldsymbol{Z}_{i-1} / (\boldsymbol{Z}_{i-1}' \boldsymbol{\beta}) + o_B(1), \end{aligned} \tag{3.21}$$

and hence under the assumptions of Theorem 3.1

$$\sup\{|F_B(\boldsymbol{x}) - F_n(\boldsymbol{x})|; \boldsymbol{x} \in \mathcal{R}^{1+p}\} = o_p(1). \tag{3.22}$$

Remark 3.2. The proof of the first part of the theorem is given later. It may be mentioned that once the first part is established, the second part follows by using the central limit theorem for weighted sums of exchangeable random variables. This argument is now quite standard in the literature. In the present case, the stationarity and ergodicity of the process helps to verify the conditions needed for applicability of this central limit theorem.

4 Simulation study.

In this section we report results of some simulation study. We consider the model (1.1) with (i) $\boldsymbol{\beta} = [0.2, 0.0, 0.1]'$ and (ii) $\boldsymbol{\beta} = [0.1, 0.0, 0.2]'$ for the illustrative purpose when the errors $\{\epsilon_i\}$'s are standard normal. We are interested in approximating the distribution of $n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$ for $n = 30$. Here $\mathcal{P}_0 = \{1\}$ and among many different choices of the weight functions we consider $u_{i-1} = v_{i-1} = 1/Y_{i-1}$. The distribution of $n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$ is simulated based on $K = 1000$ replications. The three histograms based on 1000 of $n^{1/2}(\hat{\beta}_{nj} - \beta_j)$, $0 \leq j \leq p = 2$ represent marginal distributions of $n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$ which we want to approximate via bootstrap. In addition, we compute the means and the average of the squares of the three sets of $K = 1000$ numbers; they represent the means and the mean squared errors of $n^{1/2}(\hat{\beta}_{nj} - \beta_j)$, $0 \leq j \leq p = 2$.

From many different choices of bootstrap weights, we consider $(w_{n1}, \ldots, w_{nn})$ to have a multinomial $(n, 1/n, \ldots, 1/n)$ distribution with $n = 30$. This represents the classical paired bootstrap. Next we simulate the distribution of $\sigma_n^{-1} n^{1/2}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n)$ as follows. We choose and fix $\hat{\boldsymbol{\beta}}_n$ computed from the first of the $K = 1000$ iterations; let it be denoted by $\hat{\boldsymbol{\beta}}_{n(1)}$. We then generate $B = 1000$ bootstrap samples based on the exchangeable weights $\{w_{ni}\}$'s. For the j-th sample, $1 \leq j \leq B = 1000$, we compute $\sigma_n^{-1} n^{1/2}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_{n(1)})$. The three histograms, each based on B observations represent marginal distribution of $\sigma_n^{-1} n^{1/2}(\hat{\beta}_{Bj} - \beta_{nj(1)})$, $0 \leq j \leq p = 2$. In addition, we compute the means and the average of the squares of the three sets of $B = 1000$ numbers; they represent the means and the mean squared errors of $\sigma_n^{-1} n^{1/2}(\hat{\beta}_{Bj} - \beta_{nj(1)})$, $0 \leq j \leq p = 2$. The results are tabulated as follows.

Table 1 : Means and the MSE's of the standardised estimators and its bootstrap version.

	Mean	Bootstrap mean	MSE	Bootstrap MSE
$\beta_0 = 0.1$	0.030450802	-0.027249987	0.067533492	0.031403458
$\beta_1 = 0.0$	-0.052305343	0.002382498	1.185876441	0.915267353
$\beta_2 = 0.2$	-0.459559789	0.125414934	5.158004007	1.522143784

	Mean	Bootstrap mean	MSE	Bootstrap MSE
$\beta_0 = 0.2$	0.023349895	0.078807571	0.207721701	0.169025667
$\beta_1 = 0.0$	0.002857777	-0.091139316	1.243702078	0.305486087
$\beta_2 = 0.1$	-0.216279962	-0.142388595	1.425564161	2.801528557

From the table, the values of the standardised means are quite close to zero for the first two parameters while they are away from zero for the last parameter. In fact, for this case, the bootstrap means are closer to zero in absolute value than the original standardised estimator. Similar comments can be made regarding the MSE's. The histograms of the standardised bootstrap distribution are seen to capture the main feature like skewness etc of the standardised $\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}$ very well.

5 Proofs.

In this section we give proofs of the results of Section 3. Here for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of same order, $\boldsymbol{A} \leq \boldsymbol{B}$ means element wise $a_{ij} \leq b_{ij}$. By $|\boldsymbol{A}|$, we mean the largest absolute element of $\boldsymbol{A}$. To simplify writing, we use w_i and W_i for w_{ni} and W_{ni} respectively.

Proof of Lemma 3.1. The proof is similar to the Lemma of Bose and Mukherjee (2003) and hence omitted.

Proof of Lemma 3.2. Note that

$$\sigma_n^{-1} n^{1/2} (\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) = (n^{-1} \boldsymbol{Z}' \boldsymbol{W} \boldsymbol{U} \boldsymbol{Z})^{-1} (\sigma_n n^{1/2})^{-1} \boldsymbol{Z}' \boldsymbol{W} \boldsymbol{U} (\boldsymbol{Y} - \boldsymbol{Z} \boldsymbol{\beta}).$$

Now

$$n^{-1} \boldsymbol{Z}' \boldsymbol{W} \boldsymbol{U} \boldsymbol{Z} = n^{-1} \sum_{i=1}^{n} u_{i-1} \boldsymbol{Z}_{i-1} \boldsymbol{Z}'_{i-1} + n^{-1} \sum_{i=1}^{n} (w_i - 1) u_{i-1} \boldsymbol{Z}_{i-1} \boldsymbol{Z}'_{i-1}.$$

Note that $E_B\{n^{-1}\sum_{i=1}^n(w_i-1)v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\}=0$. Using (3.17), the bootstrap variance of each element of the random matrix $n^{-1}\sum_{i=1}^n(w_i-1)v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}$ goes to zero. Hence

$$n^{-1}\boldsymbol{Z}'\boldsymbol{W}\boldsymbol{U}\boldsymbol{Z}=n^{-1}\sum_{i=1}^n u_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}+o_B(1)=O_p(1)+o_B(1).$$

Next, using (2.1) and (2.2), write $(\sigma_n n^{1/2})^{-1}\boldsymbol{Z}'\boldsymbol{W}\boldsymbol{U}(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\beta})$ as

$$\begin{aligned}&(\sigma_n n^{1/2})^{-1}\boldsymbol{Z}'\boldsymbol{W}\boldsymbol{U}(\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\beta})\\ =\ &(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n u_{i-1}\eta_i\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}+n^{-1/2}\sum_{i=1}^n W_i\{\eta_i u_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}\}.\end{aligned}$$

Note that $\sigma_n^{-1}=O(1)$ and using the martingale CLT $n^{-1/2}\sum_{i=1}^n v_{i-1}\eta_i\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}=O_p(1)$. Also, using a proof similar to (5.2) and (5.4) below, $n^{-1/2}\sum_{i=1}^n W_i\{\eta_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}\}=O_B(1)$ under the assumed condition. Hence the proof follows.

For the next lemma, define

$$\boldsymbol{U}_B=n^{-1/2}\sum_{i=1}^n W_i\Big\{v_{i-1}(\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1})/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2-E[v_0(\boldsymbol{Z}_0\boldsymbol{Z}'_0)/(\boldsymbol{Z}'_0\boldsymbol{\beta})^2]\Big\}$$

and

$$\boldsymbol{V}_B=n^{-1/2}\sum_{i=1}^n W_i\eta_i v_{i-1}\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}).$$

Lemma 5.1. *Suppose that the model assumptions hold.*

(i) If for some $\delta>2$

$$E\{v_0Y_{-j}Y_{-k}\}^\delta<\infty,\ \forall j,\ k,\in\mathcal{P}_0, \tag{5.1}$$

then

$$\boldsymbol{U}_B=O_B(1). \tag{5.2}$$

(ii) If for some $\delta>2$

$$E\{v_0Y_{-j}\}^\delta<\infty,\ \forall j\in\mathcal{P}_0, \tag{5.3}$$

then

$$\boldsymbol{V}_B=O_B(1). \tag{5.4}$$

Proof. For the proof, we verify conditions of Lemma 4.6 of Praestgaard and Wellner (1993) for the asymptotic normality of a weighted sum of the exchangeable r.v.'s $\{W_i; 1\le i\le n\}$.

Since $\boldsymbol{U}_B$ is a random matrix, for simplicity, we can show the asymptotic normality of the random vector

$$n^{-1/2}\sum_{i=1}^{n} W_i\Big\{v_{i-1}\boldsymbol{Z}_{i-1}Y_{i-1-j}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2 - E[v_0\boldsymbol{Z}_0Y_{-j}/(\boldsymbol{Z}'_0\boldsymbol{\beta})^2]\Big\},$$

where $1 \le j \le p$ is fixed, by applying the Cramer-Wold device on it; hence the 'constant weights' (with respect to the bootstrap distribution) are given by

$$a_i = v_{i-1}Y_{i-1-j}(\boldsymbol{Z}'_{i-1}l)/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2 - E[v_0Y_{-j}(\boldsymbol{Z}'_0l)/(\boldsymbol{Z}'_0\boldsymbol{\beta})^2],$$

where $\boldsymbol{l} \in \mathcal{R}^{p+1}$. Most of the conditions on the weights can be verified easily using the stationarity and ergodicity of $\{a_i\}$ and the existence δ-th moment. For example, under (5.1), for $\epsilon > 0$,

$$\begin{aligned} P[\max\{|a_i|; 1 \le i \le n\} > n^{1/2}\epsilon] &\le \sum_{i=1}^{n} P[|a_i| > n^{1/2}\epsilon] \\ = \quad nP[|a_1| > n^{1/2}\epsilon] &\le n\epsilon^{-\delta}n^{-\delta/2}E|a_1|^{\delta} = o(1), \end{aligned}$$

and hence condition (4.21) of Praestgaard and Wellner (1993) is satisfied. The conditions on $\{W_i\}$'s can be verified using assumption (3.18) and hence (5.2) holds. Similarly, (5.4) can be proved under (5.3)

For the proof of Theorem 3.2, note that (3.19) implies (3.11), (5.1) and (5.3). For the ease of writing, the following notations are used.

$$\begin{aligned} a_i &= \eta_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2, & \hat{a}_i &= \eta_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr})^2, \\ b_i &= \eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2, & \hat{b}_i &= \eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{\beta}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})^2, \\ A_i &= v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2, & \hat{A}_i &= v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{pr})^2 \\ B_i &= w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2, & \hat{B}_i &= w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\hat{\boldsymbol{\beta}}_{Bpr})^2. \end{aligned}$$

With this notation we get from (2.6) and (2.9),

$$n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = (n^{-1}\sum_{i=1}^{n}\hat{A}_i)^{-1}n^{-1/2}\sum_{i=1}^{n}\hat{a}_i, \tag{5.5}$$

$$\sigma_n^{-1}n^{1/2}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n) = (n^{-1}\sum_{i=1}^{n}\hat{B}_i)^{-1}(\sigma_n n^{1/2})^{-1}\sum_{i=1}^{n}\hat{b}_i. \tag{5.6}$$

$$n^{-1}\sum_{i=1}^{n} B_i = (\sigma_n/n^{1/2})\boldsymbol{U}_B + E[v_0(\boldsymbol{Z}_0\boldsymbol{Z}_0')/(\boldsymbol{Z}_0'\boldsymbol{\beta})^2]n^{-1}\sum_{i=1}^{n}(w_i-1) + n^{-1}\sum_{i=1}^{n} A_i, \tag{5.7}$$

and

$$(\sigma_n n^{1/2})^{-1}\sum_{i=1}^{n} b_i = \boldsymbol{V}_B + (\sigma_n n^{1/2})^{-1}\sum_{i=1}^{n} a_i. \tag{5.8}$$

We will also use the mean value theorem where for $u,\ v > 0$,

$$1/u^2 - 1/v^2 = -2(u-v)/\chi^3, \tag{5.9}$$

with χ being the intermediate point and the two-step Taylor's formula where for $u,\ v > 0$,

$$1/u^2 - 1/v^2 = -2(u-v)/v^3 + 3(u-v)^2/\xi^4, \tag{5.10}$$

with ξ being the intermediate point. Note that the intermediate point, say ξ, satisfies $0 < 1/\xi^c \leq (1/v^c)\{1+(v/u)^c\}$ for any fixed $c > 0$ chosen to facilitate our manipulation.

From Lemma 2 of Mukherjee (2006), if $\boldsymbol{U} = [u_1,\ldots,u_k]'$, $\boldsymbol{V} = [v_1,\ldots,v_k]'$ and $\boldsymbol{W}$ are vectors with all entries nonnegative then

$$\boldsymbol{W}'\boldsymbol{V}/\boldsymbol{W}'\boldsymbol{U} \leq (v_1/u_1) + \cdots (v_k/u_k), \tag{5.11}$$

where, by definition, $v_j/u_j = 0$ if $u_j = 0 = v_j$. Therefore, when (5.9) or (5.10) is used with $u = \hat{\boldsymbol{\beta}}_{Bpr}'\boldsymbol{Z}_{i-1}$ and $v = \boldsymbol{\beta}'\boldsymbol{Z}_{i-1}$, then by (5.11), the intermediate points ξ_{ni} satisfy

$$\begin{aligned} 1/\xi_{ni}^c &\leq (1/\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^c[1+\{\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}/\boldsymbol{Z}_{i-1}'\hat{\boldsymbol{\beta}}_{Bpr}\}^c] \\ &\leq (1/\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^c[1+\{(\beta_0/\hat{\beta}_{0\,Bpr}) + \cdots + (\beta_p/\hat{\beta}_{p\,Bpr})\}^c], \end{aligned} \tag{5.12}$$

where $\hat{\beta}_{j\,Bpr}$ is the j-th entry of $\hat{\boldsymbol{\beta}}_{Bpr}$, $0 \leq j \leq p$ and by the assumption (3.20), $1 + \{(\beta_0/\hat{\beta}_{0\,Bpr}) + \cdots + (\beta_p/\hat{\beta}_{p\,Bpr})\}^c = O_B(1)$.

Lemma 5.2. *Suppose that the model assumptions, (3.14) and (3.20) hold. If, in addition,*

$$E(v_0) < \infty,\ E\{v_0 Y_{-j}Y_{-k}\} < \infty,\ \textit{and}\ E\{v_0 Y_{-j}Y_{-k}Y_{-l}\} < \infty,\ \forall j,\ k,\ l \in \mathcal{P}_0, \tag{5.13}$$

then

$$n^{-1}\{\sum_{i=1}^{n}\hat{A}_i - \sum_{i=1}^{n} A_i\} = o_p(1), \tag{5.14}$$

$$n^{-1/2}\{\sum_{i=1}^{n}\hat{a}_i - \sum_{i=1}^{n} a_i\} = o_p(1), \tag{5.15}$$

$$n^{-1}\{\sum_{i=1}^{n}\hat{B}_i - \sum_{i=1}^{n} B_i\} = o_B(1), \tag{5.16}$$

and

$$(\sigma_n n^{1/2})^{-1}\{\sum_{i=1}^{n}\hat{b}_i - \sum_{i=1}^{n} b_i\} = -2(\boldsymbol{V}_B^* + \boldsymbol{T}_n)n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) + 3(\sigma_n n^{1/2})^{-1}\boldsymbol{S}_B, \tag{5.17}$$

where

$$\boldsymbol{V}_B^* = n^{-1}\sum_{i=1}^{n} W_i\eta_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2 = o_B(1),$$

$$\boldsymbol{T}_n = (\sigma_n n)^{-1}\sum_{i=1}^{n}\eta_i v_{i-1}(\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1})/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^2 = o_p(1)$$

and $\boldsymbol{S}_B = O_B(1)$.

Proof. For (5.14) with factor n^{-1} we use (5.9) whereas for (5.15) with factor $n^{-1/2}$, we use the two-step expansion (5.10) with $u = \hat{\boldsymbol{\beta}}'_{pr}\boldsymbol{Z}_{i-1}$ and $v = \boldsymbol{\beta}'\boldsymbol{Z}_{i-1}$ as in (5.12); the details are similar to Bose and Mukherjee (2003) and hence omitted. In a similar fashion, for (5.16) with factor n^{-1} we use (5.9) to get

$$\begin{aligned} & n^{-1}\sum_{i=1}^{n}\hat{B}_i - n^{-1}\sum_{i=1}^{n} B_i \\ = \ & -2n^{-1/2}n^{-1}\sum_{i=1}^{n} w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\{\boldsymbol{Z}'_{i-1}n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta})\}/\chi_{ni}^3. \end{aligned}$$

Writing $n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta}) = [b_{B0}, \cdots, b_{Bp}]'$, one can take the elements of $n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})$ outside the sum with respect to 'i'. Note that $b_{Bj} = O_B(1)$ and using (5.12) with $c = 3$ for the intermediate points, $1/\chi_{ni}^3 \leq (\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^{-3}O_B(1)$. Moreover, entry wise,

$$0 \leq n^{-1}\sum_{i=1}^{n} w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}'_{i-1}\boldsymbol{Z}'_{i-1}/(\boldsymbol{Z}'_{i-1}\boldsymbol{\beta})^3 = O_B(1)$$

since its expectation equals $E_B(w_1)Ev_0\boldsymbol{Z}_0\boldsymbol{Z}'_0\boldsymbol{Z}'_0/(\boldsymbol{Z}'_0\boldsymbol{\beta})^3$ which is finite under (5.13). Hence (5.16) follows. Next using (5.10)

$$(\sigma_n n^{1/2})^{-1}\{\sum_{i=1}^{n}\hat{b}_i - \sum_{i=1}^{n} b_i\}$$

$$\begin{aligned}
&= -2(\sigma_n n)^{-1}\sum_{i=1}^{n}\eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\{\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}\}\{\boldsymbol{Z}_{i-1}' n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})\}/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^3\\
&+ 3(\sigma_n n^{3/2})^{-1}\sum_{i=1}^{n}\eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\{\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}\}\{\boldsymbol{Z}_{i-1}' n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})\}^2/\xi_{ni}^4\\
&= -2(\sigma_n n)^{-1}\sum_{i=1}^{n}\{\eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\boldsymbol{Z}_{i-1}'/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^2\} n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})\\
&+ 3(\sigma_n n^{3/2})^{-1}\sum_{i=1}^{n}\eta_i w_i v_{i-1}\boldsymbol{Z}_{i-1}\{\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}\}\{\boldsymbol{Z}_{i-1}' n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})\}^2/\xi_{ni}^4\\
&= -2(\boldsymbol{V}_B^* + \boldsymbol{T}_n) n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta}) + 3(\sigma_n n^{1/2})^{-1}\boldsymbol{S}_B, \text{ say.}
\end{aligned}$$

It is easy to see that $\boldsymbol{V}_B^* = O_B(1)$ and $\boldsymbol{T}_n = o_p(1)$. To show $\boldsymbol{S}_B = O_B(1)$, first we take the elements of $n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr}-\boldsymbol{\beta})$ outside the sum with respect to i. Then using (5.12) with $c = 4$, we can get one $(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^{-4}O_B(1)$ factor from ξ_{ni}^{-4} which after cancellation with $\{\boldsymbol{Z}_{i-1}'\boldsymbol{\beta}\}$ of the numerator basically gives

$$|\boldsymbol{S}_B| \le n^{-1}\sum_{i=1}^{n}|\eta_i| w_i v_{i-1}|\boldsymbol{Z}_{i-1}\boldsymbol{Z}_{i-1}'\boldsymbol{Z}_{i-1}|/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})^{-3}\times O_B(1) = O_p(1)\times O_B(1),$$

which is $O_B(1)$ under (5.13).

Proof of Theorem 3.1. From (5.5) and using (5.14) and (5.15),

$$\begin{aligned}
n^{1/2}(\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}) &= (n^{-1}\sum_{i=1}^{n}\hat{A}_i)^{-1} n^{-1/2}\sum_{i=1}^{n}\hat{a}_i\\
&= (n^{-1}\sum_{i=1}^{n}A_i)^{-1} n^{-1/2}\sum_{i=1}^{n}a_i + o_p(1)\\
&= \left[E\{v_0\boldsymbol{Z}_0\boldsymbol{Z}_0'/(\boldsymbol{Z}_0'\boldsymbol{\beta})^2\}\right]^{-1}\left[n^{-1/2}\sum_{i=1}^{n}\eta_i v_{i-1}\boldsymbol{Z}_{i-1}/(\boldsymbol{Z}_{i-1}'\boldsymbol{\beta})\right] + o_p(1)
\end{aligned} \tag{5.18}$$

and hence by the martingale CLT, the result follows.

Proof of Theorem 3.2. From (5.5) and (5.6)

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n &= (\hat{\boldsymbol{\beta}}_B-\boldsymbol{\beta}) - (\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta})\\
&= [\sum_{i=1}^{n}\hat{B}_i]^{-1}[\sum_{i=1}^{n}\hat{b}_i] - [\sum_{i=1}^{n}\hat{A}_i]^{-1}[\sum_{i=1}^{n}\hat{a}_i]
\end{aligned}$$

$$= \Big\{[\sum_{i=1}^n \hat{B}_i]^{-1}[\sum_{i=1}^n \hat{b}_i] - [\sum_{i=1}^n B_i]^{-1}[\sum_{i=1}^n b_i]\Big\} - \Big\{[\sum_{i=1}^n \hat{A}_i]^{-1}[\sum_{i=1}^n \hat{a}_i] - [\sum_{i=1}^n A_i]^{-1}[\sum_{i=1}^n a_i]\Big\}$$
$$+\Big\{[\sum_{i=1}^n B_i]^{-1}[\sum_{i=1}^n b_i] - [\sum_{i=1}^n A_i]^{-1}[\sum_{i=1}^n a_i]\Big\}.$$

Recall that under (3.19), (3.11), (5.1) and (5.3) hold. Next using Lemma 5.2, (5.7) and (5.8),

$$\begin{aligned}
&\sigma_n^{-1}n^{1/2}(\hat{\boldsymbol{\beta}}_B - \hat{\boldsymbol{\beta}}_n)\\
&= \Big\{[n^{-1}\sum_{i=1}^n \hat{B}_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n \hat{b}_i] - [n^{-1}\sum_{i=1}^n B_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i]\Big\}\\
&- \Big\{[n^{-1}\sum_{i=1}^n \hat{A}_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n \hat{a}_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i]\Big\}\\
&+ \Big\{[n^{-1}\sum_{i=1}^n B_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i]\Big\}\\
&= \Big\{[n^{-1}\sum_{i=1}^n B_i + o_B(1)]^{-1}[-2(\boldsymbol{V}_B^* + \boldsymbol{T}_n)n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) + 3(\sigma_n n^{1/2})^{-1}\boldsymbol{S}_B\\
&\quad +(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i] - [n^{-1}\sum_{i=1}^n B_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i]\Big\}\\
&- \Big\{[n^{-1}\sum_{i=1}^n \hat{A}_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n \hat{a}_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i]\Big\}\\
&+ \Big\{(\sigma_n/n^{1/2})\boldsymbol{U}_B + E[v_0(\boldsymbol{Z}_0\boldsymbol{Z}_0')/(\boldsymbol{Z}_0'\boldsymbol{\beta})^2]n^{-1}\sum_{i=1}^n(w_i-1) + n^{-1}\sum_{i=1}^n A_i]^{-1}\\
&\quad \times[\boldsymbol{V}_B + (\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i]\Big\}.
\end{aligned}$$

Since $\boldsymbol{V}_B^* = O_B(1)$ and $\boldsymbol{T}_n = o_p(1)$,

$$[n^{-1}\sum_{i=1}^n B_i + o_B(1)]^{-1}[-2(\boldsymbol{V}_B^* + \boldsymbol{T}_n)n^{1/2}(\hat{\boldsymbol{\beta}}_{Bpr} - \boldsymbol{\beta}) + 3(\sigma_n n^{1/2})^{-1}\boldsymbol{S}_B + (\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i]$$
$$-[n^{-1}\sum_{i=1}^n B_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n b_i] = o_B(1).$$

Using (5.14) and (5.15)

$$[n^{-1}\sum_{i=1}^n \hat{A}_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n \hat{a}_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i] = o_p(1).$$

From (3.17), $E_B[n^{-1}\sum_{i=1}^n (w_i - 1)]^2 = o(1)$. Also $\sigma_n/n^{1/2} = o(1)$. Since $n^{-1}\sum_{i=1}^n A_i = O_p(1)$ and $n^{-1/2}\sum_{i=1}^n a_i = O_p(1)$, it follows from Lemma 5.1,

$$\begin{aligned} &[(\sigma_n/n^{1/2})\boldsymbol{U}_B + E[v_0(\boldsymbol{Z}_0\boldsymbol{Z}_0')/(\boldsymbol{Z}_0'\boldsymbol{\beta})^2]n^{-1}\sum_{i=1}^n (w_i - 1) + n^{-1}\sum_{i=1}^n A_i]^{-1} \\ &\quad \times[\boldsymbol{V}_B + (\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i] - [n^{-1}\sum_{i=1}^n A_i]^{-1}[(\sigma_n n^{1/2})^{-1}\sum_{i=1}^n a_i] \\ = &\quad [n^{-1}\sum_{i=1}^n A_i]^{-1}\boldsymbol{V}_B + o_B(1). \end{aligned}$$

Therefore (3.21) follows. Also, (3.22) follows from (3.15) and (3.21) as in the proof of (5.2).

References

[1] Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). ARCH modelling in finance; a review of the theory and empirical evidence. *J. Econometrics,* **52**, 5-59.

[2] Bose, A. and Mukherjee, K. (2003). Estimating the ARCH parameters by solving linear equations. *Journal of Time Series Analysis*, 24, 127-136.

[3] Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *J. of Econometrics,* **52**, 115-127.

[4] Chatterjee, S. and Bose, A. (2005). Generalized bootstrap for estimating equations. *Annals of Statistics,* **33**, 414-436.

[5] Engle, R.F. (1982). Autoregressive conditional heteroskedasticity and estimates of the variance of UK inflation. *Econometrica,* **50**, 987-1008.

[6] Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: dependence structure and central limit theorem. *Econometric theory,* **16**, 3-22.

[7] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Applications.* Academic Press, New York.

[8] Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *JRSS, Ser B,* **67**, 381-393.

[9] Mukherjee, K. (2006). Pseudo-likelihood estimation in ARCH models. To appear in *Canadian Journal of Statistics.*

[10] Nelson, D. B. (1990). Stationarity and persistence in the GARCH (1, 1) model. *Econometric Theory,* **6**, 318-334.

[11] Praestgaard, J. and Wellner, J. (1993). Exchangeably weighted bootstrap of the general empirical process. *Annals of Probability*, **21**, 2053-2086.

[12] Weiss, A. A. (1986). Asymptotic Theory for ARCH models: estimation and testing. *Econometric Theory*, **2**, 07-131.