이학석사학위논문

A Review Study on Asymptotic Normality and Parameter Change Test for Zero-inflated General Integer-valued GARCH Models

영과잉 일반 정수값 GARCH 모형의 근사 정규성과 변화점 탐지 검정의 재검토 연구

2019 년 2 월

서울대학교 대학원

통계학과

석 성 우

이학석사학위논문

A Review Study on Asymptotic Normality and Parameter Change Test for Zero-inflated General Integer-valued GARCH Models

영과잉 일반 정수값 GARCH 모형의 근사 정규성과 변화점 탐지 검정의 재검토 연구

2019 년 2 월

서울대학교 대학원

통계학과

석 성 우

A Review Study on Asymptotic Normality and Parameter Change Test for Zero-inflated General Integer-valued GARCH Models

영과잉 일반 정수값 GARCH 모형의 근사 정규성과 변화점 탐지 검정의 재검토 연구

지도교수 이 상 열

이 논문을 이학석사 학위논문으로 제출함

2018 년 12 월

서울대학교 대학원

통계학과

석 성 우

석성우의 이학석사 학위논문을 인준함

2018 년 12 월

위 원 장	이영조
부위원장	이상열
위 원	원중호

Abstract

This review study considers the problem of testing a parameter change in zero-inflated general integer-valued time series models where the conditional density of current observations is assumed to follow a zero-inflated one-parameter exponential family. This thesis focuses on the standardized residual-based CUSUM tests, based on the previous study of Lee and Lee (2018) and show that their null distributions converge weakly to the functions of Brownian bridges.

Keywords: Time series of counts, integer-valued GARCH models, zero-inflated exponential family, parameter change test, CUSUM test.

Student Number: 2017-26674

Contents

Chapter 1

Introduction

Integer-valued time series models have been studied by many researchers and applied to many applications in science, engineering, and economics. Integer-valued autoregressivee (INAR) time series models based on a binomial thinning operation are introduced by the authors such as McKenzie (1985, 2003), Alzaid and Al-Osh (1990), Al-Osh and Aly (1992). See Weiß(2008). Other models such as nonlinear integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models are also used by Heinen (2003), Ferland et al. (2006), Fokianos et al. (2009), and Neumann (2011). Among discrete distributions, the Poisson distribution has been widely used as the conditional distribution of current observations given past information but other distributions are also considered. See Davis and Wu (2009), Zhu (2011), and Christou and Fokianos (2014) who consider negative binomial INGARCH (NB-INGARCH) models, and also Zhu (2012a,b) and Lee et al. (2016) who consider zero-inated generalized Poisson and negative binomial INGARCH models. Davis and Liu (2016) consider nonlinear INGARCH models with conditional densities belonging to one-parameter exponential family. Lee and Lee (2018) recently studied the parameter change test in their models.

The change point problem has attracted much attention from researchers during the past decades since many time series often experience structural changes in their underlying models, see Csörgö and Horváth (1997) for a general review and Lee et al. (2003) for a background. The change point test for integer-valued time series has been studied by Kang and Lee (2009), Fokianos and Fried (2010, 2012), Franke et al. (2012), Fokianos et al. (2014), Kang and Lee (2014), Hudecová et al. (2016), and Diop and Kengue (2017). The CUSUM test performs well in many situations, but the estimate-based CUSUM test suffers from severe size distortions in GARCH models, see Kang and Lee (2014) and Lee et al. (2016). As a remedy, the residual-based CUSUM test has been proposed, see Lee et al. (2004) and Lee and Lee (2015). However, its performance is poor particularly when a parameter change locates in conditional mean part, see Oh and Lee (2018). Lee and Lee (2018) proposed to use the score vector-based CUSUM test and standardized residual-based CUSUM test, and Lee, Seok and Kim (2018) has extended Lee and Lee (2018) to the zero-inflated exponential family INGARCH models. For zero-inflated integer-valued models, we refer to Jazi and Lee et al. (2016), Kim and Lee (2018), and Chen et al. (2018).

The remainder of this thesis is organized as follows. Chapter 2 reviews the previous studies including the work of Lee and Lee (2018) and Lee, Seok and Kim (2018) and introduces the zero-inflated one parameter exponential family INGARCH models and establishes the asymptotic results for the CMLE and the CUSUM tests based on the residuals and standardized residuals. Chapter 3 provides concluding remarks. Finally, all proofs are provided in the Appendix.

Chapter 2

Literature review and main result

2.1 Model formulation

Let $\{Y_t, t \geq 1\}$ be the zero-inflated general nonlinear INGARCH time series of counts satisfying with the conditional distribution of the one-parameter exponential family

$$Y_t|\mathcal{F}_{t-1} \sim p(z|\eta_t), \quad X_t := E(Y_t|\mathcal{F}_{t-1}) = f_\theta(X_{t-1}, Y_{t-1}), \tag{2.1}$$

where $\mathcal{F}_t$ is the σ-field generated by $\eta_1, Y_1, \ldots, Y_t$, and $f_\theta(x, y)$ is a nonnegative bivariate function defined $[0, \infty) \times \mathbb{N}_0$ ($\mathbb{N}_0 = \mathbb{N} \cup \{0\}$), depending on the parameter $\theta \in \Theta \subset \mathbb{R}^d$, and $p(\cdot|\cdot)$ is a probability mass function given by

$$\begin{aligned} p(z|\eta) &= \{\rho + (1-\rho)q(0|\eta)\}I(z=0) + (1-\rho)q(z|\eta)I(z \geq 1) \quad \text{with} \\ q(z|\eta) &= \exp\{\eta z - A(\eta)\}h(z), \quad z \geq 0, \quad \text{and} \quad 0 \leq \rho < 1. \end{aligned}$$

Here η is the natural parameter, $A(\cdot)$ and $h(\cdot)$ are known functions, and $A'(\cdot)$ exists and is strictly increasing, and further, $\eta_t = (A')^{-1}(\frac{X_t}{1-\rho})$. We express

$B(\eta) = A'(\eta)$. Then, $(1-\rho)B(\eta_t)$ and $(1-\rho)\{B'(\eta_t)+\rho B(\eta_t)^2\}$ are the conditional mean and variance of Y_t, respectively, and $X_t = (1-\rho)B(\eta_t)$, $C(\eta_t) := (1-\rho)\{B'(\eta_t)+\rho B(\eta_t)^2\}$. To emphasize the role of θ, we also use notation $X_t(\theta)$ and $\eta_t(\theta)$ to stand for X_t and η_t. Note that although X_t necessarily depends upon ρ, the recursion in model (2.1) is designed to operate with a link function only depending on θ as done in Lee et al. (2016). In fact, we can write $X_t = X_t(\rho,\theta) = f_\theta(X_{t-1}(\rho,\theta), Y_{t-1})$. We put $\vartheta = (\rho, \theta^T)^T$ and denote the true parameter by $\vartheta_0 = (\rho_0, \theta_0^T)^T$.

As an example of model (1) with $\rho = 0$, we can consider Poisson (linear) INGARCH model, $Y_t|\mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t)$, $\lambda_t = \omega + \alpha\lambda_{t-1} + \beta Y_{t-1}$. In this case, we have $\eta_t = \log(X_t(\theta))$, $A(\eta) = e^\eta$, $B'(\eta) = e^\eta$, $B'(\eta_t) = X_t(\theta)$, and $B'(\eta) = B''(\eta)$. Moreover, we can consider the negative binomial (NB) INGARCH model, $Y_t|\mathcal{F}_{t-1} \sim \text{NB}(r, p_t)$, $X_t = \frac{r(1-p_t)}{p_t} = \omega + \alpha X_{t-1} + \beta Y_{t-1}$, where $r \in \mathbb{N}$ is assumed to be known and $Y \sim \text{NB}(r,p)$ implies

$$P(Y_t = k) = \begin{pmatrix} k+r-1 \\ r-1 \end{pmatrix} (1-p)^k p^r, \qquad k = 0, 1, 2, \ldots.$$

In this case, $\eta_t = \log(X_t(\theta)/(X_t(\theta)+r))$, $A(\eta) = -r\log(r/(1-e^\eta))$, $B'(\eta) = re^\eta/(1-e^\eta)^2$, $B'(\eta_t) = X_t(\theta)(X_t(\theta)+r)/r$, and $B''(\eta) = re^\eta(1+e^\eta)/(1-e^\eta)^3$.

In what follows, we assume

(A0) For all $x, x' \geq 0$ and $y, y' \in \mathbb{N}_0$,

$$\sup_{\theta\in\Theta} |f_\theta(x,y) - f_\theta(x',y')| \leq \omega_1|x-x'| + \omega_2|y-y'|,$$

where $\omega_1, \omega_2 \geq 0$ satisfies $\omega_1 + \omega_2 < 1$.

Davis and Liu (2016) showed that this assumption ensures the strict stationarity and ergodicity of $\{(X_t, Y_t)\}$ when ρ equals 0 and the existence of a measurable function $g_\infty^\theta : \mathbb{N}_0^\infty = \{(n_1, n_2, \ldots), n_i \in \mathbb{N}_0, i = 1, 2, \ldots\} \to [0, \infty)$

such that $X_t(\theta) = g^\theta_\infty(Y_{t-1}, Y_{t-2}, \ldots)$ a.s., which also holds for any $\rho \in (0, 1)$. The stationarity and ergodicity in model (2.1) can be shown as seen below similarly to Davis and Liu (2016), but verifying them is not straightforward.

Theorem 1 (Stationarity) *Suppose that the bivariate chain $\{(Y_t, X_t); t \in \mathbb{N}\}$ in model (2.1) satisfies* **(A0)**. *Then, it holds that*

(i) There exists a random variable Z_∞ such that for all x, $Z_n(x) \to Z_\infty$ almost surely. In particular, Z_∞ does not depend on x and has distribution π, which makes the stationary distribution of $\{X_t\}$.

(ii) The $\{X_t\}$ is geometric moment contracting (GMC) with π. Furthermore, $E_\pi X_1 < \infty$.

(iii) If $\{X_t\}$ has stationary distribution π, then $\{Y_t\}$ is also a stationary process.

Theorem 2 (Ergodicity) *Suppose that the bivariate chain $\{(Y_t, X_t); t \in \mathbb{N}\}$ in model (2.1) satisfies* **(A0)**. *Then, it holds that*

(i) There exists a measurable function $g^\theta_\infty : \mathbb{N}_0^\infty \to [0, \infty)$ such that $X_t(\theta) = g^\theta_\infty(Y_{t-1}, Y_{t-2}, \ldots)$ almost surely.

(ii) The process $\{Y_t\}$ is absolutely regular (or β-mixing) with coefficients satisfying $\beta(n) \le (a + b)^n/(1 - (a + b))$. Hence, $\{(Y_t, X_t); t \in \mathbb{N}\}$ is ergodic.

Theorems 1 and 2 play an important role in establishing the consistency and asymptotic normality of parameter estimates and the weak convergence of the CUSUM tests based on those estimates.

2.2 Conditional likelihood inference

The conditional likelihood function of model (2.1), based on the observations $Y_1, \ldots, Y_n$, is given by

$$\tilde{\mathcal{L}}(\theta) = \prod_{t=1}^{n} \left[\tilde{L}_{t0}(\theta) I(Y_t = 0) + \tilde{L}_{t1}(\theta) I(Y_t \geq 1)\right],$$

with

$$\begin{aligned}
\tilde{L}_{t0}(\vartheta) &= \rho + (1-\rho)\exp\{-A(\tilde{\eta}_t(\theta))\}h(0), \\
\tilde{L}_{t1}(\vartheta) &= (1-\rho)\exp\{\tilde{\eta}_t(\theta)Y_t - A(\tilde{\eta}_t(\theta))\}h(Y_t)
\end{aligned}$$

where $\tilde{\eta}_t(\theta) = B^{-1}(\frac{\tilde{X}_t(\theta)}{1-\rho})$ and $\tilde{X}_t(\theta) = f_\theta(\tilde{X}_{t-1}(\theta), Y_{t-1})$, $t \geq 2$, are recursively updated with an initial random variable $\tilde{X}_1$. The CMLE of ϑ_0 is defined as

$$\hat{\vartheta}_n = \operatorname*{arg\,max}_{\vartheta \in \boldsymbol{\Theta}} \log \tilde{L}(\vartheta)$$

where

$$\tilde{L}(\vartheta) := \sum_{t=1}^{n} \tilde{\ell}_t(\vartheta) = \sum_{t=1}^{n} \left[\tilde{\ell}_{t0}(\vartheta) I(Y_t = 0) + \tilde{\ell}_{t1}(\vartheta) I(Y_t \geq 1)\right]$$

with

$$\begin{aligned}
\tilde{\ell}_{t0}(\vartheta) &= \log\left[\rho + (1-\rho)\exp\{-A(\tilde{\eta}_t(\theta))\}h(0)\right], \\
\tilde{\ell}_{t1}(\vartheta) &= \log(1-\rho) + \tilde{\eta}_t(\theta)Y_t - A(\tilde{\eta}_t(\theta)) - \log h(Y_t).
\end{aligned}$$

We impose some regularity conditions, wherein V and $\xi \in (0,1)$ stand for a generic integrable random variable and constant, respectively; symbol $\|\cdot\|$ denotes the L^1 norm for matrices and vectors; and $E(\cdot)$ is taken under ϑ_0. Further, we use notation $\tilde{\eta}_t = \tilde{\eta}_t(\theta)$ for simplicity. To ensure the strong consistency and asymptotic normality of $\hat{\theta}_n$, we assume the following conditions:

(A1) $E\left(\sup_{\theta\in\Theta} X_1^2(\theta)\right) < \infty$ and $E\left(\sup_{\theta\in\Theta} \tilde{X}_1^2(\theta)\right) < \infty$.

(**A2**) For all t, $\sup_{\theta\in\Theta}\sup_{0\le\delta\le 1} B'((1-\delta)\eta_t+\delta\tilde{\eta}_t)\ge \underline{c}$ for some constant $\underline{c}>0$, and $\sup_{\theta\in\Theta}\sup_{0\le\delta\le 1} B((1-\delta)\eta_t+\delta\tilde{\eta}_t)\ge \underline{d}$ for some constant $\underline{d}>0$.

(**A3**) $X_t(\theta)=X_t(\theta_0)$ a.s. implies $\theta=\theta_0$.

(**A4**) For all t, there exists constant $M>0$,

$$\sup_{\theta\in\Theta}\frac{B(\eta_t)}{B'(\eta_t)}\le M$$

(**A5**)

$$E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial X_1(\theta)}{\partial\theta}\right\|^4\right)<\infty \quad\text{and}\quad E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2 X_1(\theta)}{\partial\theta\partial\theta^T}\right\|^2\right)<\infty.$$

(**A6**) For all t,

$$\sup_{\theta\in\Theta}\left\|\frac{\partial\tilde{X}_t(\theta)}{\partial\theta}-\frac{\partial X_t(\theta)}{\partial\theta}\right\|\le V\gamma^t \text{ and } \sup_{\theta\in\Theta}\left\|\frac{\partial^2\tilde{\eta}_t}{\partial\theta\partial\theta^T}-\frac{\partial^2\eta_t}{\partial\theta\partial\theta^T}\right\|\le V\gamma^t \text{ a.s.}$$

(**A7**) For all t, $\sup_{\theta\in\Theta}|B'(\tilde{\eta}_t)-B'(\eta_t)|\le V\gamma^t$ a.s. and $\sup_{\theta\in\Theta}|C(\tilde{\eta}_t)-C(\eta_t)|\le V\gamma^t$ a.s.

(**A8**) ϑ_0 is an interior point in the compact parameter space $\mathbf{\Theta}:=[0,\rho^*]\times\Theta$ for some constant $0<\rho^*<1$. $\Theta\in\mathbb{R}^d$.

(**A9**) $E\left\{Y_1\sup_{\theta\in\Theta}|\eta_1(\theta)|\right\}<\infty$.

(**A10**)

$$E\left[\sup_{\theta\in\Theta}\left\|B'(\eta_1)\left(\frac{\partial\eta_1}{\partial\theta}\cdot\frac{\partial\eta_1}{\partial\theta^T}\right)\right\|\right]<\infty,\ E\left[\sup_{\theta\in\Theta}\left\|(Y_1-B(\eta_1))\frac{\partial^2\eta_1}{\partial\theta\partial\theta^T}\right\|\right]<\infty.$$

(**A11**) For all t, $\sup_{\theta\in\Theta}C(\eta_t)^{-3/2}C'(\eta_t)\le K$ for some $K>0$.

(**A12**) $\nu^T\frac{\partial X_1(\theta)}{\partial\theta}=0$ a.s. (or equivalently, $\nu^T\frac{\partial\eta_1(\theta)}{\partial\theta}=0$ a.s.) if and only if $\nu=0$.

(**A13**) For any $\theta\in\Theta$ and $\mathbf{y}\in\mathbb{N}_0^\infty$, $f_\infty^\theta(\mathbf{y})\ge x_\theta^*\in\mathcal{R}(B)$, where $\mathcal{R}(B)$ is the range of $B(\eta)$. Moreover, $x_\theta^*\ge x^*\in\mathcal{R}(B)$ for all θ.

Davis and Liu (2016) derived the asymptotic properties of the CMLE. The proposition below can be proven using Lemmas 1 and 2 in the Appendix, in a manner similar to that seen in their Theorems 1 and 2.

Theorem 3 (Strong Consistency) *Suppose that conditions* **(A0)**–**(A3)** *hold. Then, the conditional maximum likelihood estimator (CMLE) $\hat{\vartheta}_n$ is strongly consistent. That is, as $n \to \infty$,*

$$\hat{\vartheta}_n \longrightarrow \vartheta_0 \;\; a.s.,$$

Theorem 4 (Asymptotic Normality) *Suppose that conditions* **(A0)**–**(A13)** *hold. Then, the conditional maximum likelihood estimator (CMLE) $\hat{\vartheta}_n$ is asymptotically normal. That is, as $n \to \infty$,*

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} N(0, I(\vartheta_0)^{-1}),$$

where

$$I(\vartheta_0) = E\left(\frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta}\frac{\partial \ell_t(\theta_0)}{\partial \vartheta^T}\right) = -E\left(\frac{\partial^2 \ell_t(\vartheta_0)}{\partial \vartheta \partial \vartheta^T}\right) = E\left(B^{'}(\eta_t(\theta_0))\frac{\partial \eta_t(\theta_0)}{\partial \theta}\frac{\partial \eta_t(\theta_0)}{\partial \theta^T}\right)$$

with

$$\begin{aligned} \ell_{t0}(\vartheta) &= \log\left[\rho + (1-\rho)\exp\{-A(\eta_t(\theta))\}h(0)\right], \\ \ell_{t1}(\vartheta) &= \log(1-\rho) + \eta_t(\theta)Y_t - A(\eta_t(\theta)) - \log h(Y_t). \end{aligned}$$

2.3 Change point test

In this section, we study the residual-based CUSUM tests used to assess the hypotheses

$$H_0 : \theta \text{ does not change over } Y_1, \cdots, Y_n \;\; vs. \;\; H_1 : \text{not } H_0.$$

We consider the two types of residuals: $\epsilon_{t,1} = Y_t - X_t(\theta_0)$ and $\epsilon_{t,2} = (Y_t - X_t(\theta_0))/\sqrt{C(\eta_t(\theta_0))}$. The former is considered by Franke et al. (2012), Kang and Lee (2014), and Lee et al. (2016a,b) in some Poisson AR models, whereas the latter is newly considered here. Since $\{\epsilon_{t,i}, \mathcal{F}_t\}$, $i = 1, 2$, are stationary ergodic martingale difference sequences, using a functional central limit theorem, we can derive

$$\sup_{0<s<1} \frac{1}{\sqrt{n}\tau_i} \left| \sum_{t=1}^{[ns]} \epsilon_{t,i} - \frac{k}{n} \sum_{t=1}^{n} \epsilon_{t,i} \right| \xrightarrow{w} \sup_{0\leq s\leq 1} |\mathbf{B}_1^\circ(s)|, \tag{2.2}$$

where $\tau_1^2 = Var(\epsilon_{1,1})$ and $\tau_2^2 = 1$. Note that $\epsilon_{t,i}$ is not observable, but it is possible to compute $\hat{\epsilon}_{t,1} = Y_t - \hat{X}_t$ or $\hat{\epsilon}_{t,2} = (Y_t - \hat{X}_t)/\sqrt{C(\hat{\eta}_t)}$, where $\hat{X}_t = f_{\hat{\theta}_n}(\hat{X}_{t-1}, Y_{t-1}), \hat{\eta}_t = B^{-1}(\hat{X}_t/(1-\rho))$ for $t \geq 2$ and $\hat{X}_1$ is an arbitrarily chosen value. We thus consider the tests

$$T_n^{res,i} = \max_{1\leq k\leq n} \frac{1}{\sqrt{n}\hat{\tau}_{n,i}} \left| \sum_{t=1}^{k} \hat{\epsilon}_{t,i} - \frac{k}{n} \sum_{t=1}^{n} \hat{\epsilon}_{t,i} \right|, \tag{2.3}$$

where $\hat{\tau}_{n,1}^2 = \frac{1}{n}\sum_{t=1}^n \hat{\epsilon}_{t,1}^2$ and $\hat{\tau}_{n,2}^2 = \frac{1}{n}\sum_{t=1}^n \hat{\epsilon}_{t,2}^2$. Then, we can obtain the following theorem, the proof of which is similar to that of Kang and Lee (2014) and is omitted for brevity.

Theorem 5 (Residual-based CUSUM test) *Suppose that conditions* **(A0)**–**(A13)** *hold. Then, under* H_0, *as* $n \to \infty$,

$$T_n^{res,i} \xrightarrow{d} \sup_{0\leq s\leq 1} |\mathbf{B}_1^\circ(s)|, \quad i = 1, 2.$$

Chapter 3

Conclusion

In this study, we formulated the zero-inflated general integer-valued GARCH models, and established the asymptotic property for CMLE of the models. Also, we considered CUSUM tests based on residuals, and derived their limiting null distributions under certain conditions. This thesis is based on the work of Lee, Seok and Kim (2018), and empirical studies are now on investigation.

Appendix A

Proofs of Main Results

Definition 1 *A random variable X is said to be stochastic smaller than a random variable Y ($X \leq_{ST} Y$) if $F_X(x) \geq F_Y(x)$ for all $x \in \mathbb{R}$ where F_X is is the cumulative distribution function of a random variable X.*

Lemma 1 *Let two random variables Y' and Y'' follow a distribution belonging to ZIFE with the same A, h, ρ and μ(counting measure), but with natural parameter η' and η'', respectively.*
If $\eta' \leq \eta''$, then $Y' \leq_{ST} Y''$.

Proof. Let $q(z|\eta)$ be the probability mass function of one-parameter exponential family. Then, by proposition 6 of Davis and Liu (2016),

If $\begin{cases} Z' \sim q(z|\eta') \\ Z'' \sim q(z|\eta'') \end{cases}$ and $\eta' \leq \eta''$, then $Z' \leq_{ST} Z''$. (i.e. $q(z|\eta') \geq q(z|\eta'')$).

So, if $\eta' \leq \eta''$, then $q(z|\eta') \geq q(z|\eta'')$ for all $y \in \mathbb{N}_0$. Let $p(y|\eta)$ be the probability mass function of ZIEF. Then, for all $\rho \in [0, 1)$ and $\eta' \leq \eta''$,

$$p(y|\eta') = \rho\delta_{0,y} + (1-\rho)q(y|\eta') \geq \rho\delta_{0,y} + (1-\rho)q(y|\eta'') = p(y|\eta'').$$

Thus, if $\begin{cases} Y' \sim p(z|\eta') \\ Y'' \sim p(z|\eta'') \end{cases}$ and $\eta' \leq \eta''$, then $Y' \leq_{ST} Y''$. This completes the proof. $\square$

Let $F_x(y)$ be the cumulative distribution function of ZIEF $p(y|\eta)$ with $x := E[Y] = (1-\rho)B(\eta)$.

Lemma 2 *Let $U \sim Unif(0,1)$.*
Define two random variables $\begin{cases} Y' := F_{x'}^{-1}(U) \\ Y'' := F_{x''}^{-1}(U) \end{cases}$
where $\begin{cases} x' := (1-\rho)B(\eta') = E[Y'] \\ x'' := (1-\rho)B(\eta'') = E[Y''] \end{cases}$. *Then $E|Y' - Y''| = |x' - x''|$.*

Proof. First, assume $x' \leq x''$. Then $\eta' \leq \eta''$ and $Y' \leq_{ST} Y''$. So $F_{x''}^{-1}(U) \leq F_{x'}^{-1}(U)$ for all $u \in (0,1)$. That is, $Y' \leq Y''$. Thus $E|Y' - Y''| = E(Y'' - Y') = x'' - x' = |x' - x''|$. Likewise, if $x'' \leq x'$, then $E|Y' - Y''| = E(Y' - Y'') = x' - x'' = |x' - x''|$. These establish the lemma. $\square$

Proof of Theorem 1. Use Lemma 1 and the proof of Proposition 1 of Davis and Liu (2016).
Proof of Theorem 2. Use Lemma 2 and the proof of Proposition 2 of Davis and Liu (2016).

Before proving Theorems 3 and 4, we prepare some lemmas. In what follows, we use notation $\eta_t^0 = \eta_t(\theta_0)$ and $\eta_t^n = \eta_t(\hat{\theta}_n)$.

Lemma 3 *Suppose that conditions* **(A0)**, **(A1)** *and* **(A2)** *hold. Then, we have*

$$\sup_{\theta \in \Theta} |\tilde{X}_t(\theta) - X_t(\theta)| \leq V\gamma^t, \quad \sup_{\theta \in \Theta} |\tilde{\eta}_t - \eta_t| \leq V\gamma^t \quad a.s.$$

Proof. Note that

$$\begin{aligned}|\tilde{X}_t(\theta) - X_t(\theta)| &= \Big| f_\theta(\tilde{X}_{t-1}(\theta), Y_{t-1}) - f_\theta(X_{t-1}(\theta), Y_{t-1}) \Big| \\ &\leq \omega_1 |\tilde{X}_{t-1}(\theta) - X_{t-1}(\theta)| \leq \omega_1^{t-1} |\tilde{X}_1 - X_1(\theta)|.\end{aligned}$$

Then, using the mean value theorem and **(A2)**, we have

$$\begin{aligned}|\tilde{\eta}_t - \eta_t| &= |B^{-1}(\frac{\tilde{X}_t(\theta)}{1-\rho}) - B^{-1}(\frac{X_t(\theta)}{1-\rho})| \leq \frac{\omega_1^{t-1}}{(1-\rho)B'(\eta_t^*)} |\tilde{X}_1 - X_1(\theta)| \\ &\leq \frac{\omega_1^{t-1}}{(1-\rho)\underline{c}} |\tilde{X}_1 - X_1(\theta)|,\end{aligned}$$

where $\eta_t^* = B^{-1}(\frac{X_t^*}{1-\rho})$ and X_t^* is an intermediate point between $\tilde{X}_t(\theta)$ and $X_t(\theta)$. Hence, using **(A1)**, we establish the lemma. □

Lemma 4 *Suppose that conditions* **(A0)**, **(A1)** *and* **(A2)** *hold. Then, we have*

$$\sup_{\theta\in\Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \tilde{\ell}_t(\theta) - \frac{1}{n} \sum_{t=1}^{n} \ell_t(\theta) \right| \longrightarrow 0 \quad a.s.$$

Proof. It suffices to show that for $i = 0, 1$,

$$\sup_{\theta\in\Theta} |\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)| \to 0 \quad a.s. \quad \text{as} \quad t \to \infty. \tag{A.1}$$

Since $|\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)| = [\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)]_+ + [\ell_{ti}(\theta) - \tilde{\ell}_{ti}(\theta)]_+$ for $i = 0, 1$, we first show that for $i = 0, 1$, as $t \to \infty$

$$\sup_{\theta\in\Theta} [\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)]_+ \longrightarrow 0 \quad a.s.$$

First, note that, by the mean value theorem, $\log x \leq x - 1$ and Lemma 3,

$$\begin{aligned}[\tilde{\ell}_{t0}(\theta) - \ell_{t0}(\theta)]_+ &\leq [\frac{(1-\rho)h(0)}{\rho + (1-\rho)h(0)e^{-A(\eta_t)}} (e^{-A(\tilde{\eta}_t)} - e^{-A(\eta_t)})]_+ \\ &\leq \frac{h(0)(1-\rho)}{\rho} |e^{-A(\tilde{\eta}_t)} - e^{-A(\eta_t)}| \\ &= \frac{h(0)(1-\rho)A'(\tilde{\eta}_t)}{\rho} |\tilde{\eta}_t - \eta_t| e^{-A(\eta_t^*)} \\ &\leq \frac{h(0)X_t^*(\theta)}{\rho} |\tilde{\eta}_t - \eta_t|\end{aligned}$$

for some intermediate points X_t^* between $X_t(\theta)$ and $\tilde{X}_t(\theta)$ and $\eta_t^* = B^{-1}(X_t^*/(1-\rho))$. Since

$$X_t^*(\theta) \leq X_t(\theta) + |\tilde{X}_t(\theta) - X_t(\theta)| \leq X_t(\theta) + V\gamma^t,$$

according to **(A1)** and Lemma 3, we can show that $\sup_{\theta\in\Theta}[\tilde{\ell}_{t0}(\theta) - \ell_{t0}(\theta)]_+ \longrightarrow 0 \quad a.s.$.

Second, note that, by the mean value theorem, **(A2)** and Lemma 3,

$$\begin{aligned}
[\tilde{\ell}_{t1}(\theta) - \ell_{t1}(\theta)]_+ &\leq |\tilde{\eta}_t - \eta_t| Y_t + |A(\tilde{\eta}_t) - A(\eta_t)| \\
&= \frac{Y_t + \frac{X_t^*}{1-\rho}}{B'(\eta_t^{**})(1-\rho)} |\tilde{X}_t(\theta) - X_t(\theta)| \\
&\leq \frac{(1-\rho)Y_t + X_t^*}{\underline{c}(1-\rho)^2} V\gamma^t
\end{aligned}$$

for some intermediate points X_t^{**} between $X_t(\theta)$ and $\tilde{X}_t(\theta)$ and $\eta_t^{**} = B^{-1}(X_t^{**}/(1-\rho))$. Since

$$X_t^*(\theta) \leq X_t(\theta) + |\tilde{X}_t(\theta) - X_t(\theta)| \leq X_t(\theta) + V\gamma^t,$$

according to **(A1)** and Lemma 3, we can show that $\sup_{\theta\in\Theta}[\tilde{\ell}_{t1}(\theta) - \ell_{t1}(\theta)]_+ \longrightarrow 0 \quad a.s.$.

Similarly, it can be shown that for $i = 0, 1$, as $t \to \infty$

$$\sup_{\theta\in\Theta}[\tilde{\ell}_{ti}(\theta) - \ell_{ti}(\theta)]_- \longrightarrow 0 \quad a.s.$$

This validates the lemma. □

Proof of Theorem 3. We can express

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^n \tilde{\ell}_t(\theta) - E\ell_t(\theta)\right| \leq \sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^n \tilde{\ell}_t(\theta) - \frac{1}{n}\sum_{t=1}^n \ell_t(\theta)\right| + \sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^n \ell_t(\theta) - E\ell_t(\theta)\right|.$$

The first term of the RHS of the inequality above converges almost surely to 0 by Lemma 4. Since ℓ_t is strictly stationary and ergodic, the second term also converges almost surely to 0.

Hence we obtain

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta)-E\ell_t(\theta)\right|\longrightarrow 0 \quad a.s..$$

Note that

$$\begin{aligned}0\le E\ell_t(\theta_0)-E\ell_t(\hat{\theta}_n) &\le (E\ell_t(\theta_0)-\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta_0))+(\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\hat{\theta}_n)-E\ell_t(\hat{\theta}_n))\\ &\le 2\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\tilde{\ell}_t(\theta)-E\ell_t(\theta)\right|.\end{aligned}$$

Therefore, by **(A1)** and Lebesgue dominated convergence theorem,

$$\lim_{n\to\infty}E\ell_t(\hat{\theta}_n)=E\ell_t(\theta_0).$$

Since $E\ell_t(\theta)$ has a unique maximum at θ_0 by **(A3)**, the strong consistency can be proven. □

The first derivatives of log conditional likelihood function are obtained as follows:

$$\frac{\partial\ell_t(\vartheta)}{\partial\vartheta}=\frac{\partial\ell_{t0}(\vartheta)}{\partial\vartheta}I(Y_t=0)+\frac{\partial\ell_{t1}(\vartheta)}{\partial\vartheta}I(Y_t\ge 1)$$

with

$$\begin{aligned}\frac{\partial\ell_{t0}(\vartheta)}{\partial\rho} &= \frac{1-h(0)e^{-A(\eta_t)}-\frac{X_tB(\eta_t)}{(1-\rho)B'(\eta_t)}h(0)e^{-A(\eta_t)}}{\rho+(1-\rho)h(0)e^{-A(\eta_t)}}\\ \frac{\partial\ell_{t0}(\vartheta)}{\partial\theta} &= -\frac{\frac{B(\eta_t)}{B'(\eta_t)}h(0)e^{-A(\eta_t)}}{\rho+(1-\rho)h(0)e^{-A(\eta_t)}}\frac{\partial X_t}{\partial\theta}\\ &= -Z_{t0}(\vartheta)\frac{\partial X_t}{\partial\theta}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial\ell_{t1}(\vartheta)}{\partial\rho} &= -\frac{1}{1-\rho}+\frac{\{Y_t-B(\eta_t)\}X_t}{(1-\rho)^2B'(\eta_t)}\\ \frac{\partial\ell_{t1}(\vartheta)}{\partial\theta} &= \frac{Y_t-B(\eta_t)}{(1-\rho)B'(\eta_t)}\frac{\partial X_t}{\partial\theta}\\ &= Z_{t1}(\vartheta)\frac{\partial X_t}{\partial\theta}.\end{aligned}$$

Also, the second derivatives are as follows:

$$\frac{\partial^2 \ell_t(\vartheta)}{\partial\vartheta\partial\vartheta^T} = \frac{\partial^2 \ell_{t0}(\vartheta)}{\partial\vartheta\partial\vartheta^T} I(Y_t = 0) + \frac{\partial^2 \ell_{t1}(\vartheta)}{\partial\vartheta\partial\vartheta^T} I(Y_t \geq 1)$$

with

$$\begin{aligned}
& \frac{\partial^2 \ell_{t0}(\vartheta)}{\partial\rho^2} \\
= \quad & \Bigg\{ h(0)^2 \left[\frac{B'(\eta_t)B(\eta_t)^2 + B''(\eta_t)B(\eta_t) - B'(\eta_t)^2}{B'(\eta_t)^2} \frac{B(\eta_t)^2}{B'(\eta_t)} - \frac{(B'(\eta_t) + B(\eta_t)^2)^2}{B'(\eta_t)^2} \right] \\
& e^{-2A(\eta_t)} + h(0) \left[\frac{2(B'(\eta_t) + B(\eta_t)^2)}{B'(\eta_t)} + \frac{B'(\eta_t)B(\eta_t)^2 + B''(\eta_t)B(\eta_t) - B'(\eta_t)^2}{B'(\eta_t)^3} \right. \\
& \left. \frac{\rho B(\eta_t)^2}{1-\rho} \right] e^{-A(\eta_t)} - 1 \Bigg\} \times \frac{1}{(\rho + (1-\rho)h(0)e^{-A(\eta_t)})^2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell_{t0}(\vartheta)}{\partial\theta\partial\theta^T} &= \frac{(1-\rho)h(0)[\rho(B'(\eta_t) - B(\eta_t)^2) + (1-\rho)h(0)B'(\eta_t)e^{-A(\eta_t)}]}{[\rho + (1-\rho)h(0)e^{-A(\eta_t)}]^2} \frac{\partial\eta_t}{\partial\theta} \frac{\partial\eta_t}{\partial\theta^T} \\
&+ \frac{(1-\rho)h(0)B(\eta_t)e^{-A(\eta_t)}}{\rho + (1-\rho)h(0)e^{-A(\eta_t)}} \frac{\partial^2\eta_t}{\partial\theta\partial\theta^T} \\
\frac{\partial^2 \ell_{t0}(\vartheta)}{\partial\rho\partial\theta} &= \Bigg\{ \rho \frac{B'(\eta_t)B(\eta_t)^2 + B''(\eta_t)B(\eta_t) - B'(\eta_t)^2}{B'(\eta_t)^2} \\
&+ (1-\rho)\left[1 + \frac{B''(\eta_t)B(\eta_t) - 2B'(\eta_t)^2}{B'(\eta_t)^2} h(0)e^{-A(\eta_t)} \right] \Bigg\} \\
&\times \frac{h(0)B(\eta_t)e^{-A(\eta_t)}}{[\rho + (1-\rho)h(0)e^{-A(\eta_t)}]^2} \frac{\partial\eta_t}{\partial\theta}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 \ell_{t1}(\vartheta)}{\partial\rho^2} &= -\frac{1}{(1-\rho)^2} + \frac{[Y_t - 2B(\eta_t)]B(\eta_t)}{(1-\rho)^2 B'(\eta_t)} \\
&- \frac{[Y_t - B(\eta_t)]B(\eta_t)[B''(\eta_t)B(\eta_t) - B'(\eta_t)^2]}{(1-\rho)^3 B'(\eta_t)^3} \\
\frac{\partial^2 \ell_{t1}(\vartheta)}{\partial\theta\partial\theta^T} &= -\frac{B'(\eta_t)^2 + [Y_t - B(\eta_t)]B''(\eta_t)}{(1-\rho)B'(\eta_t)^2} \frac{\partial\eta_t}{\partial\theta} \frac{\partial\eta_t}{\partial\theta^T} + \frac{Y_t - B(\eta_t)}{(1-\rho)B'(\eta_t)} \frac{\partial^2\eta_t}{\partial\theta\partial\theta^T} \\
\frac{\partial^2 \ell_{t1}(\vartheta)}{\partial\rho\partial\theta} &= \frac{1}{(1-\rho)} \left[(Y_t - 2B(\eta_t)) - \frac{(Y_t - B(\eta_t))B(\eta_t)B''(\eta_t}{B'(\eta_t)^2}) \right] \frac{\partial\eta_t}{\partial\theta}.
\end{aligned}$$

Lemma 5 *Under* **(A2)** *and* **(A4)**, *we have*

(i) $|B(\tilde{\eta}_t) - B(\eta_t)| = \frac{1}{(1-\rho)} \left| \tilde{X}_t(\theta) - X_t(\theta) \right|;$

(ii) $\left| \frac{B(\tilde{\eta}_t)}{B'(\tilde{\eta}_t)} - \frac{B(\eta_t)}{B'(\eta_t)} \right| \leq \frac{1}{\underline{c}^2}(B'(\eta_t) + B(\eta_t)) |B(\tilde{\eta}_t) - B(\eta_t)|;$

(iii) $\left| \frac{B(\tilde{\eta}_t)^2}{B'(\tilde{\eta}_t)} - \frac{B(\eta_t)^2}{B'(\eta_t)} \right| \leq \frac{1}{\underline{c}^2}[(B'(\eta_t) + B(\eta_t))B'(\eta_t) + B(\eta_t)^2] |B(\tilde{\eta}_t) - B(\eta_t)|;$

(iv) $\left| e^{-A(\tilde{\eta}_t)} - e^{-A(\eta_t)} \right| \leq M |B(\tilde{\eta}_t) - B(\eta_t)|;$

(v)
$$\left| \frac{B(\tilde{\eta}_t)}{B'(\tilde{\eta}_t)} e^{-A(\tilde{\eta}_t)} - \frac{B(\eta_t)}{B'(\eta_t)} e^{-A(\eta_t)} \right|$$
$$\leq \frac{1}{\underline{c}^2}(B'(\eta_t) + B(\eta_t)) |B(\tilde{\eta}_t) - B(\eta_t)| + M^2 |B(\tilde{\eta}_t) - B(\eta_t)|;$$

(vi)
$$\left| \frac{B(\tilde{\eta}_t)^2}{B'(\tilde{\eta}_t)} e^{-A(\tilde{\eta}_t)} - \frac{B(\eta_t)^2}{B'(\eta_t)} e^{-A(\eta_t)} \right|$$
$$\leq \frac{1}{\underline{c}^2}[(B'(\eta_t) + B(\eta_t))B'(\eta_t) + B(\eta_t)^2] |B(\tilde{\eta}_t) - B(\eta_t)|$$
$$+ M B(\eta_t) |B(\tilde{\eta}_t) - B(\eta_t)|.$$

Proof. Use the mean value theorem. □

Lemma 6 *Suppose that* **(A0)**-**(A2)** *and* **(A4)** *hold. Then,*

(i) $\sup_{\vartheta\in\Theta} |Z_{t0}(\vartheta)| \leq \frac{Mh(0)}{\delta_L}$ *and* $\sup_{\vartheta\in\Theta} |Z_{t1}(\vartheta)| \leq \frac{Y_t}{\delta_L \underline{c}} + \frac{M}{\delta_L};$

(ii)
$$\sup_{\vartheta\in\Theta} \left| \tilde{Z}_{t0}(\vartheta) - Z_{t0}(\vartheta) \right|$$
$$\leq \frac{h(0)(1-\delta_L)}{\delta_L^2 (1-h(0))^2 \underline{c}^2}[(1-h(0))(B'(\eta_t) + B(\eta_t)) + M^2 \underline{c}^2] |B(\tilde{\eta}_t) - B(\eta_t)|;$$

(iii)
$$\sup_{\vartheta\in\Theta} \left| \tilde{Z}_{t1}(\vartheta) - Z_{t1}(\vartheta) \right|$$
$$\leq \frac{Y_t}{\delta_L \underline{c}^2} \left| B'(\tilde{\eta}_t) - B'(\eta_t) \right| + \frac{1}{\delta_L \underline{c}^2}(B'(\eta_t) + B(\eta_t)) |B(\tilde{\eta}_t) - B(\eta_t)|.$$

Proof. Use Lemma 5. □

Lemma 7 *Suppose that* **(A0)**-**(A13)** *hold. Then, under H_0, as $n \to \infty$,*

$$\text{(i)} \quad \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_t(\vartheta_0)}{\partial \vartheta} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_t(\vartheta_0)}{\partial \vartheta} \right\| = o_P(1);$$

$$\text{(ii)} \quad \sup_{\vartheta \in \Theta} \left\| \frac{\partial^2 \tilde{\ell}_t(\vartheta)}{\partial \vartheta \partial \vartheta^T} - \frac{\partial^2 \ell_t(\vartheta)}{\partial \vartheta \partial \vartheta^T} \right\| \longrightarrow 0 \quad a.s. \text{ as } t \to \infty;$$

$$\text{(iii)} \quad -\frac{1}{n} \frac{\partial^2 \tilde{L}_n(\vartheta_n^*)}{\partial \vartheta \partial \vartheta^T} \longrightarrow I(\vartheta_0) \quad a.s.,$$

where ϑ_n^* is the intermediate point $\hat{\vartheta}_n$ and ϑ_0.

Proof. (i) First, by Lemma 6 and **(A5)**, **(A6)**, for $i = 0, 1$,

$$\begin{aligned}
&\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{ti}(\vartheta_0)}{\partial \theta} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{ti}(\vartheta_0)}{\partial \theta} \right\| \\
&\leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| \tilde{Z}_{ti}(\vartheta_0) \frac{\partial \tilde{X}_t}{\partial \theta}(\theta_0) - Z_{ti}(\vartheta_0) \frac{\partial X_t}{\partial \theta}(\theta_0) \right\| \\
&\leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| (\tilde{Z}_{ti}(\vartheta_0) - Z_{ti}(\vartheta_0)) \frac{\partial \tilde{X}_t}{\partial \theta}(\theta_0) \right\| + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left\| Z_{ti}(\vartheta_0) (\frac{\partial X_t}{\partial \theta}(\theta_0) - \frac{\partial \tilde{X}_t}{\partial \theta}(\theta_0)) \right\| \\
&= o_P(1).
\end{aligned}$$

Second, by Lemma 5,

$$\begin{aligned}
&\left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \tilde{\ell}_{t1}(\vartheta_0)}{\partial \rho} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \ell_{t1}(\vartheta_0)}{\partial \rho} \right| \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left| \frac{\partial \tilde{\ell}_{t1}(\vartheta_0)}{\partial \rho} - \frac{\partial \ell_{t1}(\vartheta_0)}{\partial \rho} \right| \\
&\leq \frac{1}{\sqrt{n}} \frac{1}{(1-\rho)} \sum_{t=1}^{n} Y_t \left| \frac{B(\tilde{\eta}_t)}{B'(\tilde{\eta}_t)} - \frac{B(\eta_t)}{B'(\eta_t)} \right| + \frac{1}{\sqrt{n}} \frac{1}{(1-\rho)} \sum_{t=1}^{n} \left| \frac{B(\tilde{\eta}_t)^2}{B'(\tilde{\eta}_t)} - \frac{B(\eta_t)^2}{B'(\eta_t)} \right| \\
&= o_P(1).
\end{aligned}$$

Finally, by Lemma 5,

$$
\begin{aligned}
&\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\tilde{\ell}_{t0}(\vartheta_0)}{\partial\rho}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial\ell_{t0}(\vartheta_0)}{\partial\rho}\right| \\
&\leq \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|\frac{\partial\tilde{\ell}_{t0}(\vartheta_0)}{\partial\rho}-\frac{\partial\ell_{t0}(\vartheta_0)}{\partial\rho}\right| \\
&\leq \frac{1}{\sqrt{n}}\frac{1}{\delta_L^2(1-h(0))^2}\sum_{t=1}^{n}\left[h(0)\left|e^{-A(\tilde{\eta}_t)}-e^{-A(\eta_t)}\right|\right. \\
&+\rho h(0)\left|\frac{B(\tilde{\eta}_t)^2}{B'(\tilde{\eta}_t)}e^{-A(\tilde{\eta}_t)}-\frac{B(\eta_t)^2}{B'(\eta_t)}e^{-A(\eta_t)}\right| \\
&\left.+(1-\rho)h(0)^2\left|\frac{B(\tilde{\eta}_t)^2}{B'(\tilde{\eta}_t)}-\frac{B(\eta_t)^2}{B'(\eta_t)}\right|\right]=o_P(1).
\end{aligned}
$$

Thus, the proof is completed.

(ii) It's similar to the proof of Proposition 5 of Lee et al. (2016a). It is easy to show that for $i=0,1$, as $t\to\infty$,

$$
\sup_{\vartheta\in\Theta}\left|\frac{\partial^2\tilde{\ell}_{ti}(\vartheta)}{\partial\rho^2}-\frac{\partial^2\ell_{ti}(\vartheta)}{\partial\rho^2}\right|\longrightarrow 0 \quad a.s.
$$

$$
\sup_{\vartheta\in\Theta}\left\|\frac{\partial^2\tilde{\ell}_{ti}(\vartheta)}{\partial\theta\partial\theta^T}-\frac{\partial^2\ell_{ti}(\vartheta)}{\partial\theta\partial\theta^T}\right\|\longrightarrow 0 \quad a.s.
$$

$$
\sup_{\vartheta\in\Theta}\left\|\frac{\partial^2\tilde{\ell}_{ti}(\vartheta)}{\partial\rho\partial\theta}-\frac{\partial^2\ell_{ti}(\vartheta)}{\partial\rho\partial\theta}\right\|\longrightarrow 0 \quad a.s.
$$

Thus, the proof is completed.

(iii) This can be proven almost identically to the proof of Proposition 5 of Lee et al. (2016a). □

Proof of Theorem 4. Since $\{\partial\ell_t(\theta_0)/\partial\theta;\mathcal{F}_t\}$ forms a martingale difference sequence, we can show that $\frac{1}{\sqrt{n}}\partial\ell_t(\theta_0)/\partial\theta$ converges weakly to $N(0,I(\theta_0))$ by using a martingale central limit theorem and the Cramér-Wold device. Then, by using Taylor's theorem and Lemma 5, we can prove the theorem. $\square$

Proof of Theorem 5. We consider $T_n^{res,2}$ only (the proof of $T_n^{res,1} \xrightarrow{w} \sup_{0\le s\le 1} |\mathbf{B}_1^\circ(s)|$ is similar to that of Kang and Lee (2014)).

First, note that by **(A2)**,

$$\begin{aligned} C(\eta_t) &= (1-\rho)\{B'(\eta_t) + \rho B(\eta_t)^2\} \\ &\ge (1-\rho)(\underline{c} + \rho \underline{d}^2) \\ &:= \underline{f} > 0 \end{aligned}$$

We write

$$\begin{aligned} \hat{\epsilon}_{t,2} - \epsilon_{t,2} &= \frac{Y_t - \hat{X}_t}{\sqrt{C(\hat{\eta}_t)}} - \frac{Y_t - X_t(\theta_0)}{\sqrt{C(\eta_t^0)}} \\ &= (X_t(\theta_0) - \hat{X}_t)\left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^0)}}\right) \\ &\quad + \epsilon_{t,1}\left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^0)}}\right) + \frac{1}{\sqrt{C(\eta_t^0)}}(X_t(\theta_0) - \hat{X}_t)) \\ &:= R_{t,1} + R_{t,2} + R_{t,3}. \end{aligned}$$

In view of (2.2), it suffices to show that

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^k (\hat{\epsilon}_{t,2} - \epsilon_{t,2}) - \frac{k}{n}\sum_{t=1}^n (\hat{\epsilon}_{t,2} - \epsilon_{t,2}) \right| = o_P(1), \tag{A.2}$$

i.e., for $i = 1, 2, 3$,

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^k R_{t,i} - \frac{k}{n}\sum_{t=1}^n R_{t,i} \right| = o_P(1). \tag{A.3}$$

Firstly, we express

$$\max_{1\le k\le n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^k R_{t,1} - \frac{k}{n}\sum_{t=1}^n R_{t,1} \right| \le \frac{2}{\sqrt{n}}\sum_{t=1}^n |R_{t,1}| \le I_{n,1} + I_{n,2} + I_{n,3},$$

where

$$I_{n,1} = \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \left| \left(X_t(\theta_0) - X_t(\hat{\theta}_n) \right) \left(\frac{1}{\sqrt{C(\eta_t^0)}} - \frac{1}{\sqrt{C(\eta_t^n)}} \right) \right|,$$
$$I_{n,2} = \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \left| \left(X_t(\theta_0) - X_t(\hat{\theta}_n) \right) \left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^n)}} \right) \right|,$$
$$I_{n,3} = \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \left| \left(X_t(\hat{\theta}_n) - \hat{X}_t \right) \left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^0)}} \right) \right|.$$

Using the mean value theorem with intermediate points $\theta^*_{n,1}$ and $\theta^*_{n,2}$ between $\hat{\theta}_n$ and θ_0, we have

$$\begin{aligned}
& I_{n,1} \\
= \;& \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left| (\hat{\theta}_n - \theta_0)^T \frac{\partial X_t(\theta^*_{n,1})}{\partial \theta} \cdot \frac{C'(\eta_t(\theta^*_{n,2}))}{C(\eta_t(\theta^*_{n,2}))^{3/2}} \frac{1}{(1-\rho) B'(\eta_t(\theta^*_{n,2}))} \right| \\
& \left| (\hat{\theta}_n - \theta_0)^T \frac{\partial X_t(\theta^*_{n,2})}{\partial \theta} \right| \\
\leq \;& n \|\hat{\theta}_n - \theta_0\|^2 \frac{1}{\underline{c}(1-\rho)} \frac{1}{n\sqrt{n}} \sum_{t=1}^{n} \left| \sup_{\theta \in \Theta} \frac{C'(\eta_t)}{C(\eta_t)^{3/2}} \right| \cdot \left\| \sup_{\theta \in \Theta} \frac{\partial X_t(\theta)}{\partial \theta} \right\|^2 \\
= \;& O_p(1) \cdot o_P(1) = o_P(1),
\end{aligned}$$

where we have used Theorem 4 and **(A2)**, **(A5)** and **(A11)**. Since $\hat{\eta}_t$ can be represented as $\tilde{\eta}_t(\hat{\theta}_n) = B^{-1}(\tilde{X}_t(\hat{\theta}_n/(1-\rho)))$ with $\tilde{X}_1(\hat{\theta}_n) = \hat{X}_1$, we have

$$\begin{aligned}
I_{n,2} \;\leq\; & \frac{1}{\sqrt{n}} \frac{1}{\underline{f}^{3/2}} \sum_{t=1}^{n} \left| \left(X_t(\theta_0) - X_t(\hat{\theta}_n) \right) \left(C(\eta_t^n) - C(\hat{\eta}_t) \right) \right| \\
\leq\; & \sqrt{n} \|\hat{\theta}_n - \theta_0\| \frac{V}{n \underline{f}^{3/2}} \sum_{t=1}^{n} \gamma^t \left\| \frac{\partial X_t(\theta^*_{n,1})}{\partial \theta} \right\| = O_p(1) \cdot o_P(1) = o_P(1)
\end{aligned}$$

with intermediate point $\theta^*_{n,1}$ between $\hat{\theta}_n$ and θ_0, due to **(A2)**, **(A5)** and **(A7)**. Furthermore, note that $|\hat{X}_t - X_t(\hat{\theta}_n)| \leq V\gamma^t$ a.s. since owing to **(A0)**,

$$\begin{aligned}
|\hat{X}_t - X_t(\hat{\theta}_n)| \;=\; & \left| f_{\hat{\theta}_n}(\hat{X}_{t-1}, Y_{t-1}) - f_{\hat{\theta}_n}(X_{t-1}(\hat{\theta}_n), Y_{t-1}) \right| \\
\leq\; & \omega_1 |\hat{X}_{t-1} - X_{t-1}(\hat{\theta}_n)| \leq \omega_1^{t-1} |\hat{X}_1 - X_1(\hat{\theta}_n)|.
\end{aligned}$$

Then, by using this and **(A2)**,

$$I_{n,3} \leq \frac{2}{\sqrt{n}} \sum_{t=1}^{n} V\gamma^t \left\| \sup_{\theta \in \Theta} \frac{2}{\sqrt{C(\eta_t)}} \right\| \leq \frac{4V}{\sqrt{\underline{f}}\sqrt{n}} \sum_{t=1}^{n} \gamma^t = o_P(1).$$

Thus, (A.3) holds for $i = 1$.

Secondly, we express

$$\max_{1 \leq k \leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} R_{t,2} - \frac{k}{n} \sum_{t=1}^{n} R_{t,2} \right| \leq \frac{2}{\sqrt{n}} \sum_{t=1}^{n} |R_{t,2}| \leq II_{n,1} + II_{n,2},$$

where

$$II_{n,1} = \max_{1 \leq k \leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} \epsilon_{t,1} \left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^n)}} \right) - \frac{k}{n} \sum_{t=1}^{n} \epsilon_{t,1} \left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^n)}} \right) \right|,$$

$$II_{n,2} = \max_{1 \leq k \leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=1}^{k} \epsilon_{t,1} \left(\frac{1}{\sqrt{C(\eta_t^n)}} - \frac{1}{\sqrt{C(\eta_t^0)}} \right) - \frac{k}{n} \sum_{t=1}^{n} \epsilon_{t,1} \left(\frac{1}{\sqrt{C(\eta_t^n)}} - \frac{1}{\sqrt{C(\eta_t^0)}} \right) \right|.$$

Similarly to the case of $I_{n,2}$, we have

$$\begin{aligned} II_{n,1} &\leq \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \left| \epsilon_{t,1} \left(\frac{1}{\sqrt{C(\hat{\eta}_t)}} - \frac{1}{\sqrt{C(\eta_t^n)}} \right) \right| \\ &\leq \frac{1}{\sqrt{n}} \frac{1}{\underline{f}^{3/2}} \sum_{t=1}^{n} |\epsilon_{t,1}(C(\hat{\eta}_t) - C(\eta_t^n))| \leq \frac{V}{\sqrt{n}} \frac{1}{\underline{f}^{3/2}} \sum_{t=1}^{n} |\epsilon_{t,1}|\gamma^t = o_P(1). \end{aligned}$$

Using Taylor's theorem, we have

$$C(\eta_t^n)^{-1/2} = C(\eta_t^0)^{-1/2} - \frac{1}{2} Z_t(\theta_0)(\hat{\theta}_n - \theta_0)^T \frac{\partial \eta_t^0}{\partial \theta} - \frac{1}{2}(\hat{\theta}_n - \theta_0)^T \left(\zeta_t(\theta_n^*) - \zeta_t(\theta_0) \right),$$

where θ_n^* is an intermediate point between $\hat{\theta}_n$ and θ_0, and $\zeta_t(\theta) = C'(\eta_t)C(\eta_t)^{-3/2} \frac{\partial \eta_t}{\partial \theta}$, so that

$$II_{n,2} \leq II'_{n,2} + II''_{n,2},$$

where

$$\begin{aligned} II'_{n,2} &= \sqrt{n}\|\hat{\theta}_n - \theta_0\| \max_{1 \leq k \leq n} \frac{k}{n} \left| \frac{1}{k} \sum_{t=1}^{k} \epsilon_{t,1}\zeta_t(\theta_0) - \frac{1}{n} \sum_{t=1}^{n} \epsilon_{t,1}\zeta_t(\theta_0) \right|, \\ II''_{n,2} &= \sqrt{n}\|\hat{\theta}_n - \theta_0\| \frac{2}{n} \sum_{t=1}^{n} |\epsilon_{t,1}| \, \|\zeta_t(\theta_n^*) - \zeta_t(\theta_0)\|. \end{aligned}$$

Since $\{\epsilon_{t,1}\zeta_t(\theta_0)\}$ is ergodic, $\sqrt{n}\|\hat{\theta}_n - \theta_0\| = O_P(1)$, and $E|\epsilon_{t,1}|\|\zeta_t(\theta_0)\| < \infty$ owing to **(A1)**, **(A2)**, **(A4)** and **(A11)**, we have $II'_{n,2} = o_P(1)$. Moreover, because

$$II''_{n,2} \leq \sqrt{n}\|\hat{\theta}_n - \theta_0\| \frac{2}{n} \sum_{t=1}^{n} |\epsilon_{t,1}| \sup_{\|\theta-\theta_0\| \leq \|\hat{\theta}_n - \theta_0\|} \|\zeta_t(\theta) - \zeta_t(\theta_0)\|$$

and $E\sup_{\theta\in\Theta} \|\zeta_t(\theta)\|^2 < \infty$ owing to **(A2)**, **(A4)** and **(A11)**, using **(A1)** and the dominated convergence theorem, we can have $II''_{n,2} = o_P(1)$ (cf., Proposition 5 in Lee, S., Lee, Y. and Chen, C. W. (2016)). Thus, (A.3) holds for $i = 2$.

Finally, using Taylor's theorem, we have

$$X_t(\hat{\theta}_n) = X_t(\theta_0) + (\hat{\theta}_n - \theta_0)^T \frac{\partial X_t(\theta_0)}{\partial\theta} + (\hat{\theta}_n - \theta_0)^T \left(\frac{\partial X_t(\theta_n^*)}{\partial\theta} - \frac{\partial X_t(\theta_0)}{\partial\theta} \right)$$

for some θ_n^* between $\hat{\theta}_n$ and θ_0. Then, similarly to the case of $II_{n,2}$, we can show that (A.3) holds for $i = 3$. Hence, (A.2) is verified. $\square$

Bibliography

Al-Osh, M. A. and Aly, E-E. A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Communications in Statistics-Theory & Methods* **21**(9), 2483-2492.

Alzaid, A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability* **27**, 314-324.

Berkes, I., Horvth, L. and Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p,q) models. *Statistics & Probability Letters* **70**(4), 263-273.

Chen, C. W. and Lee, S. (2016). Generalized Poisson autoregressive models for time series of counts. *Computational Statistics & Data Analysis* **99**, 51-67.

Christou, V. and Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis* **35**(1), 55-78.

Csörgö, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis.* Vol. 18. John Wiley & Sons Inc.

Davis, R. A. and Liu, H. (2016). Theory and inference for a class of observation-driven models with application to time series of counts. *Statistica Sinica* **26**(4), 1673-1707.

Davis, R. A. and Wu, R. (2009). A negative binomial model for time series of counts. *Biometrika* **96**(3), 735-749.

Diop, M. L. and Kengne, W. (2017). Testing parameter change in general integer-valued time series. *Journal of Time Series Analysis* **38**(6), 880-894.

Doukhan, P., Fokianos, K. and Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statistics & Probability Letters* **82**(5), 942-948.

Doukhan, P., Fokianos, K. and Tjøstheim, D. (2013). Correction to "On weak dependence conditions for Poisson autoregressions". *Statistics & Probability Letters* **83**(8), 1926-1927.

Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis* **27**(6), 923-942.

Fokianos, K. and Fried, R. (2010). Interventions in INGARCH processes. *Journal of Time Series Analysis* **31**(3), 210-225.

Fokianos, K. and Fried, R. (2012). Interventions in log-linear Poisson autoregression. *Statistical Modelling* **12**(4), 299-322.

Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression. *Journal of American Statistical Association* **104**(488), 1430-1439.

Fokianos, K., Gombay, E. and Hussein, A. (2014). Retrospective change detection for binary time series models. *Journal of Statistical Planning & Inferences* **145**, 102-112.

Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**(4), 605-637.

Franke, J., Kirch, C. and Kamgaing, J. T. (2012). Changepoints in times series of counts. *Journal of Time Series Analysis* **33**(5), 757-770.

Hudecová, Š. (2013). Structural changes in autoregressive models for binary time series. *Journal of Statistical Planning & Inferences* **143**(10), 1744-1752.

Hudecová, Š, Hudecová, M. and Meintanis, S. (2016). Change detection in INARCH time series of counts. *Nonparametric Statistics In: Cao R, Gonzalez Manteiga W, Romo J* **175**, 47-58.

Kang, J. and Lee, S. (2009). Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *Journal of Time Series Analysis* **30**(2), 239-258.

Kang, J. and Lee, S. (2014). Parameter change test for Poisson autoregressive models. *Scandinavian Journal of Statistics* **41**(4), 1136-1152.

Kim, H. and Lee, S. (2017). On first-order integer-valued autoregressive process with Katz family innovations. *Journal of Statistical Computation & Simulation* **87**(3), 546-562.

Lee, S. and Lee, J. (2015). Parameter change test for nonlinear time series models with GARCH type errors. *Journal of Korean Mathematical Society* **52**(23), 503-553.

Lee, S., Ha, J., Na, O. and Na, S. (2003). The cusum test for parameter change in time series models. *Scandinavian Journal of Statistics* **30**(4), 781-796.

Lee, S., Lee, Y. and Chen, C. W. (2016). Parameter change test for zero-inflated generalized Poisson autoregressive models. *Statistics* **50**(3), 1-18.

Lee, Y., Lee, S. and Tjøstheim, D. (2018). Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. *TEST* **27**(1), 52-69.

Lee, S., Seok, S. and Kim, D. (2018). Asymptotic normality and parameter change teest for zero-inflated general integer-valued GARCH models. Unpublished manuscript.

McKenzie, E. (1985). Some simple models for discrete variate time series1. *Journal of American Water Resource Association* **21**(4), 645-650.

McKenzie, E. (2003). Ch. 16. Discrete variate time series. *Handbook of statistics* **21**, 573-606.

Neumann, M. H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17**(4), 1268-1284.

Oh, H. and Lee, S. (2017a). On score vector- and residual-based CUSUM tests in ARMA-GARCH models. *Statistical Methods & Applications. Online published.*

Oh, H. and Lee, S. (2017b). Modified residual CUSUM test for location-scale time series models with heteroscedasticity. *Submitted for publication.*

Weiß, C. H. (2008). Thinning operations for modeling time series of counts-a survey. *AStA-Advances Statistical Analysis* **92**(3), 319-341.

Zhu, F. (2011). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis* **32**(1), 54-67.

Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *Journal of Mathematical Analysis & Applications* **389**(1), 58-71.

Zhu, F. (2012b). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning & Inferences* **142**(4), 826-839.

국문초록

이 재검토 연구에서는 영과잉 일반 정수값 시계열 모형의 모수 변화 검정의 문제를 다루었다. 일반 정수값 시계열 모형은 관측값의 조건부 밀도 함수가 영과잉 일변량 지수족을 따르는 모형이다. 이 논문은 Lee and Lee (2018)의 연구에 기초하여 표준화된 잔차 기반 CUSUM 검정에 초점을 두고 있다. 또한, 표준화된 잔차 기반 CUSUM 검정의 통계량이 귀무가설 하에서 브라우니안 브릿지의 함수로 분포 수렴함을 보였다.

주요어: 이산형 시계열, 정수값 GARCH 모형, 영과잉 지수족, 모수 변화 검정, CUSUM 검정.

학번: 2017-26674