Testing for the Conditional Geometric Mixture Distribution

JIN SEO CHO
School of Economics
Yonsei University
Email: jinseocho@yonsei.ac.kr

JIN SEOK PARK
Department of Economics
Yonsei University
Email: js-ism@naver.com

SANG WOO PARK
Department of Economics
Yonsei University
Email: psw199010@gmail.com

June 20, 2018

Abstract

This study examines the mixture hypothesis of conditional geometric distributions using a likelihood ratio (LR) test statistic based on that used for unconditional geometric distributions. As such, we derive the null limit distribution of the LR test statistic and examine its power performance. In addition, we examine the interrelationship between the LR test statistics used to test the geometric and exponential mixture hypotheses. We also examine the performance of the LR test statistics under various conditions and confirm the main claims of the study using Monte Carlo simulations.

Keywords: mixture of conditional geometric distributions, likelihood ratio test, unobserved heterogeneity, Gaussian stochastic process.

JEL Classification: C12, C41, C80.

Acknowledgements: The author is most grateful to the editor, Jin Lee. Cho acknowledges the support of the Yonsei University Future-leading Research Initiative of 2017 (2017-22-0090).

1 Introduction

Duration data have been a popular research topic. For example, Van den Berg and Ridder (1998) empirically examine unemployment and job duration in the Netherlands using search theory, and Kennan (1985) examines contract strike duration data in U.S. manufacturing industries.

Unobserved heterogeneity is unavoidable in most empirical duration data analyse, which means handling it correctly is crucial to obtaining correct inferences from the data. Van den Berg and Ridder (1998) specify a finite mixture model for their data, and Kennan (1985) employs various tests to confirm the absence of unobserved heterogeneity in his data.

Empirical duration data are often available in the form of grouped observations. Although data analyses are conducted assuming continuous observations, available observations are measured in days, weeks, months, and so on. This implies that analyzing grouped duration data using a continuous model can lead to a misspecified model estimation, and that inferencing based on unobserved heterogeneity can be misleading.

Therefore, the main goal of this study is to infer unobserved heterogeneity for grouped duration data. We achieve this goal by testing a finite mixture hypothesis of two conditional geometric distributions. According to search theory (*e.g.*, Van den Berg and Ridder, 1998), the equilibrium duration conditionally follows an exponential distribution on other conditioning variables, unless unobserved heterogeneity is involved. Furthermore, as we discuss below, grouping exponential data yields geometrically distributed observations. Therefore, a finite mixture of two conditional geometric distributions can be a proper model for grouped duration data in the presence of unobserved heterogeneity.

The approach of the current study extends the methodology in Cho and Han (2009) by applying the method in Cho and White (2007). Cho and Han (2009) provide a methodology for testing a mixture hypothesis of two unconditional geometric distributions, applying the likelihood ratio (LR) test statistic. The current work extends their approach by supposing a mixture of two conditional geometric distributions and applying the same LR testing principle in Cho and Han (2009). In this way, we extend the applicability of the geometric mixture hypothesis.

Another goal of this study is to examine the interrelationship between the null limit distributions of LR test statistics. Cho and White (2010) derive the null limit distribution of the LR test statistic that tests the mixture hypothesis of two exponential distributions. We achieve our second goal by examining how their null limit distribution is associated with ours, which is obtained from the LR test statistic that tests the geometric mixture hypothesis. We suppose different data sets with different grouping bin sizes and examine how the null limit distribution of the LR test statistic responds as the size decreases. From this, we obtain a

regular interrelationship between the null limit distributions.

The remainder of the paper proceeds as follows. Section 2 provides the DGP and model environments, as well as the null limit distribution of the LR test statistic that tests the conditional geometric mixture hypothesis. In the same section, we examine the association between the null limit distributions of the LR test statistics used to test the geometric and exponential mixture hypotheses. In Section 3, we conduct Monte Carlo experiments and examine the finite performance of the LR test statistic. Section 4 provides concluding remarks. All mathematical proofs are available in the Appendix.

2 Mixtures of Conditional Geometric Distributions

2.1 Motivation, Data Generating Process (DGP), and Model

Economic theories on duration data are often associated with exponential distributions. For example, using search theory, Van den Berg and Ridder (1998) show that unemployment and job duration follow exponential distributions. Specifically, if we let D_t be unemployment or job duration, the following probability density function (PDF) becomes the conditional distribution of $D_t|\mathbf{X}_t$:

$$f_o(d|\mathbf{x};\boldsymbol{\beta}_*,\delta_*) \equiv \delta_* h(\mathbf{x};\boldsymbol{\beta}_*)\exp\{-\delta_* h(\mathbf{x};\boldsymbol{\beta}_*)d\}, \tag{1}$$

according to search theory, where $\mathbf{X}_t$ is a $k \times 1$ vector of conditioning variables.

Nevertheless, if the duration data are contaminated by unobserved heterogeneity, the conditional exponential distribution yields a misspecified model, which is a common problem in most empirical studies.

Therefore, models for unobserved heterogeneity are often specified as well, often using mixture models. For example, Nickell (1979) and Van den Berg and Ridder (1998) employ a finite mixture for unobserved heterogeneity, and Lancaster (1979) assumes a gamma distribution. Please refer to Lancaster (1992) for other mixture assumptions related to unobserved heterogeneity. Finite mixture models specify the distribution of $D_t|\mathbf{X}_t$ as

$$f_a(d|\mathbf{x};\boldsymbol{\beta},\delta_1,\delta_2) \equiv \pi f_o(d|\mathbf{x};\boldsymbol{\beta},\delta_1) + (1-\pi) f_o(d|\mathbf{x};\boldsymbol{\beta},\delta_2), \tag{2}$$

where $(\pi,\boldsymbol{\beta},\delta) \in [0,1] \times \mathbf{B} \times \Delta$, and $\mathbf{B} \times \Delta$ is a convex and compact set in $\mathbb{R}^{1+d}$ $(d \in \mathbb{N})$. Note that (1) is the DGP for $D_t|\mathbf{X}_t$ in the absence of unobserved heterogeneity, whereas (2) is a mixture model for $D_t|\mathbf{X}_t$ that accommodates unobserved heterogeneity.

Furthermore, most duration data used in empirical studies are grouped observations. For example, Kennan (1985) examines contract strike duration data in U.S. manufacturing industries that are measured in

days, and Van den Berg and Ridder (1998) examine unemployment and job duration data in the Netherlands, which are measured in months. Thus, estimating the parameters using (2) can be misleading, even though model (2) is correctly specified for $D_t|\mathbf{X}_t$ (*e.g.*, Ryu, 1995).

We capture this grouping feature by letting $(Y_t, \mathbf{X}_t')' \in \mathbb{N} \times \mathbb{R}^k$ be a set of available observations, where $Y_t \equiv \lceil D_t \rceil$ and $\lceil x \rceil := \min\{a \in \mathbb{N} : a \geq x\}$. Here, model (2) is misspecified for $Y_t|\mathbf{X}_t$. Given that $Y_t \in \mathbb{N}$, we need to employ a conditional probability mass function (PMF) rather than the conditional PDF. Note that if an exponential random variable is grouped according to our plan, it follows a geometric distribution. More specifically, if $D_t|\mathbf{X}_t$ is distributed according to (1), the conditional cumulative distribution function (CDF) of $Y_t|\mathbf{X}_t$ is obtained as

$$F(y|\mathbf{x}; \boldsymbol{\beta}_*, \delta_*) = 1 - \exp\{-\delta_* h(\mathbf{x}; \boldsymbol{\beta}_*) y\},$$

so that the conditional PMF of $Y_t|\mathbf{X}_t$ is obtained as

$$F(y|\mathbf{x}; \boldsymbol{\beta}_*, \delta_*) - F(y-1|\boldsymbol{\beta}_*, \delta_*) = [1 - \exp\{-\delta_* h(\mathbf{x}; \boldsymbol{\beta}_*)\}] \times [\exp\{-\delta_* h(\mathbf{x}; \boldsymbol{\beta}_*)\}]^{y-1}. \tag{3}$$

By this feature, a finite mixture model of two geometric distributions becomes a proper model for $Y_t|\mathbf{X}_t$ in the presence of unobserved heterogeneity:

$$\mathcal{M}_a \equiv \{g_a(\,\cdot\,|\,\cdot\,; \pi, \boldsymbol{\beta}, \delta_1, \delta_2) : (\pi, \boldsymbol{\beta}, \delta_1, \delta_2) \in [0,1] \times \mathbf{B} \times \Delta \times \Delta\},$$

where for $y \in \mathbb{N}$, $g_a(y|\mathbf{x}; \pi, \boldsymbol{\beta}, \delta_1, \delta_2) \equiv \pi g_o(y|\mathbf{x}; \boldsymbol{\beta}, \delta_1) + (1-\pi) g_o(y|\mathbf{x}; \boldsymbol{\beta}, \delta_2)$ and

$$g_o(y|\mathbf{x}; \boldsymbol{\beta}, \delta) \equiv [1 - \exp\{-\delta h(\mathbf{x}; \boldsymbol{\beta})\}] \times [\exp\{-\delta h(\mathbf{x}; \boldsymbol{\beta})\}]^{y-1}.$$

The main goal of this study is to test the mixture hypothesis of conditional geometric distributions. Then, we associate this result with that of conditional exponential distributions obtained in the absence of data grouping.

Before proceeding, several remarks are provided on our approach and its association with the literature. First, Cho and White (2007, 2010) and Cho and Han (2009) examine the null limit distribution of the likelihood ratio (LR) test statistic used to test the mixture hypothesis of regular conditional distributions. In particular, Cho and White (2010) examine a mixture hypothesis of conditional exponential/Weibull distributions, whereas Cho and Han (2009) examine a mixture of unconditional geometric distributions. They both derive the null limit distributions of the LR test statistics under their specific environments. The current study links their independent studies by supposing a mixture model of conditional geometric distributions, which have a regular interrelationship between their null limit distributions. Second, we reparameterize the

original model $\mathcal{M}_a$ for analytical convenience. Cho and White (2010) point out that if $\mathcal{M}_a$ is reparameterized as

$$\mathcal{M}'_a \equiv \{g_a(\cdot\,|\,\cdot\,;\pi,\boldsymbol{\beta},\alpha_1\delta_*,\alpha_2\delta_*) : (\pi,\boldsymbol{\beta},\alpha_1,\alpha_2) \in [0,1]\times\mathbf{B}\times A\times A\},$$

deriving the null limit distribution of the LR test statistic becomes more straightforward, where A is such that $\Delta = \{\alpha\delta_* : \alpha \in A\}$. By the invariance principle, the same LR test statistic is obtained from both $\mathcal{M}'_a$ and $\mathcal{M}_a$. We follow their convention and discuss their usage without loss of generality whenever it is convenient for our analysis.

2.2 Likelihood Ratio Test for Unobserved Heterogeneity

Following Cho and White (2007, 2010) and Cho and Han (2009), the goal of this study is achieved by estimating the parameters of interest. The absence of unobserved heterogeneity is also presented using the same parameters. Specifically, if we let $(\pi_*,\boldsymbol{\beta}_*,\delta_{1*},\delta_{2*})$ be the parameter describing the DGP of $Y_t|\mathbf{X}_t$, we have that $g_a(\cdot\,|\,\cdot\,;\pi_*,\boldsymbol{\beta}_*,\delta_{1*},\cdot) = g_o(\cdot\,|\,\cdot\,;\boldsymbol{\beta}_*,\delta_*)$, provided that $\pi_* = 1$ and $\delta_{1*} = \delta_*$. Thus, the conditional geometric PMF becomes the DGP of $Y_t|\mathbf{X}_t$. Here, δ_{2*} is irrelevant to the DGP, that is, δ_{2*} is not identified. Analogously, if $\delta_{1*} = \delta_{2*} = \delta_*$, $g_a(\cdot\,|\,\cdot\,;\cdot,\boldsymbol{\beta}_*,\delta_{1*},\delta_{2*}) = g_o(\cdot\,|\,\cdot\,;\boldsymbol{\beta}_*,\delta_*)$, so that π_* is not identified. Finally, if $\pi_* = 0$ and $\delta_{2*} = \delta_*$, $g_a(\cdot\,|\,\cdot\,;\cdot,\boldsymbol{\beta}_*,\delta_{1*},\delta_{2*}) = g_o(\cdot\,|\,\cdot\,;\boldsymbol{\beta}_*,\delta_*)$ and δ_{1*} is not identified. This case is parallel to that in which $\pi_* = 1$ and $\delta_{1*} = \delta_*$. On the other hand, if $\pi_* \in (0,1)$ and $\delta_{1*} \neq \delta_{2*}$, the mixture model must be appropriate for the distribution of $Y_t|\mathbf{X}_t$. Therefore, the following is stated as our proper set of null and alternative hypotheses:

$$\mathcal{H}_0 : \pi_* = 1 \text{ and } \delta_{1*} = \delta_*;\ \ \delta_{1*} = \delta_{2*} = \delta_*;\ \ \pi_* = 0 \text{ and } \delta_{2*} = \delta_*, \quad \text{versus}$$

$$\mathcal{H}_1 : \pi_* \in (0,1) \text{ and } \delta_{1*} \neq \delta_{2*}.$$

The current study tests $\mathcal{H}_0$ versus $\mathcal{H}_1$ using the following LR test statistic:

$$LR_n(A) \equiv 2\left\{\sum_{t=1}^{n} \ln g_a(Y_t|\mathbf{X}_t;\hat{\pi}_n,\hat{\boldsymbol{\beta}}_n,\hat{\delta}_{1n},\hat{\delta}_{2n}) - \sum_{t=1}^{n} \ln g_o(Y_t|\mathbf{X}_t;\hat{\boldsymbol{\beta}}_{on},\hat{\delta}_{on})\right\},$$

where n is the sample size, and $(\hat{\boldsymbol{\beta}}_{on},\hat{\delta}_{on})$ and $(\hat{\pi}_n,\hat{\boldsymbol{\beta}}_n,\hat{\delta}_{1n},\hat{\delta}_{2n})$ are the maximum-likelihood estimators (MLEs) obtained under the null and alternative model assumptions, respectively. Here, the LR test statistic is indexed by the parameter space A for α, which is part of $\mathcal{M}'_a$. As discussed below, the null limit distribution of the LR test statistic is influenced by the parameter space A. We accommodate this feature by letting the LR test statistics be indexed by the parameter space A.

Here, we examine the LR test statistic because other test statistics are difficult to apply in this context. Note that Davies's (1977, 1987) identification problem is present in multiple ways under $\mathcal{H}_0$. That is, if $\pi_* = 1$ (resp. $\pi_* = 0$), δ_{2*} (resp. δ_{1*}) is not identified, which yields Davies's (1977, 1987) identification problem. Furthermore, if $\delta_{1*} = \delta_{2*}$, π_* is not identified, implying Davies's (1977, 1987) identification problem in a different manner. This trifold identification problem makes it difficult to apply Wald's (1943) testing principle, although applying the LR testing principle is straightforward. Thus, we examine the LR test statistic in this study.

Multifold identification problems are often observed in the literature. For example, Cho and Ishida (2012), Baek, Cho, and Phillips (2015), and Cho and Phillips (2018) test polynomial model hypotheses using power transformations, and multifold identification problems arise under their null hypotheses. As another example, Cho and White (2011a, 2011b), Cho, Ishida, and White (2011, 2014), and White and Cho (2011) test correct model assumptions using artificial neural network models, and they too observe multifold identification problems under their correct model assumptions. The aforementioned studies handle multifold identification problems by applying the LR testing principle when testing their hypotheses. The current study extends the existing literature by applying the LR testing principle to testing the mixture hypothesis.

2.2.1 Asymptotic Null Distribution

Cho and White (2007) examine the LR test statistic in a general context and apply the methodology in Andrews (1999, 2001) to show that the LR test statistic weakly converges to a function of a Gaussian process under $\mathcal{H}_0$. Their result also holds for our problem. The following theorem reveals the null limit distribution of the LR test statistic.

Theorem 1. *Given Assumptions 1 to 4 in the Appendix, if* $\inf A > 1/2$*,*

$$LR_n(A) \Rightarrow \mathcal{LR}(A) \equiv \sup_{\alpha \in A} \left(\max[0, \mathcal{G}(\alpha)]\right)^2 \tag{4}$$

under $\mathcal{H}_0$*, where* $\mathcal{G}(\cdot)$ *is a Gaussian process such that for each* α *and* $\alpha' \in A$*,*

$$\mathbb{E}[\mathcal{G}(\alpha)\mathcal{G}(\alpha')] = \frac{\rho(\alpha, \alpha')}{\sqrt{\rho(\alpha, \alpha)}\sqrt{\rho(\alpha', \alpha')}}, \tag{5}$$

where $\rho(\alpha, \alpha') \equiv A(\alpha, \alpha') - \boldsymbol{B}(\alpha)'\boldsymbol{C}^{-1}\boldsymbol{B}(\alpha')$*;*

$$A(\alpha, \alpha') \equiv \mathbb{E}\left[\frac{[Q_t(\alpha) - Q_t(1)][Q_t(\alpha') - Q_t(1)]}{[1 - Q_t(1)][Q_t(1) - Q_t(\alpha)Q_t(\alpha')]}\right];$$

$$\boldsymbol{B}(\alpha) \equiv \mathbb{E}\left[\left[\frac{[Q_t(\alpha) - Q_t(1)]}{[Q_t(1) - 1][Q_t(\alpha) - 1]}\right]\left[\begin{array}{c} \delta_* h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \\ \delta_* \nabla_{\boldsymbol{\beta}} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \end{array}\right]\right];$$

$$\boldsymbol{C} \equiv \mathbb{E}\left[\frac{Q_t(1)\delta_*^2}{[Q_t(1)-1]^2}\begin{bmatrix} h(\boldsymbol{X}_t;\boldsymbol{\beta}_*)^2 & h(\boldsymbol{X}_t;\boldsymbol{\beta}_*)\nabla'_{\boldsymbol{\beta}}h(\boldsymbol{X}_t;\boldsymbol{\beta}_*) \\ h(\boldsymbol{X}_t;\boldsymbol{\beta}_*)\nabla_{\boldsymbol{\beta}}h(\boldsymbol{X}_t;\boldsymbol{\beta}_*) & \nabla_{\boldsymbol{\beta}}h(\boldsymbol{X}_t;\boldsymbol{\beta}_*)\nabla'_{\boldsymbol{\beta}}h(\boldsymbol{X}_t;\boldsymbol{\beta}_*) \end{bmatrix}\right];$$

and $Q_t(\alpha) \equiv \exp[\alpha\delta_* h(\boldsymbol{X}_t;\boldsymbol{\beta}_*)]$. □

A number of remarks are warranted. First, the null limit distribution of the LR test statistic depends on the parameter space A. Given that the functional of the Gaussian process $\mathcal{G}(\cdot)$ is maximized over A, if we use a different parameter space for A, a different null limit distribution is obtained.

Second, it is more convenient to use $\mathcal{M}'_a$ than $\mathcal{M}_a$ to obtain the null limit distribution. Here, the result in Theorem 1 is derived from the Gaussian process $\mathcal{G}(\cdot)$ defined on A, which is associated with $\mathcal{M}'_a$. That is why we reparameterize $\mathcal{M}_a$ as $\mathcal{M}'_a$.

Third, the covariance structure in (5) generalizes the result in Cho and Han (2009). If we let $h(\mathbf{X}_t;\boldsymbol{\beta}_*) \equiv 1$ so that for each α and α', we can denote $Q_t(\alpha)$, $Q_t(\alpha')$, and $Q_t(1)$ as $1-p$, $1-p'$, and $1-p_*$ respectively, it follows that

$$\rho(\alpha,\alpha') = \frac{(p-p_*)^2(p'-p_*)^2}{pp'(1-p_*)p_*[(1-p_*)-(1-p)(1-p')]}.$$

Note that this covariance structure is identical to that of Cho and Han (2009, p. 51), treating their theorem 1 as a special case of Theorem 1 here.

Fourth, Theorem 1 is consistent with other results in the literature. Table 1 summarizes the relevant literature on testing mixture hypotheses under different DGP and model conditions. The studies in Table 1 also provide methodologies that consistently yield the asymptotic critical values of their test statistics. When characterizing the null limit distributions of their test statistics, these studies all exploit Gaussian processes, as Theorem 1 does.

Fifth, the null limit distribution of the LR test statistic is not distribution-free. Note that the covariance structure of $\mathcal{G}(\cdot)$ is affected by the distribution of $\mathbf{X}_t$: for different distributions of $\mathbf{X}_t$, different functional forms are obtained for $\rho(\cdot,\cdot)$. In addition, the distribution of $\mathbf{X}_t$ is often unknown. In this case, the closed form of $\rho(\cdot,\cdot)$ is difficult to obtain. Furthermore, the functional form of $\rho(\cdot,\cdot)$ depends on $h(\mathbf{X}_t;\boldsymbol{\beta})$. Different specifications for $h(\mathbf{X}_t;\boldsymbol{\beta})$ lead to various functional forms for $\rho(\cdot,\cdot)$, and thus, to different asymptotic critical values. This aspect implies that the asymptotically conservative critical values advocated by Davies (1977) and Piterbarg (1996) are difficult to apply here, because their critical values are derived using the functional form of $\rho(\cdot,\cdot)$.

Sixth, the mixture hypothesis in our context needs to be tested nonparametrically. The weighted bootstrap method of Hansen (1996) is useful for this purpose. Cho, Cheong, and White (2011) implement the weighted bootstrap method for the LR test statistic described in Cho and White (2010). The procedure of

the weighted bootstrap method is described here for the current study to be self-contained:

(1) For each grid point of $\alpha \in A$, compute $\hat{S}_{nt}(\alpha) \equiv \{\hat{D}_{nt}(\alpha)\}^{-\frac{1}{2}} \hat{W}_{nt}(\alpha)$, where

$$\hat{W}_{nt}(\alpha) \equiv [1 - \hat{R}_{nt}(\alpha)] - \hat{\mathbf{U}}'_{nt} \left[\sum_{t=1}^{n} \hat{\mathbf{U}}_{nt} \hat{\mathbf{U}}'_{nt}\right]^{-1} \left[\sum_{t=1}^{n} \hat{\mathbf{U}}_{nt}[1 - \hat{R}_{nt}(\alpha)]\right],$$

$$\hat{D}_{nt}(\alpha) \equiv n^{-1} \sum_{t=1}^{n} [1 - \hat{R}_{nt}(\alpha)]^2 - n^{-1} \sum_{t=1}^{n} [1 - \hat{R}_{nt}(\alpha)] \hat{\mathbf{U}}'_{nt} \left[\sum_{t=1}^{n} \hat{\mathbf{U}}_{nt} \hat{\mathbf{U}}'_{nt}\right]^{-1} \sum_{t=1}^{n} \hat{\mathbf{U}}_{nt}[1 - \hat{R}_{nt}(\alpha)],$$

$$\hat{R}_{nt}(\alpha) \equiv g_o(Y_t | \mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \alpha\hat{\delta}_{on}) / g_o(Y_t | \mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on}),$$

and $\hat{\mathbf{U}}_{nt} \equiv \nabla_{(\boldsymbol{\beta},\delta)} \ln[g_o(Y_t | \mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on})]$.

(2) Generate $Z_t^j \sim$ IID $N(0,1)$, for $t = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, J$, and compute the empirical distribution of

$$\mathcal{LR}_n^j(A) \equiv \sup_{\alpha \in A} \left(\max \left[0, \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \hat{S}_{nt}(\alpha) Z_t^j \right] \right)^2$$

by iterating J times.

(3) The empirical p-value is computed as $\hat{p}_n^J \equiv \frac{1}{J} \sum_{j=1}^{J} I[LR_n(A) < \mathcal{LR}_n^j(A)]$.

According to Hansen (1996), $\hat{p}_n^J$ asymptotically follows a uniform distribution on $[0,1]$ under $\mathcal{H}_0$. Otherwise, it should converge to zero in probability. Thus, if $\hat{p}_n^J$ is less than the level of significance, we reject the null hypothesis.

2.2.2 Asymptotic Power of the LR Test Statistic

The LR test statistic has asymptotic power when the mixture model is correctly specified.

Theorem 2. *Given Assumptions 1 to 4 and $\mathcal{H}_1$, if $\mathcal{M}_a$ is correctly specified and* $\inf A > 1/2$*, then for any sequence $\{c_n\}$ such that $c_n = o(n)$, $\mathbb{P}[LR_n(A) \geq c_n] \to 1$ as $n \to \infty$.* □

Although Theorem 2 holds straightforwardly by the Kullback–Leibler information criterion (KLIC), we prove it in the Appendix.

Despite its consistency, a careful interpretation of Theorem 2 is needed. First, if $\mathcal{M}_a$ is misspecified, the consistency of Theorem 2 may not hold. Cho and White (2008) examine a general theory of testing the mixture hypothesis of misspecified models and provide regularity conditions under which the LR test statistic has a nondegeneracy property and asymptotic power under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. Theorem 2 is valid only when $\mathcal{M}_a$ is correctly specified. Second, the LR test statistic may appear to have unobserved

heterogeneity, even under $\mathcal{H}_0$, because $h(\cdot;\beta)$ is misspecified. Note that $h(\cdot;\beta)$ contributes to $\mathbb{E}[Y_t|\mathbf{X}_t]$. Therefore, we recommend testing for the correct conditional mean specification first to avoid this consistent type-I error, before applying the LR test statistic.

2.3 Asymptotic Behavior of the LR Test Statistic Local to the Exponential Distribution

There is a regular interrelationship between the null limit distribution in Theorem 1 and that of the LR test statistic used to test for the conditional exponential mixture hypothesis given in Cho and White (2010, theorem 1). In Section 2.1, we showed that geometrically distributed observations are generated by grouping exponentially distributed observations by supposing a unitary bin size. In this subsection, we suppose a different bin size and derive the null limit distribution of the LR test statistic. Thus, we examine the null limit distribution when the grouping bin size is extremely small. In this case, the null limit distribution of the LR test statistic can be thought of as an approximation of the LR test statistic used to test the conditional exponential mixture hypothesis.

This aspect implies that if the grouping bin size is sufficiently small, the asymptotic critical values in Cho and White (2010) can be used for our inference purpose without implementing the weighted bootstrap method. Note that Cho and White (2010) show that the null limit distribution of their LR test statistic is asymptotically distribution-free, providing a straightforward simulation method that delivers the asymptotic critical values. Thus, applying their asymptotic critical values can be efficient for our inferencing purpose.

Here, we consider different bin sizes. For the same D_t given in (1), if we suppose that the grouping bin size is given as $\Delta > 0$, then

$$F(d|\mathbf{x};\boldsymbol{\beta}_*,\delta_*) - F(d-\Delta|\boldsymbol{\beta}_*,\delta_*) = [1-\exp\{-\delta_*\Delta h(\mathbf{x};\boldsymbol{\beta}_*)\}] \times [\exp\{-\delta_* h(\mathbf{x};\boldsymbol{\beta}_*)\}]^{d-\Delta}. \quad (6)$$

If we further let $\Delta \equiv 1/\omega$, (6) is converted into the following geometric random variable:

$$F[k/\omega|\mathbf{x};\boldsymbol{\beta}_*,\delta_*] - F[(k-1)/\omega|\boldsymbol{\beta}_*,\delta_*] = [1-\exp\{-\delta_* h(\mathbf{x};\boldsymbol{\beta}_*)/\omega\}] \times [\exp\{-\delta_* h(\mathbf{x};\boldsymbol{\beta}_*)/\omega\}]^{k-1},$$

where $k = 1, 2, \ldots$ Note that this is the conditional PMF of $Y_t^\omega \equiv \lceil D_t \times \omega \rceil$ on $\mathbf{X}_t$: for $y = 1, 2, \ldots$, the conditional PMF of Y_t^ω is

$$g_o^\omega(y|\mathbf{x};\boldsymbol{\beta}_*,\delta_*) \equiv [1-\exp\{-\delta_*^\omega h(\mathbf{x};\boldsymbol{\beta}_*)\}] \times [\exp\{-\delta_*^\omega h(\mathbf{x};\boldsymbol{\beta}_*)\}]^{y-1},$$

where $\delta_*^\omega \equiv \delta_*/\omega$. The only difference between this and the conditional distribution in (3) is the adjustment of the location parameter δ_* to accommodate the influence of ω. If $\omega = 1$, the same conditional PMF is

obtained from (6), and the LR test statistic $LR_n(A)$ in Section 2.2 is a special case of the following LR test statistic:

$$LR_n^{\omega}(A) \equiv 2 \left\{ \sum_{t=1}^{n} \ln g_a^{\omega}(Y_t^{\omega}|\mathbf{X}_t; \hat{\pi}_n, \hat{\boldsymbol{\beta}}_n, \hat{\delta}_{1n}^{\omega}, \hat{\delta}_{2n}^{\omega}) - \sum_{t=1}^{n} \ln g_o^{\omega}(Y_t^{\omega}|\mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on}^{\omega}) \right\},$$

where for $y \in \mathbb{N}$, $g_a(y|\mathbf{x}; \pi, \boldsymbol{\beta}, \delta_1^{\omega}, \delta_2^{\omega}) \equiv \pi g_o^{\omega}(y|\mathbf{x}; \boldsymbol{\beta}, \delta_1^{\omega}) + (1-\pi) g_o^{\omega}(y|\mathbf{x}; \boldsymbol{\beta}, \delta_2^{\omega})$, and $(\hat{\pi}_n, \hat{\boldsymbol{\beta}}_n, \hat{\delta}_{1n}^{\omega}, \hat{\delta}_{2n}^{\omega})$ and $(\hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on}^{\omega})$ are the MLEs obtained by maximizing the alternative likelihood function $\sum_{t=1}^{n} \ln g_a^{\omega}(\, Y_t^{\omega}|\mathbf{X}_t; \cdot)$ and the null likelihood function $\sum_{t=1}^{n} \ln g_o^{\omega}(Y_t^{\omega}|\mathbf{X}_t; \cdot)$, respectively. Note that $LR_n(A)$ is obtained from $LR_n^{\omega}(A)$ by letting ω be unity.

We contain the null limit distribution of $LR_n^{\omega}(A)$ in the following theorem.

Theorem 3. *Given Assumptions 1 to 4 in the Appendix, if* $\inf A > 1/2$*, for each* ω*,*

$$LR_n^{\omega}(A) \Rightarrow \mathcal{LR}_{\omega}(A) \equiv \sup_{\alpha \in A} \left(\max[0, \mathcal{G}_{\omega}(\alpha)]\right)^2 \tag{7}$$

under $\mathcal{H}_0$*, where* $\mathcal{G}_{\omega}(\cdot)$ *is a Gaussian process such that for each* α *and* $\alpha' \in A$*,*

$$\mathbb{E}[\mathcal{G}_{\omega}(\alpha)\mathcal{G}_{\omega}(\alpha')] = \frac{\rho_{\omega}(\alpha, \alpha')}{\sqrt{\rho_{\omega}(\alpha, \alpha)}\sqrt{\rho_{\omega}(\alpha', \alpha')}},$$

where $\rho_{\omega}(\alpha, \alpha') \equiv A_{\omega}(\alpha, \alpha') - \boldsymbol{B}_{\omega}(\alpha)' \boldsymbol{C}_{\omega}^{-1} \boldsymbol{B}_{\omega}(\alpha')$*;*

$$A_{\omega}(\alpha, \alpha') \equiv \mathbb{E}\left[\frac{[Q_t(\alpha/\omega) - Q_t(1/\omega)][Q_t(\alpha'/\omega) - Q_t(1/\omega)]}{[Q_t(1/\omega) - 1][Q_t(\alpha/\omega)Q_t(\alpha'/\omega) - Q_t(1/\omega)]} \right];$$

$$\boldsymbol{B}_{\omega}(\alpha) \equiv \mathbb{E}\left[\frac{\delta_*^{\omega}[Q_t(\alpha/\omega) - Q_t(1/\omega)]}{[Q_t(1/\omega) - 1][Q_t(\alpha/\omega) - 1]} \begin{bmatrix} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \\ \nabla_{\boldsymbol{\beta}} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \end{bmatrix} \right]; \quad \textit{and}$$

we let $\boldsymbol{C}_{\omega}$ *be defined as*

$$\mathbb{E}\left[\frac{(\delta_*^{\omega})^2 Q_t(1/\omega)}{[Q_t(1/\omega) - 1]^2} \begin{bmatrix} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*)^2 & h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \nabla_{\boldsymbol{\beta}}' h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \\ h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \nabla_{\boldsymbol{\beta}} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) & \nabla_{\boldsymbol{\beta}} h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \nabla_{\boldsymbol{\beta}}' h(\boldsymbol{X}_t; \boldsymbol{\beta}_*) \end{bmatrix} \right]. \qquad \square$$

Note that the covariance structure $\rho_{\omega}(\alpha, \alpha')$ is identical to $\rho(\omega, \omega')$ in Theorem 1 if ω is equal to one. Because Theorem 2 treats a more general case than Theorem 1, Theorem 2 is proved in the Appendix, and we omit proving Theorem 1.

The goal of this section is achieved by letting the grouping bin size Δ tend to zero (or $\omega \to \infty$). In the following theorem, we provide the null limit distribution of the LR test statistic obtained by letting $\omega \to \infty$.

Theorem 4. *Given Assumptions 1 to 4 in the Appendix,* if $\inf A > 1/2$ *and there is no conditioning variable* $\boldsymbol{X}_t$*, then*

$$\lim_{\omega \to \infty} \mathcal{LR}_{\omega}(A) \Rightarrow \mathcal{LR}_{\infty}(A) \equiv \sup_{\alpha \in A} \left(\max[0, \mathcal{G}_{\infty}(\alpha)]\right)^2 \tag{8}$$

under $\mathcal{H}_0$, where $\mathcal{G}_\infty(\cdot)$ is a standard Gaussian process such that for each α and $\alpha' \in A$,

$$\mathbb{E}[\mathcal{G}_\infty(\alpha)\mathcal{G}_\infty(\alpha')] = \frac{(2\alpha-1)^{1/2}(2\alpha'-1)^{1/2}}{\alpha+\alpha'-1}. \qquad \square$$

Theorem 4 yields a regular relationship between the null limit distributions of the LR test statistics used to test the geometric and exponential mixture hypotheses. The covariance structure $\mathbb{E}[\mathcal{G}_\infty(\alpha)\mathcal{G}_\infty(\alpha')]$ is identical to that obtained by Cho and White (2010, theorem $1(i)$) as the null limit distribution of the LR test statistic used to test the exponential mixture hypothesis. Therefore, if the grouping bin size Δ is sufficiently small, the null limit distribution given by Cho and White (2010) can approximate that of the LR test statistic used to test the geometric mixture hypothesis.

The reasoning behind Theorem 4 is straightforward. As ω approaches infinity, the probability mass generated by grouping continuous data decreases, such that the CDF of Y_t^ω approaches that of the exponential random variable. As a result, the asymptotic critical value of the LR test statistic used to test the geometric mixture hypothesis approaches that of the LR test statistic used to test the exponential mixture hypothesis.

Theorem 4 is comparable to the quasi-maximum likelihood (QML) estimation of grouped exponential random variables. Ryu (1995) analyzes the QML estimator obtained by maximizing the quasi-likelihood function of an exponential distribution and grouped exponential random observations. The QML estimator is not consistent for the parameters in the DGP, although the bias is asymptotically negligible if the size of bin Δ is sufficiently small. This result agrees with that in Theorem 4.

Nevertheless, a regular interrelationship between the null limit distributions of the LR test statistics is not established when the conditioning variable $\mathbf{X}_t$ exists. By estimating the unknown parameter $\boldsymbol{\beta}_*$ in $h(\mathbf{X}_t;\boldsymbol{\beta}_*)$, it introduces an estimation error that modifies the asymptotic covariance structure of the Gaussian process $\mathcal{G}_\infty(\cdot)$. Specifically, when the regularity conditions hold such that we can apply Lebesgue's dominated convergence theorem, it follows that

$$\lim_{\omega\to\infty}\rho_\omega(\alpha,\alpha') = \frac{(\alpha-1)(\alpha'-1)}{\alpha+\alpha'-1} - \frac{(\alpha-1)(\alpha'-1)}{\alpha\alpha'}\mathbf{B}_\infty'\mathbf{C}_\infty^{-1}\mathbf{B}_\infty,$$

where

$$\mathbf{B}_\infty \equiv \mathbb{E}\begin{bmatrix} 1 \\ \nabla_{\boldsymbol{\beta}}\ln h(\mathbf{X}_t;\boldsymbol{\beta}_*) \end{bmatrix} \quad \text{and}$$

$$\mathbf{C}_\infty \equiv \mathbb{E}\begin{bmatrix} 1 & \nabla_{\boldsymbol{\beta}}'\ln h(\mathbf{X}_t;\boldsymbol{\beta}_*) \\ \nabla_{\boldsymbol{\beta}}\ln h(\mathbf{X}_t;\boldsymbol{\beta}_*) & \nabla_{\boldsymbol{\beta}}\ln h(\mathbf{X}_t;\boldsymbol{\beta}_*)\nabla_{\boldsymbol{\beta}}'\ln h(\mathbf{X}_t;\boldsymbol{\beta}_*) \end{bmatrix}.$$

Note that this limit covariance structure leads to a different correlation structure that of Cho and White (2010, theorem $1(ii)$), who derive the same null limit distribution for the LR test statistic, irrespective of the

existence of the conditioning variable $\mathbf{X}_t$. This implies that their null limit distribution cannot be obtained by $\lim_\infty \mathcal{LR}_\omega(A)$.

3 Monte Carlo Experiments

3.1 Testing Using the Weighted Bootstrap Method

In this section, we conduct Monte Carlo experiments to examine the level and power properties of the LR test statistic.

For our experiments, we consider the following DGP and model conditions. First, for the level property, we specifically consider the following DGPs:

- DGP I: $Y_t \sim$ IID $\mathbb{G}[1 - \exp\{-\frac{1}{2}\}]$;
- DGP II: $Y_t|X_t \sim$ IID $\mathbb{G}[1 - \exp\{-1 + \exp[-\exp(X_t)]\}]$,

where $X_t \sim$ IID $N(0,1)$, and $\mathbb{G}(p_*)$ denotes the geometric distribution with parameter p_*, such that if $Y_t \sim \mathbb{G}(p_*)$ then $P(Y_t = 0) = p_*$.

Second, for these DGPs, we estimate the following models, respectively:

- Model I: $\pi\mathbb{G}[1 - \exp\{-\alpha_1\delta_*\}] + (1-\pi)\mathbb{G}[1 - \exp\{-\alpha_2\delta_*\}]$;
- Model II:

$$\pi\mathbb{G}[1 - \exp\{-\alpha_1\delta_*[1 - \exp(-\exp(\beta X_t))]\}] + (1-\pi)\mathbb{G}[1 - \exp\{-\alpha_2\delta_*[1 - \exp(-\exp(\beta X_t))]\}],$$

where we consider two different parameter spaces for $\alpha_1 \equiv \delta_1/\delta_*$ and $\alpha_2 \equiv \delta_2/\delta_*$, namely, $A_1 \equiv [3/4, 5/4]$ and $A_2 \equiv [3/4, 6/4]$. We conduct the experiments by estimating Models I and II using the observations generated by DGPs I and II, respectively. We implement the weighted bootstrap method described in Section 2.2.1 to examine the performance of the LR test statistic.

Remarks are warranted in implementing this experiment. First, the values of δ_* are 0.5 and 1.0 for Model I and II, respectively. However, they are unknown, α_1 and α_2 are not obtained directly. Instead, we estimate δ_* using the null model, denoting it as $\hat{\delta}_{on}$ and letting α_1 and α_2 be $\delta_1/\hat{\delta}_{on}$ and $\delta_2/\hat{\delta}_{on}$, respectively. For a finite sample size, A_1 and A_2 are not estimated precisely by this estimate, but the uncertainty conveyed by this estimate disappears as the sample size n tends to infinity.

Second, because DGP I and Model I do not contain conditioning variables, we can also test the mixture hypothesis using the methodology of Cho and Han (2009). We compare the performance of the LR

test statistic using the weighted bootstrap method with that of their methodology. Theorem 2 of Cho and Han (2009) shows that the asymptotic distribution of the LR test statistic can be obtained consistently by simulating

$$\widehat{\mathcal{LR}}_m(A) \equiv \sup_{\alpha \in A} \max[0, \widehat{\mathcal{G}}_m(\alpha)]^2$$

many times, where

$$\widehat{\mathcal{G}}_m(\alpha) \equiv \left\{ \frac{(1-\hat{p}_{on}) - [1-\hat{p}_{on}(\alpha)]^2}{(1-\hat{p}_{on})} \right\}^{1/2} \sum_{k=0}^{m} \left\{ \frac{[1-\hat{p}_{on}(\alpha)]}{\sqrt{1-\hat{p}_{on}}} \right\}^k Z_k,$$

$\hat{p}_{on} \equiv 1 - \exp\{-\hat{\delta}_{on}\}$, $\hat{p}_{on}(\alpha) \equiv 1 - \exp\{-\alpha\hat{\delta}_{on}\}$, and $Z_k \sim$ IID $N(0,1)$. For our comparison, we let m be 50 and denote the LR test statistic evaluated by $\widehat{\mathcal{LR}}_{50}(A)$ as $LR_n^\star(A)$.

Table 2 reports the finite sample level properties of the LR test statistic. We summarize these properties as follows.

(1) As the sample size n increases, the empirical levels approach the nominal levels (1%, 5%, and 10%). This feature is observed when applying both the weighted bootstrap method and the asymptotic critical values of Cho and Han (2009), implying that they are both asymptotically valid testing procedures under the null hypothesis.

(2) The empirical rejection rates obtained from the weighted bootstrap method are similar to those obtained from the asymptotic critical values of Cho and Han (2009). This feature implies that applying the weighted bootstrap method may be more appealing because it is applicable even when conditioning variables exist in the model.

(3) If the parameter space A is small, there is a tendency for the LR test statistic to yield more precise nominal levels. That is, the empirical rejection rate from A_1 is closer to the nominal level than that from A_2. Therefore, choosing a smaller parameter space can reduce finite sample level distortions.

Third, we consider the following DGPs for the power properties of the LR test statistic:

- DGP III: $Y_t \sim$ IID $\frac{1}{2}\mathbb{G}[1-\exp\{-0.3\}] + \frac{1}{2}\mathbb{G}[1-\exp\{-0.7\}]$;
- DGP IV: $Y_t|X_t \sim$ IID $\mathbb{G}[1-\exp\{1-\exp[-\exp(X_t)]\}]$;
- DGP V: $Y_t|(X_t, Z_t) \sim$ IID $\mathbb{G}[1-\exp\{-\exp(Z_t+X_t)\}]$; and
- DGP VI: $Y_t|X_t \sim$ IID $\frac{1}{2}\mathbb{G}\left[1-\exp\left\{-\exp\left(-\frac{2}{7}+X_t\right)\right\}\right] + \frac{1}{2}\mathbb{G}\left[1-\exp\left\{-\exp\left(\frac{5}{7}+X_t\right)\right\}\right]$,

where $(X_t, Z_t)' \sim$ IID $N(\mathbf{0}, \mathbf{I}_2)$. DGPs III and IV are estimated using Model I. We examine these DGPs to compare the performance of the LR test statistic evaluated by the weighted bootstrap method and the

asymptotic critical values of Cho and Han (2009). We estimate Model II using the observations generated from DGPs V and VI. As noted above, we cannot obtain the asymptotic critical values for this model because the marginal distribution of X_t is assumed to be unknown. Here, Model I and Model II are misspecified for DGP IV and DGPs V and VI, respectively.

We report the finite sample properties of the LR test statistic in Table 3. These properties are summarized as follows:

(1) The unobserved heterogeneity is consistently detected by the LR test statistic. As the sample size n increases, the empirical rejection rates approach unity for every case under consideration.

(2) For DGPs III and IV, the LR test statistics using the weighted bootstrap method have similar power patterns to the LR test statistic evaluated by the asymptotic critical values of Cho an Han (2009).

(3) Even when the distributional assumption of unobserved heterogeneity is incorrect, it is consistently detected by the LR test statistic. However, this does not necessarily imply that the LR test statistic is able to detect any distributional misspecification. As mentioned above, Cho and White (2008) consider a set of conditions under which the LR test statistic is able to detect unobserved heterogeneity consistently. Unless their conditions are met, the LR test may not be consistent for the unobserved heterogeneity.

3.2 Testing Using the Approximated Critical Values

In this section, we conduct a Monte Carlo simulation to examine how the empirical distribution of the LR test statistic is affected by letting the grouping bin size Δ converge to zero. According to Theorem 4, if $\omega \to \infty$, the empirical null limit distribution of the LR test statistic approaches that of the LR statistic used to test the exponential mixture hypothesis. We verify this property by means of a simulation.

We proceed with our experiments in the following order. First, we consider the following eight DGPs: for $\omega = 0.1$, 0.2, 5.0, and 10.0,

• DGP I*: $Y_t^\omega \sim$ IID $\mathbb{G}[1 - \exp\{-1/(2\omega)\}]$.

Second, for each ω, we compute the LR test statistic using the observations generated by DGP I* and Models 0* and I*, where

• Model 0*: $\mathbb{G}[1 - \exp(-\delta/\omega)]$;
• Model I*: $\pi\mathbb{G}[1 - \exp(-\delta_1/\omega)] + (1 - \pi)\mathbb{G}[1 - \exp(-\delta_2/\omega)]$.

For each ω, the total number of observations is 20,000. We also consider two parameter spaces for α_1 and α_2: A_1 and A_2. Therefore, we calculate eight (2 parameter spaces $\times$ 4 ω's) LR test statistics for each experiment.

Third, we obtain the null limit distribution of the LR test statistic that tests the exponential mixture hypothesis in the study of Cho and White (2010, theorem 2(*i*)). Following their methodology, we generate the null limit distribution.

Finally, we compare the empirical distributions of the our LR test statistics with the null limit distribution of Cho and White (2010, theorem 2(*i*)).

The results of the experiments are shown in Table 4 and Figure 1. Table 4 reports the empirical rejection rates evaluated using the null limit distribution of Cho and White (2010, theorem 2(*i*)). Figure 1 shows their null limit distribution, along with the empirical distributions of the LR test statistics. The results are summarized as follows:

(1) As ω increases, the empirical distributions of the LR test statistics are well approximated by the null limit distribution. For each $\omega = 0.5$, 2.0, and 10.0, the overall empirical distributional shapes are close to that of the null limit distribution. However, if $\omega = 0.1$, the empirical distribution of the LR test statistic is quite different from that of the null limit distribution. This implies that for a moderately large ω, the null limit distribution can be usefully exploited to test the geometric mixture hypothesis.

(2) If the associated parameter space A is small, the null limit distribution better approximates the empirical distributions of the LR test statistics. Note that the empirical distribution of the LR test statistic combined with A_1 is better approximated by the null limit distribution than that combined with A_2. This indicates that the intended approximation is more useful if the LR test statistic is combined with a smaller parameter space of A.

4 Conclusion

This study examines the mixture hypothesis of conditional geometric distributions using a likelihood ratio (LR) test statistic that extends that used for an unconditional geometric distribution by Cho and Han (2009). We derive the null limit distribution of the LR test statistic and examine its power performance. In addition, we examine the interrelationship between the LR test statistics used to test the geometric and exponential mixture hypotheses. We also examine the performance of the LR test statistics under various circumstances and confirm the main claims of the study using Monte Carlo simulations.

Appendix 1: Assumptions

Here, we apply the regularity conditions in Cho and White (2010) to our data context and provide regularities for the claims in the text.

Assumption 1. *(i)* $\{(Y_t^\omega, \boldsymbol{X}_t')'\}$ *is a strictly stationary geometric* β*-mixing process defined on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* β*-mixing coefficients* $\beta_\tau \leq c\rho^\tau$*, for some* $c > 0$ *and* $\rho \in [0, 1)$*, where* Y_t^ω *is* $\mathbb{N}$*-valued,* $\boldsymbol{X}_t$ *is* $\mathbb{R}^k-$*valued,* $k \in \mathbb{N}$*, and* $\boldsymbol{X}_t$ *does not contain a constant term;*

(ii) For $t = 1, 2, \ldots$*, conditional on* $\boldsymbol{X}_t$*,* Y_t^ω *has the following mass: for some* $(\pi_*, \boldsymbol{\beta}_*, \delta_{1*}^\omega, \delta_{2*}^\omega) \in \mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}^+$ $(d \in \mathbb{N})$*,*

$$m^\omega(y|\boldsymbol{X}_t; \pi_*, \boldsymbol{\beta}_*, \delta_{1*}^\omega, \delta_{2*}^\omega) \equiv \pi_* g_o^\omega(y|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_{1*}^\omega) + (1 - \pi_*) g_o^\omega(y|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_{2*}^\omega),$$

where $g_o^\omega(y|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega) \equiv [1 - \exp\{-\delta_*^\omega h(\boldsymbol{X}_t; \boldsymbol{\beta}_*)\}] \times [\exp\{-\delta_*^\omega h(\boldsymbol{X}_t; \boldsymbol{\beta}_*)\}]^{y-1}$*; and for each* $\boldsymbol{\beta} \in \boldsymbol{B} \subset \mathbb{R}^d$*,* $h(\,\cdot\,; \boldsymbol{\beta}) : \mathbb{R}^k \to \mathbb{R}^+$ *is a Borel measurable function;*

(iii) $m^\omega(\,\cdot\,|\boldsymbol{X}_t; \pi_*, \boldsymbol{\beta}_*, \delta_{1*}^\omega, \delta_{2*}^\omega) = p(\,\cdot\,|\boldsymbol{X}_t, Y_{t-1}^\omega, \boldsymbol{X}_{t-2}, Y_{t-2}^\omega, \cdots)$ *almost surely, where* $p(\,\cdot\,|\boldsymbol{X}_t, Y_{t-1}^\omega, \boldsymbol{X}_{t-2}, Y_{t-2}^\omega, \cdots)$ *is the conditional probability mass function of* Y_t^ω *given* $\boldsymbol{X}_t$*,* Y_{t-1}^ω*,* $\boldsymbol{X}_{t-2}$*,* Y_{t-2}^ω*, ...* □

Assumption 2. *(i)* $h(\boldsymbol{X}_t; \,\cdot\,)$ *is four-times continuously differentiable almost surely;*

(ii) $(\pi_*, \boldsymbol{\beta}_*, \delta_{1*}^\omega, \delta_{2*}^\omega) \in [0, 1] \times \boldsymbol{B} \times D \times D$*, and* $\boldsymbol{B} \times D \times D$ *is a convex compact subset of* $\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}^+$*.*

□

For each α and α' in A, we let

$$\mathbf{A}^\omega(\alpha, \alpha') \equiv \begin{bmatrix} \mathbb{E}[M_t^\omega(\alpha) M_t^\omega(\alpha')] - 1 & \mathbb{E}[M_t^\omega(\alpha) W_t^\omega] & \mathbb{E}[M_t^\omega(\alpha) \mathbf{S}_t'] \\ \mathbb{E}[W_t^\omega M_t^\omega(\alpha')] & \mathbb{E}[W_t^{\omega 2}] & \mathbb{E}[W_t^\omega \mathbf{S}_t^{\omega\prime}] \\ \mathbb{E}[\mathbf{S}_t^\omega M_t^\omega(\alpha')] & \mathbb{E}[\mathbf{S}_t^\omega W_t^\omega] & \mathbb{E}[\mathbf{S}_t^\omega \mathbf{S}_t^{\omega\prime}] \end{bmatrix},$$

where $W_t^\omega \equiv \nabla_\delta^2 g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)/g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)$, for each $\alpha \in A$, $M_t^\omega(\alpha) \equiv g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \alpha\delta_*^\omega)$ $/g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)$, and $\mathbf{S}_t^\omega \equiv \nabla_{(\boldsymbol{\beta}, \delta^\omega)} g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)$ $/g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)$. Furthermore, we also let

$$\mathbf{B}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*}) \equiv \mathbb{E}[\nabla_{(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)} \ell_t^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*}) \nabla'_{(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)} \ell_t^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})],$$

where we let $\ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2) \equiv \ln[\pi g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}, \alpha_1\zeta_o^\omega) + (1 - \pi) g_o^\omega(Y_t^\omega|\mathbf{X}_t; \boldsymbol{\beta}, \alpha_2\zeta_o^\omega)]$, $A \equiv \{\alpha : \alpha\zeta_o^\omega \in D\}$, and ζ_o^ω is defined as in Assumption 3.

Assumption 3. *(i)* $(\boldsymbol{\beta}_o, \zeta_o^\omega) \equiv \arg\max_{(\boldsymbol{\beta}, \delta^\omega) \in \boldsymbol{B} \times D} \mathbb{E}[\ln g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}, \delta^\omega)]$ *exists and is unique, and for each* $(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2) \in [0, 1] \times \boldsymbol{B} \times A \times A$*,* $\mathbb{E}[\ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)]$ *exists and is finite;*

(ii) For every $(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})$, $\lambda_{\min}\{\mathbf{B}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})\} \geq 0$ *such that*

(a) if $\lambda_{\min}\{\mathbf{B}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})\} > 0$, $\lambda_{\max}\{\mathbf{B}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})\} < \infty$*;*

(b) if $\lambda_{\min}\{\mathbf{B}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})\} = 0$*, for any* $\epsilon > 0$, $\lambda_{\min}\{\mathbf{A}^\omega(\pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*}\,)\} > 0$ *and* $\lambda_{\max}\{\mathbf{A}^\omega(\alpha, \alpha)\} < \infty$ *uniformly in* $\alpha \in A(\epsilon) \equiv \{\alpha \in A : |\alpha - 1| \geq \epsilon\}$*, where* $\lambda_{\min}(\cdot)$ *and* $\lambda_{\max}(\cdot)$ *are the minimum and maximum eigenvalues of the given matrix, respectively.* □

Assumption 4. *There exists a sequence of strictly stationary and ergodic random variables* $\{M_t\}$ *such that for some* $\epsilon > 0$*,*

(a) $\mathbb{E}[M_t^{1+\epsilon}] < \Delta < \infty$*;*

(b) $\sup_{(\pi,\boldsymbol{\beta},\alpha_1,\alpha_2)} |\nabla_j \ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2) \nabla_k \ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)| \leq M_t$*;*

(c) $\sup_{(\pi,\boldsymbol{\beta},\alpha_1,\alpha_2)} |\nabla_{j,k} \ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)| \leq M_t$*;*

(d) $|\nabla_{i_1} g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)/g_o^\omega(Y_t^\omega|\, \boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)|^4 \leq M_t$*;*

(e) $|\nabla_{i_1}\nabla_{i_2} g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)/g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)|^2 \leq M_t$*;*

(f) $|\nabla_{i_1}\nabla_{i_2}\nabla_{i_3} g_o^\omega\ (Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)/g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}_*, \delta_*^\omega)|^2 \leq M_t$*; and*

(g) $\sup_{(\boldsymbol{\beta},\delta)} |\nabla_{i_1}\nabla_{i_2}\nabla_{i_3}\nabla_{i_4} g_o^\omega(Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}, \delta^\omega)/g_o^\omega\ (Y_t^\omega|\boldsymbol{X}_t; \boldsymbol{\beta}, \delta^\omega)| \leq M_t$*, where* $j, k \in \{\pi, \alpha_1, \alpha_2, \beta_1, \cdots, \beta_d\}$ *and* $i_1, \cdots, i_4 \in \{\delta, \beta_1, \cdots, \beta_d\}$. □

Appendix 2: Proofs

Proof of Theorem 1: The proof of Theorem 1 is completed by letting $\omega = 1.0$ in the proof of Theorem 3. ■

We provide the following supplementary lemma to prove the consistency of the LR test.

Lemma 1. *If Assumptions 1 to 4 are satisfied,* $\sup_{(\pi,\boldsymbol{\beta},\alpha_1,\alpha_2)} |n^{-1}\sum \ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1,\ \alpha_2) - \mathbb{E}[\ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)]| \overset{a.s.}{\to} 0$. □

Proof of Lemma 1: First, note that $\ell_t^\omega(\cdot)$ is differentiable by Assumption 2(i). Thus, it is continuous on $[0,1] \times B \times A \times A$, which is a compact and convex subset of $\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}^+$. Second, for some positive, stationary, and ergodic random variable, M_t, it follows that $\|\nabla_{(\pi,\boldsymbol{\beta},\alpha_1,\alpha_2)} \ell_t^\omega(\pi, \boldsymbol{\beta}, \alpha_1, \alpha_2)\|_\infty < M_t$ by Assumption 4. Finally, we can, therefore, apply Ranga Rao's (1962) uniform law of large numbers to $\{n^{-1}\sum \ell_t^\omega(\,\cdot\,)\}$. This completes the proof. ■

Proof of Theorem 2: We apply the proof of theorem 3 in Cho and White (2010). The Kullback–Leibler information criterion implies that $\mathbb{E}[\ln g_a(Y_t|\mathbf{X}_t;\ \pi_*, \boldsymbol{\beta}_*, \alpha_{1*},\ \alpha_{2*})] > \mathbb{E}[\ln g_o(Y_t|\mathbf{X}_t; \boldsymbol{\beta}_o, \zeta_o)]$ under $\mathcal{H}_a$,

where $\zeta_o = \zeta_o^\omega$ such that $\omega = 1$. Therefore, Lemma A implies that there exists $n^*(\varepsilon)$ with probability one such that if $n \geq n^*(\varepsilon)$, then $|G_{1n}| < \varepsilon$, $|G_{2n}| < \varepsilon$, $|H_{1n}| < \varepsilon$, and $|H_{2n}| < \varepsilon$, where

$$G_{1n} \equiv n^{-1} \sum \left\{ \ln[g_a(Y_t|\mathbf{X}_t; \hat{\pi}_n, \hat{\boldsymbol{\beta}}_n, \hat{\alpha}_{1n}, \hat{\alpha}_{2n})] - \ln[g_a(Y_t|\mathbf{X}_t; \pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})] \right\};$$

$$G_{2n} \equiv n^{-1} \sum \ln[g_a(Y_t|\mathbf{X}_t; \pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})] - \mathbb{E}\{\ln[g_a(Y_t|\mathbf{X}_t; \pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})]\};$$

$$H_{1n} \equiv n^{-1} \sum \left\{ \ln[g_o(Y_t|\mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on})] - \ln[g_o(Y_t|\mathbf{X}_t; \boldsymbol{\beta}_o, \zeta_o)] \right\}; \quad \text{and}$$

$$H_{2n} \equiv n^{-1} \sum \ln[g_o(Y_t|\mathbf{X}_t; \boldsymbol{\beta}_o, \zeta_o)] - \mathbb{E}\{\ln[g_o(Y_t|\mathbf{X}_t; \boldsymbol{\beta}_o, \zeta_o)]\}.$$

From this, it follows that $|(G_{1n} + G_{2n}) - (H_{1n} + H_{2n})| \leq \eta$, where $\eta \equiv 4\varepsilon$. If we let

$$\hat{\Lambda}_n \equiv n^{-1} \sum \left\{ \ln[g_a(Y_t|\mathbf{X}_t; \hat{\pi}_n, \hat{\boldsymbol{\beta}}_n, \hat{\alpha}_{1n}, \hat{\alpha}_{2n})] - \ln[g_o(Y_t|\mathbf{X}_t; \hat{\boldsymbol{\beta}}_{on}, \hat{\delta}_{on})] \right\} \quad \text{and}$$

$$\Lambda_* \equiv \mathbb{E}\{\ln[g_a(Y_t|\mathbf{X}_t; \pi_*, \boldsymbol{\beta}_*, \alpha_{1*}, \alpha_{2*})]\} - \mathbb{E}\{\ln[g_o(Y_t|\mathbf{X}_t; \boldsymbol{\beta}_o, \zeta_o)]\},$$

then we obtain $\Lambda_* - \eta \leq \hat{\Lambda}_n \leq \Lambda_* + \eta$. Thus, for some $\delta_1 \in (0, \Lambda_* - \eta)$ and $\delta_2 \in (\Lambda_* + \eta, \infty)$, if $n > n^*(\varepsilon)$, then $\delta_1 < \hat{\Lambda}_n < \delta_2$. From $LR_n(A) = 2n\hat{\Lambda}_n$, we find that $LR_n(A) = O_p(n)$, but not $o_p(n)$. This completes the proof. ■

Proof of Theorem 3: We derive the desired weak convergence of the LR statistic by verifying the conditions of theorem 6(a) of Cho and White (2007). First, our Assumption 1 is sufficient for their assumption A1; second, our Assumptions 1(ii) and 2 satisfy their assumption A2; third, Assumption 4(i) is sufficient for their assumptions A3 and A4; fourth, their assumptions A5(ii and iii) are satisfied by our Assumption 4; and finally, our Assumption 4 relaxes their A6(iv) because it is not necessary to impose the positive definite matrix assumption on $\mathbf{A}(\alpha, \alpha')$ for every $(\alpha, \alpha')(\neq (1,1))$ to obtain the desired result. Therefore, the desired weak convergence follows from their theorem 6(a).

Next, we derive the covariance structure (5) using the formula in lemma 1(b) of Cho and White (2007). First, note that

$$M_t^\omega(\alpha) = \left[\frac{1 - Q_t(\alpha/\omega)}{1 - Q_t(1/\omega)}\right] \left[\frac{Q_t(1/\omega)}{Q_t(\alpha/\omega)}\right]^{Y_t^\omega},$$

so that

$$\begin{aligned}\mathbb{E}[M_t^\omega(\alpha) M_t^\omega(\alpha')|\mathbf{X}_t] - 1 &= \left[\frac{[Q_t(\alpha/\omega) - 1][Q_t(\alpha'/\omega) - 1]}{Q_t(1/\omega) - 1}\right] \sum_{y=1}^{\infty} \left[\frac{Q_t(1/\omega)}{Q_t(\alpha/\omega) Q_t(\alpha'/\omega)}\right]^y - 1 \\ &= \frac{[Q_t(1/\omega) - Q_t(\alpha/\omega)][Q_t(1/\omega) - Q_t(\alpha'/\omega)]}{[Q_t(1/\omega) - 1][Q_t(\alpha/\omega) Q_t(\alpha'/\omega) - Q_t(1/\omega)]},\end{aligned}$$

and

$$\mathbb{E}[M_t^\omega(\alpha) M_t^\omega(\alpha')] - 1 = A_\omega(\alpha, \alpha') \tag{9}$$

using the law of iterated expectation: $\mathbb{E}[M_t^\omega(\alpha)M_t^\omega(\alpha')] = \mathbb{E}[\mathbb{E}[M_t^\omega(\alpha)M_t^\omega(\alpha')|\ \mathbf{X}_t]]$. Here, we can apply the infinite geometric sum formula because $\inf A > 1/2$. Note that for any α and $\alpha' \in A$, $Q_t(1/\omega)/\{Q_t(\alpha/\omega)Q_t(\alpha'/\omega)\} = \exp[(1-\alpha-\alpha')\delta_*^\omega h(\mathbf{X}_t;\boldsymbol{\beta}_*)] \in (0,1)$ because $\delta_*^\omega > 0$ and $h(\,\cdot\,;\boldsymbol{\beta}_*) > 0$ by Assumption $1(ii)$.

Second, we consider $\mathbb{E}[M_t^\omega(\alpha)\mathbf{S}_t^\omega]$. For each α,

$$\mathbf{S}_t^\omega = \delta_*^\omega \left[\frac{Q_t(1/\omega)}{Q_t(1/\omega)-1} - Y_t^\omega\right]\mathbf{D}_t^\omega,$$

where

$$\mathbf{D}_t^\omega \equiv \begin{bmatrix} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \\ \nabla_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \end{bmatrix},$$

so that

$$\begin{aligned}\mathbb{E}[M_t^\omega(\alpha)\mathbf{S}_t^\omega|\mathbf{X}_t] &= \delta_*^\omega\,[Q_t(\alpha/\omega)-1]\sum_{y=1}^{\infty}\left[\frac{1}{Q_t(\alpha/\omega)}\right]^y\left[\frac{Q_t(1/\omega)}{Q_t(1/\omega)-1} - y\right]\mathbf{D}_t^\omega \\ &= \frac{\delta_*^\omega[Q_t(\alpha/\omega)-Q_t(1/\omega)]}{[Q_t(1/\omega)-1][Q_t(\alpha/\omega)-1]}\mathbf{D}_t^\omega.\end{aligned}$$

Here, we can apply the infinite geometric sum formula because for each $\alpha \in A$, $1/Q_t(\alpha/\omega) \in (0,1)$. This implies that

$$\mathbb{E}[M_t^\omega(\alpha)\mathbf{S}_t^\omega] = \mathbb{E}\left\{\frac{\delta_*^\omega[Q_t(\alpha/\omega)-Q_t(1/\omega)]}{[Q_t(1/\omega)-1][Q_t(\alpha/\omega)-1]}\begin{bmatrix} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \\ \nabla_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \end{bmatrix}\right\}$$

by the law of iterated expectation, and that

$$\mathbb{E}[M_t^\omega(\alpha)\mathbf{S}_t^\omega] = \mathbf{B}_\omega(\alpha). \tag{10}$$

Third, we consider $\mathbb{E}[\mathbf{S}_t^\omega\mathbf{S}_t^{\omega\prime}]$. Note that

$$\begin{aligned}\mathbb{E}[\mathbf{S}_t^\omega\mathbf{S}_t^{\omega\prime}|\mathbf{X}_t] &= (\delta_*^\omega)^2[Q_t(1/\omega)-1]\sum_{y=1}^{\infty}\left[\frac{Q_t(1/\omega)}{Q_t(1/\omega)-1} - y\right]^2\left[\frac{1}{Q_t(1/\omega)}\right]^y\mathbf{D}_t^\omega\mathbf{D}_t^{\omega\prime} \\ &= \frac{(\delta_*^\omega)^2 Q_t(1/\omega)}{[Q_t(1/\omega)-1]^2}\mathbf{D}_t^\omega\mathbf{D}_t^{\omega\prime}.\end{aligned}$$

This implies that

$$\mathbb{E}[\mathbf{S}_t^\omega\mathbf{S}_t^{\omega\prime}] = \mathbb{E}\left\{\frac{(\delta_*^\omega)^2 Q_t(1/\omega)}{[Q_t(1/\omega)-1]^2}\begin{bmatrix} h(\mathbf{X}_t;\boldsymbol{\beta}_*)^2 & h(\mathbf{X}_t;\boldsymbol{\beta}_*)\nabla'_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \\ h(\mathbf{X}_t;\boldsymbol{\beta}_*)\nabla_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*) & \nabla_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*)\nabla'_{\boldsymbol{\beta}} h(\mathbf{X}_t;\boldsymbol{\beta}_*) \end{bmatrix}\right\}$$

by the law of iterated expectation, so that

$$\mathbb{E}[\mathbf{S}_t^\omega\mathbf{S}_t^{\omega\prime}] = \mathbf{C}_\omega. \tag{11}$$

By substituting (9), (10), and (11) into

$$\rho_\omega(\alpha,\alpha') \equiv \mathbb{E}[M_t^\omega(\alpha)M_t^\omega(\alpha')] - 1 - \mathbb{E}[M_t^\omega(\alpha)\mathbf{S}_t]'\{\mathbb{E}[\mathbf{S}_t\mathbf{S}_t']\}^{-1}\mathbb{E}[M_t^\omega(\alpha')\mathbf{S}_t]$$

in lemma 1(b) of Cho and White (2007), we obtain that $\rho_\omega(\alpha,\alpha') \equiv A_\omega(\alpha,\alpha') - \mathbf{B}_\omega'\mathbf{C}_\omega^{-1}\mathbf{B}_\omega(\alpha')$. This is the desired covariance structure (5) and completes the proof. ■

Proof of Theorem 4: From the given assumption that $\mathbf{X}_t$ is absent, $Q_t(1/\omega) = \exp(\delta_*/\omega)$ and $Q_t(\alpha/\omega) = \exp(\alpha\delta_*/\omega)$. Thus,

$$\begin{aligned} A_\omega(\alpha,\alpha') &= \frac{[\exp(\alpha\delta_*/\omega) - \exp(\delta_*/\omega)][\exp(\alpha'\delta_*/\omega) - \exp(\delta_*/\omega)]}{[\exp(\delta_*/\omega) - 1][\exp((\alpha+\alpha')\delta_*/\omega) - \exp(\delta_*/\omega)]} \\ &\xrightarrow{\omega\to\infty} \frac{(\alpha-1)(\alpha'-1)}{\alpha+\alpha'-1}; \end{aligned}$$

$$\mathbf{B}_\omega(\alpha) = \frac{\delta_*[\exp(\alpha\delta_*/\omega) - \exp(\delta_*/\omega)]}{\omega[\exp(\delta_*/\omega) - 1][\exp(\alpha\delta_*/\omega) - 1]} \xrightarrow{\omega\to\infty} \left(\frac{\alpha-1}{\alpha}\right)$$

$$\mathbf{C}_\omega \equiv \frac{\delta_*^2 Q_t(1/\omega)}{\omega^2[Q_t(1/\omega) - 1]^2} \xrightarrow{\omega\to\infty} 1.$$

Thus, it follows that

$$\lim_{\omega\to\infty} \rho_\omega(\alpha,\alpha') = \frac{(\alpha-1)(\alpha'-1)}{\alpha+\alpha'-1} - \left(\frac{\alpha-1}{\alpha}\right)\left(\frac{\alpha'-1}{\alpha'}\right) = \frac{(\alpha-1)^2(\alpha'-1)^2}{\alpha\alpha'(\alpha+\alpha'-1)}$$

and

$$\lim_{\omega\to\infty} \frac{\rho_\omega(\alpha,\alpha')}{\sqrt{\rho_\omega(\alpha,\alpha)}\sqrt{\rho_\omega(\alpha',\alpha')}} = \frac{(2\alpha-1)^{1/2}(2\alpha'-1)^{1/2}}{\alpha+\alpha'-1}.$$

This completes the proof. ■

References

Andrews, D.W.K. (1999). Estimation when a parameter is on a boundary, Econometrica 67, 543–563.

Andrews, D.W.K. (2001). Testing when a parameter is on the boundary of the maintained hypothesis, Econometrica 69, 683–734.

Baek, Y., Cho, J., and Phillips, P.C.B. (2015). Testing linearity using power transforms of regressors, Journal of Econometrics 187, 376–384.

Chen, H. and Chen, J. (2001). The likelihood ratio test for homogeneity in the finite mixture models, Canadian Journal of Statistics 29, 201–216.

Chernoff, H. and Lander, E. (1995). Asymptotic distribution of the likelihood ratio test that a mixture of two binomials is a single binomial, Journal of Statistical Planning and Inference 43, 19–40.

Cho, J., Cheong, T., and White, H, (2011). Experience with the weighted bootstrap in testing for unobserved heterogeneity in exponential and weibull duration models, Journal of Economic Theory and Econometrics 22, 60–91.

Cho, J. and Han, C. (2009). Testing for the mixture hypothesis of geometric distributions, Journal of Economic Theory and Econometrics 20, 31–55.

Cho, J. and Ishida, I. (2012). Testing for the effects of omitted power transformations, Economics Letters 117, 287–290.

Cho, J., Ishida, I., and White, H. (2011). Revisiting tests for neglected nonlinearity using artificial neural networks, Neural Computation 23, 1133–1186.

Cho, J., Ishida, I., and White, H. (2014). Testing for neglected nonlinearity using twofold unidentified models under the null and hexic expansions, in Essays in Nonlinear Time Series Econometrics, Festschrift in Honor of Timo Teräsvirta, 3–27, eds., N. Haldrup, M. Meitz, and P. Saikkonen, Oxford University Press.

Cho, J. and Phillips, P.C.B. (2018). Sequentially testing polynomial model hypotheses using power transforms of regressors, Journal of Applied Econometrics 33, 141–159.

Cho, J. and White, H, (2007). Testing for regime switching, Econometrica 75, 1671–1720.

Cho, J. and White, H, (2008). Testing the mixture hypothesis with misspecified exponential family distributions, Discussion Paper, Yonsei University.

Cho, J. and White, H. (2010). Testing for unobserved heterogeneity in exponential and Weibull duration models, Journal of Econometrics 157, 458–480.

Cho, J. and White, H. (2011a). Generalized runs tests for the IID hypothesis, Journal of Econometrics 162, 326–344.

Cho, J. and White, H. (2011b). Testing correct model specification using extreme learning machines, Neurocomputing 74, 2552–2565.

Davies, R. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247–254.

Davies, R. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 74, 33–43.

Hansen, B. (1996). Inference when a nuisance parameter is not identified under the null hypothesis, Econometrica 64, 413–430.

Kennan, J. (1985). The duration of contract strikes in U.S. manufacturing, Journal of Econometrics 28, 5–28.

Lancaster, T. (1979). Econometric methods for the duration of unemployment, Econometrica 47, 939–956.

Lancaster, T. (1992). The Econometric Analysis of Transition Data, Cambridge University Press.

Liu, X., Pasarica, C., and Shao, Y. (2003). Testing homogeneity in gamma mixture models, Scandinavian Journal of Statistics 30, 227–239.

Nickell, S. (1979). Estimating the probability of leaving unemployment, Econometrica 47, 1249–1266.

Piterbarg, V. (1996). Asymptotic Methods in the Theory of Gaussian Processes and Fields, Translations of Mathematical Monographs, 148, American Mathematical Society.

Ranga Rao, R. (1962). Relations between weak and uniform convergence of measures with applications, Annals of Mathematical Statistics 33, 659–680.

Ryu, K. (1995). Analysis of a continuous proportional hazard model using discrete duration data, Econometric Review 14, 299–313.

Van den Berg, G. and Ridder, G. (1998). An empirical equilibrium search model of the labor market, Econometrica, 66, 1183–1221.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large, Transactions of the American Mathematical Society 54, 426–482.

White, H. and Cho, J. (2011). Higher-order approximations for testing neglected nonlinearity, Neural Computation 24, 273–287.

Authors	Mixtures
Chernoff and Lander (1995)	Unconditional binomial distributions
Chen and Chen (2001)	Unconditional normal and Poisson distributions
Liu, Pasarica and Shao (2003)	Unconditional gamma distributions
Cho and White (2007)	Conditional Normal distributions
Cho and Han (2009)	Unconditional geometric distributions
Cho and White (2010)	Conditional Exponential and Weibull distreibutions

Table 1: LITERATURE FOR TESTING THE MIXTURE HYPOTHESIS. This table provides the literature providing the null limit distribution of statistics testing the mixture hypothesis.

DGP: $Y_t \sim$ IID $\mathbb{G}[1-\exp\{-\frac{1}{2}\}]$							
Model 0 versus Model I							
Statistics	Levels $\backslash$ n	500	1,000	2,000	5,000	10,000	20,000
$LR_n(A_1)$	1%	0.30	0.50	0.65	1.00	0.90	0.95
	5%	2.05	3.20	3.45	3.80	3.95	4.05
	10%	6.55	7.00	7.75	8.25	8.50	9.15
$LR_n(A_2)$	1%	0.80	0.90	0.80	0.90	1.10	0.75
	5%	3.05	3.85	4.60	4.30	4.80	4.40
	10%	7.35	7.65	8.65	8.45	8.95	8.50
$LR^\star_n(A_1)$	1%	0.25	0.65	0.60	1.10	0.90	1.05
	5%	2.15	3.10	4.05	4.35	4.40	4.65
	10%	6.30	7.20	8.30	8.55	9.25	9.95
$LR^\star_n(A_2)$	1%	0.70	0.75	0.80	0.60	1.25	0.90
	5%	2.75	3.25	4.25	4.05	4.65	4.10
	10%	5.90	6.55	8.10	8.05	8.55	8.40
DGP: $Y_t \vert X_t \sim$ IID $\mathbb{G}[1-\exp\{-1+\exp[-\exp(X_t)]\}]$							
Model $0'$ versus Model II							
Statistics	Levels $\backslash$ n	500	1,000	2,000	5,000	10,000	20,000
$LR_n(A_1)$	1%	0.20	0.30	1.10	1.05	1.00	1.10
	5%	2.80	4.45	5.50	5.55	5.10	5.45
	10%	8.20	9.35	11.95	11.40	9.85	11.45
$LR_n(A_2)$	1%	0.70	0.85	1.15	0.85	1.30	1.70
	5%	4.30	4.15	5.10	6.25	6.70	6.70
	10%	8.75	8.95	10.40	12.40	12.75	12.95

Table 2: LEVELS OF THE LR TEST STATISTICS (NUMBER OF EXPERIMENT REPETITIONS: 2,000). This table shows the finite sample properties of the LR test statistics under the DGP/model assumptions and in the absence of unobserved heterogeneity. Note that the empirical rejection rates are more or less similar to the nominal significance levels. Model 0 indicates $\mathbb{G}[1-\exp\{-\delta_1\}]$, Model I indicates $\pi\mathbb{G}[1-\exp\{-\delta_1\}]+(1-\pi)\mathbb{G}[1-\exp\{-\delta_2\}]$, Model $0'$ indicates $\mathbb{G}[1-\exp\{1-\exp[-\delta_1\exp(\beta X_t)]\}]$, and Model II indicates $\pi\mathbb{G}[1-\exp\{-\delta_1(1-\exp[-\exp(\beta X_t)])\}]+(1-\pi)\mathbb{G}[1-\exp\{-\delta_2(1-\exp[-\exp(\beta X_t)])\}]$. Furthermore, we let $X_t \sim$ IID $N(0,1)$, $A_1 \equiv [3/4, 5/4]$ and $A_2 \equiv [3/4, 6/4]$. $LR_n(A)$ and $LR^\star_n(A)$ denotes the LR test statistics evaluated by the weighted bootstrap and the methodologyly in Cho and Han (2009), respectively.

DGP: $Y_t \sim$ IID $\frac{1}{2}\mathbb{G}[1-\exp\{-0.3\}] + \frac{1}{2}\mathbb{G}[1-\exp\{-0.7\}]$						
Model 0 versus Model I						
Statistics $\backslash\, n$	50	100	200	500	1,000	2,000
$LR_n(A_1)$	0.55	3.75	18.40	73.20	98.80	100.0
$LR_n(A_2)$	3.25	10.35	29.30	79.80	98.75	100.0
$LR^\star_n(A_1)$	0.45	4.00	19.55	77.45	98.90	100.0
$LR^\star_n(A_2)$	2.65	9.25	27.50	80.25	98.75	100.0
DGP: $Y_t\|X_t \sim$ IID $\mathbb{G}[1-\exp\{1-\exp[-\exp(X_t)]\}]$						
Model 0 versus Model I						
Statistics $\backslash\, n$	50	100	200	500	1,000	2,000
$LR_n(A_1)$	12.10	48.15	90.55	100.0	100.0	100.0
$LR_n(A_2)$	21.40	63.40	93.40	100.0	100.0	100.0
$LR^\star_n(A_1)$	12.00	48.50	91.70	100.0	100.0	100.0
$LR^\star_n(A_2)$	19.05	63.10	93.80	100.0	100.0	100.0
DGP: $Y_t\|(X_t, Z_t) \sim$ IID $\mathbb{G}[1-\exp\{-\exp(Z_t+X_t)\}]$						
Model 0′ versus Model II						
Statistics $\backslash\, n$	50	100	200	500	1,000	2,000
$LR_n(A_1)$	6.00	22.05	56.40	92.65	99.90	100.0
$LR_n(A_2)$	11.90	35.85	69.05	97.40	99.95	100.0
DGP: $Y_t\|X_t \sim$ IID $\frac{1}{2}\mathbb{G}[1-\exp\{-\exp(-\frac{2}{7}+X_t)\}] + \frac{1}{2}\mathbb{G}[1-\exp\{-\exp(\frac{5}{7}+X_t)\}]$						
Model 0′ versus Model II						
Statistics $\backslash\, n$	50	100	200	500	1,000	2,000
$LR_n(A_1)$	0.15	1.55	8.30	32.70	58.75	84.90
$LR_n(A_2)$	1.95	6.10	16.70	36.90	63.50	87.40

Table 3: LEVELS OF THE LR TEST STATISTICS (NUMBER OF EXPERIMENT REPETITIONS: 2,000; LEVEL OF SIGNIFICANCE: 5%). This table shows the finite sample properties of the LR test statistics under the DGP/model assumptions and in the presence of unobserved heterogeneity. Note that the empirical rejection rates converge to 100% as the sample size increases. Model 0 indicates $\mathbb{G}[1-\exp\{-\delta_1\}]$, Model I indicates $\pi\mathbb{G}[1-\exp\{-\delta_1\}]+(1-\pi)\mathbb{G}[1-\exp\{-\delta_2\}]$, Model 0′ indicates $\mathbb{G}[1-\exp\{1-\exp[-\delta_1\exp(\beta X_t)]\}]$, and Model II indicates $\pi\mathbb{G}[1-\exp\{-\delta_1(1-\exp[-\exp(\beta X_t)])\}]+(1-\pi)\mathbb{G}[1-\exp\{-\delta_2(1-\exp[-\exp(\beta X_t)])\}]$. Furthermore, we let $(X_t, Z_t)' \sim$ IID $N(\mathbf{0}, \mathbf{I}_2)$, $A_1 \equiv [3/4, 5/4]$ and $A_2 \equiv [3/4, 6/4]$. $LR_n(A)$ and $LR^\star_n(A)$ denotes the LR test statistics evaluated by the weighted bootstrap and the methodology in Cho and Han (2009), respectively.

Model 0^* versus Model I^*					
Parameter Space	Levels $\setminus$ ω	0.10	0.50	2.00	10.00
$A_1 \equiv [3/4, 5/4]$	1%	0.35	0.40	0.85	0.55
	5%	1.85	3.15	3.65	3.50
	10%	5.00	7.60	7.15	8.05
$A_2 \equiv [3/4, 6/4]$	1%	0.35	0.80	0.90	0.55
	5%	2.70	3.80	4.30	3.80
	10%	5.90	6.90	7.55	8.05

Table 4: LEVELS OF THE LR TEST STATISTICS (Number of Experiment Repetitions: 2,000; Number of Observations: 20,000). This tables shows the empirical rejection rates of the LR test statistic when it is evaluated by the null limit distribution in Cho and White (2010). Note that the nominal levels of significance (1%, 5%, and 10%) get close to the empirical rejection rates, as the grouping bin size ($\Delta \equiv 1/\omega$) reduces to zero. Model 0^* indicates $\mathbb{G}[1 - \exp\{-\delta/\omega\}]$, and Model I^* indicates $\pi\mathbb{G}[1 - \exp\{-\delta_1/\omega\}] + (1 - \pi)\mathbb{G}[1 - \exp\{-\delta_2/\omega\}]$.

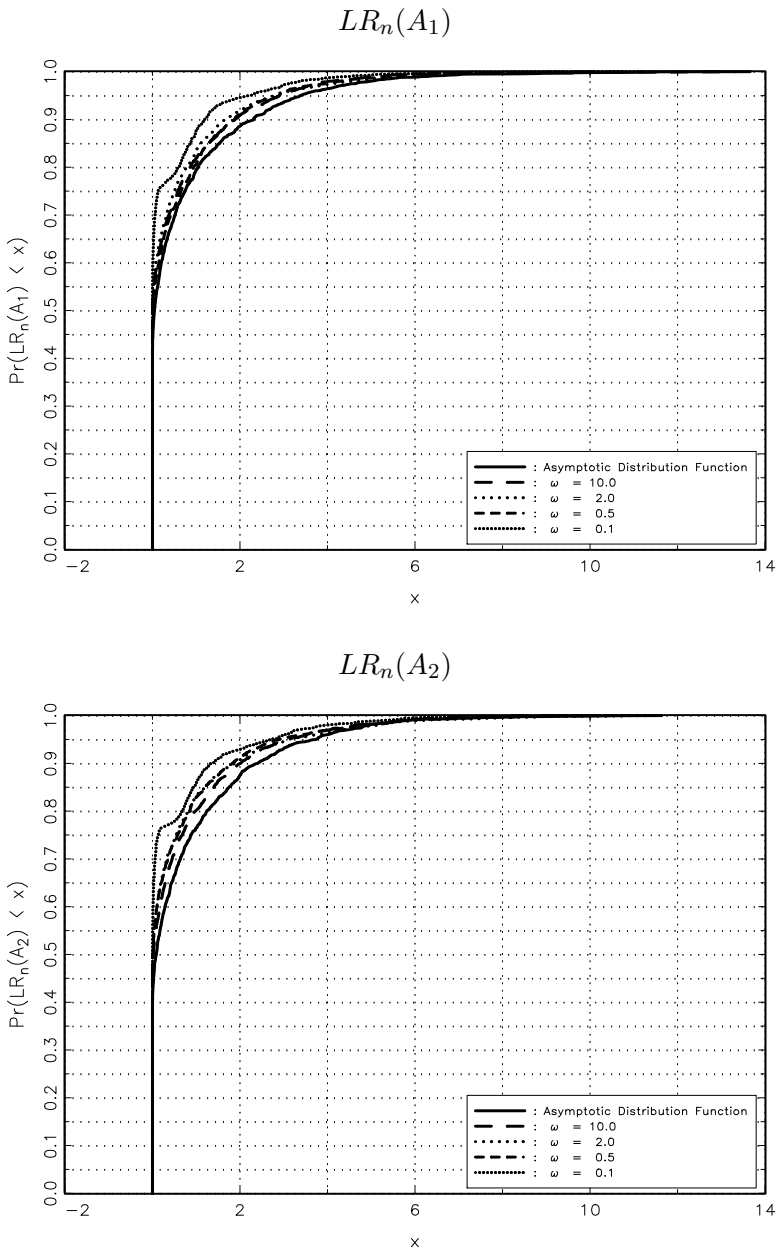

Figure 1: EMPIRICAL AND ASYMPTOTIC NULL LIMIT DISTRIBUTIONS OF THE LR TEST STATISTICS (NUMBER OF EXPERIMENT REPETITIONS: 2,000; NUMBER OF OBSERVATIONS: 20,000). The figures show the empirical distributions of the LR test statistics and the null limit distributions in Cho and White (2010). Note that the empirical null distributions of the LR test statistics approache the null limit distributions, as the grouping bin size ($\Delta \equiv 1/\omega$) reduces to zero. DGP: $Y_t^\omega \sim$ IID $\mathbb{G}[1 - \exp\{-1/(2\omega)\}]$, and the LR test statistic tests Model 0^* versus Model I^*. Here, $A_1 \equiv [3/4,\ 5/4]$ and $A_2 \equiv [3/4,\ 6/4]$.