Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models

JIN SEO CHO — School of Economics, Yonsei University

HALBERT WHITE — Department of Economics, University of California, San Diego

First version: August, 2005. This version: March, 2010

Abstract

We examine use of the likelihood ratio (LR) statistic to test for unobserved heterogeneity in duration models, based on mixtures of exponential or Weibull distributions. We consider both the uncensored and censored duration cases. The asymptotic null distribution of the LR test statistic is not the standard chi-square, as the standard regularity conditions do not hold. Instead, there is a nuisance parameter identified only under the alternative, and a null parameter value on the boundary of the parameter space, as in Cho and White (2007a). We accommodate these and provide methods delivering consistent asymptotic critical values. We conduct a number of Monte Carlo simulations, comparing the level and power of the LR test statistic to an information matrix (IM) test due to Chesher (1984) and Lagrange multiplier (LM) tests of Kiefer (1985) and Sharma (1987). Our simulations show that the LR test statistic generally outperforms the IM and LM tests. We also revisit the work of van den Berg and Ridder (1998) on unemployment durations and of Ghysels, Gourieroux, and Jasiak (2004) on interarrival times between stock trades, and, as it turns out, affirm their original informal inferences.

Key Words: Unobserved Heterogeneity, Mixture Models, Likelihood Ratio Test, Search Theory, Interarrival Times.

Subject Class: C12, C22, C24, C41, C80, J22, J64.

Acknowledgements: The authors are most grateful to the Co-editor, Takeshi Amemiya, and three anonymous referees for their helpful comments. Also, we gratefully acknowledge the empirical data kindly provided by Eric Ghysels and Joann Jasiak and helpful discussions with Robert Davies, Estate Khmaladze, Peter Thomson, David Vere-Jones, and the workshop participants at NZESG05 (Auckland), NZSA (Dunedin), statistics seminar of VUW, and FEMES2006 (Beijing). Taul Cheong provided excellent research assistance for our Monte Carlo simulations. Finally, the first author appreciates the research support from Korea University under Grant No. K0823641.

1 Introduction

Econometric specifications for duration data are often based on exponential or Weibull distributions. In labor economics, Lancaster (1979) exploits these distributions to analyze unemployment spells. In financial econometrics, Engle and Russell (1998) and Engle (2000) exploit exponential and Weibull distributions to model interarrival times of stock transactions based upon market microstructure theory of Easley and O'Hara (1992) and O'Hara (1995). The properties and analyses of various duration models in economics are well reviewed in Kiefer (1988), Lancaster (1992), and Hong and Liu (2007). The popularity of the exponential and Weibull distributions is not restricted solely to economics. They are also widely applied in epidemiology, and especially clinical trials, to model the time to occurrence of significant milestones, such as death or recovery.

The presence of unobserved heterogeneity creates serious challenges for duration models. As Heckman and Singer (1984) point out, estimated parameters can be quite sensitive to the presence of unobserved heterogeneity. Thus, testing for unobserved heterogeneity often accompanies parameter estimation. For this, Lancaster (1979) and Kalbfleisch and Prentice (1980) assume a conventional gamma distribution for the unobserved heterogeneity and test for its presence by measuring the variance of the gamma distribution. Chesher (1984) and Lancaster (1985) propose an information matrix (IM) test (White, 1982; 1994, ch.11), as the information matrix equality holds in the absence of unobserved heterogeneity. Kiefer (1985), Sharma (1987), and Prieger (2000, 2003) propose Lagrange multiplier (LM) tests exploiting the fact that the exponential and Weibull distributions can be represented using Laguerre polynomials.

In order to obtain estimates less sensitive to unobserved heterogeneity, researchers have developed a variety of flexible specifications. Heckman and Singer (1984) exploit a discrete mixture distribution for heterogeneity and estimate parameters using nonparametric maximum-likelihood estimation. Honoré (1990) assumes a Weibull distribution for the durations but does not impose a specific distribution on the heterogeneity, obtaining consistent parameter estimates. Meyer (1990) develops an estimation theory without specifying the baseline hazard, but, for convenience, retains the gamma distribution for heterogeneity. For the most part, however, the theories in this literature specify hazard or conditional mean functions in which the unknown coefficients multiply the explanatory variables, implementing a form of linearity. As yet, it is unknown how these theories may need to be modified when specifying general nonlinear models for the conditional mean. As we see in our empirical time-series application, a conditional mean equation embodying linearity can easily be misspecified. On the other hand, Horowitz (1999) provides a non-parametric estimation procedure under the condition that the distribution of heterogeneity is highly smooth; but if the associated heterogeneity is a discrete mixture, then this procedure may not work. There thus remains a need for tests of unobserved heterogeneity, whether continuous or discrete, applicable when the hazard function or conditional mean does not necessarily embody linearity.

The main goal of this paper is therefore to develop convenient test statistics having power comparable to or better than standard test statistics for unobserved heterogeneity and applicable to flexibly specified models. To achieve our goal, we develop log-likelihood ratio (LR) test statistics based upon a discrete mixture of exponential or Weibull distributions. We consider both the uncensored and censored duration cases. To develop our tests, we apply results of Cho and White (2007a), who build on work of Andrews (1999, 2001). Cho and White (2007a) use discrete mixtures to develop a test for regime-switching in a time series context and derive the asymptotic distribution of the LR statistic under the null hypothesis of a single regime. Here, we obtain the asymptotic null distributions of our LR statistics under the hypothesis of no heterogeneity. As pointed out in the literature, the null distribution of the LR statistic in such situations is model dependent (e.g., Hartigan, 1985; Chernoff and Lander, 1995; Cho and White, 2007a). Thus, the null distributions derived for the discrete mixtures of binomials in Chernoff and Lander (1995) or normals in Cho and White (2007a) cannot be applied either to the discrete mixture of exponentials or the mixture of Weibulls. We separately derive the null distributions for these two cases. In particular, we find that the censored case differs substantially from the uncensored case. We provide procedures to obtain consistent asymptotic critical values for our LR test statistics, and we conduct large scale Monte Carlo simulations under various heterogeneity assumptions, including continuous distributions for heterogeneity. As we see, our LR test statistics have well behaved levels, and they appear to be consistent not only for discrete forms of heterogeneity, but also for many continuous heterogeneous alternatives.

Another goal of this paper is to revisit the empirical analyses of van den Berg and Ridder (1998) and Ghysels, Gourieroux, and Jasiak (2004). Search theory predicts that unemployment durations follow an exponential distribution for each segmented labor market (see Yoon, 1981; van den Berg and Ridder, 1998; and the references therein). Nevertheless, most empirical papers in the literature admit that, due to unobserved heterogeneity, exponential distributions are hard to verify empirically. van den Berg and Ridder (1998) estimate reduced form equations for labor markets in the Netherlands and identify unobserved heterogeneity using LR statistics based on mixtures of exponential and Weibull distributions. Nevertheless, van den Berg and Ridder (1998) use an informal procedure to test for unobserved heterogeneity. Here we provide a formal testing procedure valid under their assumptions; as it turns out, we affirm their original inferences. Ghysels, Gourieroux, and Jasiak (2004) note that an accurate analysis of financial market liquidity needs to accommodate both conditional mean and variance at the same time; for this they propose the stochastic volatility duration (SVD) model, which extends the exponential duration model with gamma heterogeneity to the time-series context. We examine their data using the methods proposed here, and affirm the presence of unobserved heterogeneity, motivating use of the SVD model. As we discuss, correct specification of the conditional mean equation plays a key role for inference in this context.

The plan of this paper is as follows. In Section 2, we derive asymptotic null distributions for our LR statistics under the null hypothesis of no unobserved heterogeneity. We first treat the uncensored case. As we show, discrete mixtures of the exponential or Weibull distributions have different asymptotic null distributions. In particular, we represent the limiting behaviors of the LR statistic as functions of different Gaussian processes. We provide alternate representations of these processes; these yield consistent asymptotic critical values. We next discuss the censored case. As we see, the censored case differs substantially from the uncensored case. In particular, more involved methods based on those of Hansen (1996) are required to obtain consistent asymptotic critical values. In Section 3, we conduct Monte Carlo simulations comparing the performance of LR-based tests with IM and LM tests for both uncensored and censored cases. These experiments corroborate the results of Section 2 and provide useful insight into specifying key aspects of the parameter space. Section 4 contains our analysis of van den Berg and Ridder's (1998) and of Ghysels, Gourieroux, and Jasiak's (2004) data on Netherlands unemployment durations and interarrival durations of stock transactions, respectively. We provide a summary and conclusions in Section 5. The Appendix contains formal statements of the relevant assumptions and a link to proofs of our results.

2 Modeling Durations with Weibull Distributions

2.1 The Data Generating Process

Let $\{(Y_t, X_t')'\}$ be a strictly stationary geometric β–mixing stochastic process, where Y_t is scalar-valued and X_t is $\mathbb{R}^k-$ valued , $k \in \mathbb{N}$. Y_t represents a duration and must thus be non-negative. In the time-series context, for example, in the study of interarrival times for stock transactions, X_t can contain lagged values of Y_t. The β-mixing condition permits X_t to be a function of the entire history of Y_t, as is true for autoregressive conditional duration (ACD) models. (See Carrasco and Chen (2002).) As a special case, $\{(Y_t, X_t')'\}$ can be an independent identically distributed (IID) process, as is suitable for cross-section data. In this case, t indexes individuals in the sample, and the elements of X_t are duration-invariant. For either time-series or cross-section data, we assume without loss of generality that X_t does not contain a constant term, as explained below.

Throughout, we assume that our interest focuses on the conditional distribution of Y_t given X_t. For the uncensored case and in the absence of unobserved heterogeneity, we suppose that the conditional probability density function (PDF) of the durations is defined by the Weibull density

$$f(y \mid X_t; \delta^*, \beta^*, \gamma^*) = \delta^* \gamma^* g(X_t; \beta^*) y^{\gamma^* - 1} \exp(-\delta^* g(X_t; \beta^*) y^{\gamma^*}),$$

for some $(\pi^*, \delta^*, \beta^{*\prime}, \gamma^*) \in [0,1] \times D \times B \times \Gamma \subset [0,1] \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^+$, where $g(X_t; \cdot)$ satisfies the differentiability condition given in the Appendix. We leave the functional form of $g(X_t; \beta^*)$ unspecified, as this form differs from application to application. In labor economics, $g(X_t; \beta^*)$ is often Cox's (1972) proportional hazard function, and in many cases, $g(X_t; \beta^*) = \exp(X_t'\beta^*)$ is the chosen specification. In finance, $g(X_t; \beta^*)$ is often associated with the conditional mean of $Y_t \mid X_t$. In particular, ACD models specify autoregressive functions for the conditional mean of $Y_t \mid X_t$; Engle and Russell (1998) provide details. The explanatory variables X_t are assumed not to contain a constant, as an intercept can be captured by δ^* or one of the β^*'s. For example, if we let $g(X_t; \beta^*) = \exp(X_t'\beta^*)$ as in Cox's (1972) proportional hazard model, then $\delta^* g(X_t; \beta^*) = \exp(\ln(\delta^*) + X_t'\beta^*)$. Thus, $\ln(\delta^*)$ is now the coefficient of the constant term. Alternatively, if δ^* cannot be moved into $g(X_t; \beta^*)$, then we may define $g(X_t; \beta^*)$ as, e.g., $\tilde{g}(X_t'\kappa^* + \lambda^*)$ by letting $\beta^* = (\kappa^{*\prime}, \lambda^*)'$, say, so that one of the β^*'s is now the intercept. Thus, there is no loss of generality in assuming that X_t does not contain a constant. The conditional Weibull distribution reduces to a conditional exponential distribution when the duration dependence parameter γ^* is one. Engle and Russell (1998) assume this and estimate the conditional mean duration by quasi-maximum likelihood methods.

Lancaster (1979) and Ghysels, Gourieroux, and Jasiak (2004), among others, are concerned with the presence of unobserved heterogeneity in δ^*, because this results in a negative bias for the estimate of γ^* (Lancaster, 1979) and overdispersion of duration data. To deal with unobserved heterogeneity, it is standard in the literature to assume that the durations are distributed according to the density defined by

$$\int f(y \mid X_t; \delta, \beta^*, \gamma^*) h(\delta) d\mu(\delta),$$

where μ is a $\sigma-$finite measure absolutely continuous with respect to the distribution H of δ^*, so that $h \equiv dH/d\mu$ is its Radon-Nikodým density. In the case of no heterogeneity, μ is counting measure and h has a point mass at the single point, δ^*. Thus, testing the hypothesis that h has zero variance is one way to test for unobserved heterogeneity. Lancaster (1979) and Kalbfleisch and Prentice (1980) specify h to be the gamma density with unit mean. (Here, heterogeneity is continuously distributed, so μ is Lebesgue measure.) Also, Ghysels, Gourieroux, and Jasiak (2004) extend the ACD models of Engle and Russell (1998) to various stochastic volatility duration (SVD) models, treating the gamma heterogeneous duration model as a special case of the SVD model for interarrival times of stock transactions. As they show, it is important to associate the dispersion of the heterogeneously generated duration with its conditional mean for accurate financial market liquidity analysis. Heckman and Singer (1984) note that although the gamma distribution is convenient, it is somewhat *ad hoc*; on the other hand, Abbring and van den Berg (2007) show that a large class of mixed proportional hazard models have heterogeneity distributions whose tails can be well approximated by a gamma distribution.

Here, we follow Nickell (1979), Heckman and Singer (1984) and van den Berg and Ridder (1998) by considering alternatives with unobserved heterogeneity generated as a discrete mixture, with point masses at δ_1^* and δ_2^*, say. In this case, the PDF of Y_t given X_t can be written as the mixture of conditional Weibull PDFs, defined by

$$\pi^* f(y \mid X_t; \delta_1^*, \beta^*, \gamma^*) + (1 - \pi^*) f(y \mid X_t; \delta_2^*, \beta^*, \gamma^*),$$

with $(\pi^*, \delta_1^*, \delta_2^*, \beta^{*\prime}, \gamma^*) \in [0, 1] \times D \times D \times B \times \Gamma \subset [0, 1] \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^+$. Heterogeneity is absent when $\pi^* = 0$, $\pi^* = 1$, or $\delta_1^* = \delta_2^*$. We develop a likelihood ratio test for this hypothesis.

The main motivations for choosing this discrete mixture alternative are: its simplicity and the computational convenience of its associated inference; its prominence in the literature (e.g., Heckman and Singer, 1984; van den Berg and Ridder, 1998); and its previously demonstrated ability to yield tests detecting a broad range of heterogeneous alternatives. As to computational convenience, we show that under the null of no unobserved heterogeneity, the LR statistic weakly converges to a function of a Gaussian process whose covariance structure is relatively simple, so that we can straightforwardly simulate the null distribution to obtain asymptotic critical values. As to power, results of Stinchcombe and White (1998) suggest that discrete mixture models may be comprehensively revealing[1], in which case arbitrary heterogeneous alternatives would be consistently detected with a sufficient number of discrete mixtures. In many cases, even a two-point mixture may be expected to have good power. (See also Ferguson, 1983.) Thus, we expect that the mixture alternative may well deliver tests with good power even when the alternative is misspecified. Essentially, this alternative acts as an approximation to an unknown heterogeneity distribution, yielding tests capable of detecting a wide range of heterogeneous alternatives. Our Monte Carlo simulations of Section 3 verify this.

2.2 Discrete Mixtures

Continuing with our consideration of the uncensored case, we specify a model $\mathcal{M}$ as a collection of data generating processes, represented using their conditional PDFs, as in White (1994). The homogeneous null model is

$$\mathcal{M}_o \equiv \{f(\cdot \mid \cdot; \delta, \beta, \gamma) : \delta, \beta, \gamma \in D \times B \times \Gamma\}.$$

Let ζ^* be the probability limit of the quasi-maximum likelihood estimator of δ^* based on $\mathcal{M}_o$. Under mild regularity conditions and in the absence of unobserved heterogeneity, $\zeta^* = \delta^*$, the true

[1] Stinchcombe and White (1998) define a comprehensively revealing set of functions for testing correct specification of a conditional mean: When $H \subset L^q(X)$ with $X \in \mathbb{R}^k$, if for any $e \in L^p(X)$, $1/p + 1/q = 1$, there is an $h \in H$ such that $\int e \cdot h \, dP \neq 0$ then we say that H is *totally revealing*, where $L^p(X)$ is the space of functions f such that $|f|^p$ is integrable. $H \subset M_b(B)$ is said to be *comprehensively revealing* if it is totally revealing for $L^p(\mu)$ for every $q \in [1, \infty]$ and every finite signed measure μ supported on B, where $M_b(B)$ and $L^p(\mu)$ denote the set of bounded measurable functions on B and real-valued functions f such that $[\int |f(r)|^p d\text{v}\mu]^{1/p} < \infty$, respectively. Here, $\text{v}\mu$ is the variation of μ, and $\mu(A) \equiv P(X^{-1}(A))$ for any $A \in \mathcal{B}^k$, the Borel σ-algebra on $\mathbb{R}^k$.

value. Under the heterogeneous mixture, $\zeta^* = \delta^*_{o,a}$, say; this is determined by $\delta^*_1, \delta^*_2, \beta^*, \gamma^*$, and the distribution of the X_t's. The analysis becomes especially straightforward by defining $\alpha^*_1 \equiv \delta^*_1/\zeta^*$ and $\alpha^*_2 \equiv \delta^*_2/\zeta^*$ (necessarily $\zeta^* \neq 0$) and writing the heterogeneous PDF as

$$\pi^* f(y \mid X_t; \alpha^*_1\zeta^*, \beta^*, \gamma^*) + (1-\pi^*) f(y \mid X_t; \alpha^*_2\zeta^*, \beta^*, \gamma^*).$$

We parameterize the heterogeneous PDF as

$$f_a(y \mid X_t; \pi, \alpha_1, \alpha_2, \beta, \gamma) = \pi f(y \mid X_t; \alpha_1\zeta^*, \beta, \gamma) + (1-\pi) f(y \mid X_t; \alpha_2\zeta^*, \beta, \gamma), \qquad (1)$$

so that the heterogeneous alternative model is

$$\mathcal{M}_a \equiv \{f_a(\cdot \mid \cdot; \pi, \alpha_1, \alpha_2, \beta, \gamma) : (\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0,1] \times A \times A \times B \times \Gamma\},$$

where $A \equiv \{\alpha : \alpha\zeta^* \in D\}$. In what follows, we assume that A is convex and compact.

2.3 The Likelihood Ratio Statistic

For the null and alternative models the expected log-likelihoods are

$$L_o(\delta, \beta, \gamma) \equiv E[\ln f(Y_t \mid X_t; \delta, \beta, \gamma)] \quad \text{and}$$

$$L_a(\pi, \alpha_1, \alpha_2, \beta, \gamma) \equiv E[\ln f_a(Y_t \mid X_t; \pi, \alpha_1, \alpha_2, \beta, \gamma)].$$

Under the null of homogeneity, $L_o(\zeta^*, \beta^*, \gamma^*) = L_a(\pi^*, \alpha^*_1, \alpha^*_2, \beta^*, \gamma^*)$, whereas under the heterogeneous alternative, $L_o(\zeta^*, \beta^*_{o,a}, \gamma^*_{o,a}) < L_a(\pi^*, \alpha^*_1, \alpha^*_2, \beta^*, \gamma^*)$, where $\zeta^*, \beta^*_{o,a}, \gamma^*_{o,a}$ solve

$$\max_{(\delta,\beta,\gamma)\in D\times B\times\Gamma} L_o(\delta, \beta, \gamma).$$

This suggests testing the null of homogeneity based on an estimate of $L_o(\zeta^*, \beta^*_{o,a}, \gamma^*_{o,a}) - L_a(\pi^*, \alpha^*_1, \alpha^*_2, \beta^*, \gamma^*)$, as the properties of the Kullback-Leibler information criterion ensure that this difference is zero if and only if the DGP is homogeneous. We thus consider tests based on the LR statistic:

$$LR_n \equiv 2 \left\{ \sum_{t=1}^{n} \ln[f_a(Y_t \mid X_t; \hat{\pi}_n, \hat{\alpha}_{1n}, \hat{\alpha}_{2n}, \hat{\beta}_{an}, \hat{\gamma}_{an})] - \sum_{t=1}^{n} \ln[f(Y_t \mid X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)] \right\},$$

where n is the sample size, and $(\hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)$ and $(\hat{\pi}_n, \hat{\alpha}_{1n}, \hat{\alpha}_{2n}, \hat{\beta}_{an}, \hat{\gamma}_{an})$ are the maximum-likelihood estimators (MLEs) obtained under the null and alternative hypotheses respectively. That is, $(\hat{\delta}_n, \hat{\beta}_n,$ $\hat{\gamma}_n)$ solves

$$\max_{(\delta,\beta,\gamma)\in D\times B\times\Gamma} \sum_{t=1}^{n} \ln[f(Y_t \mid X_t; \delta, \beta, \gamma)],$$

$(\hat{\pi}_n, \hat{\delta}_{1n}, \hat{\delta}_{2n}, \hat{\beta}_{an}, \hat{\gamma}_{an})$ solves

$$\max_{(\pi,\delta_1,\delta_2,\beta,\gamma)\in[0,1]\times D\times D\times B\times\Gamma} \sum_{t=1}^{n} \ln[f_a(Y_t \mid X_t; \pi, \delta_1/\hat{\delta}_n, \delta_2/\hat{\delta}_n, \beta, \gamma)],$$

and $\hat{\alpha}_{1n} \equiv \hat{\delta}_{1n}/\hat{\delta}_n$, $\hat{\alpha}_{2n} \equiv \hat{\delta}_{2n}/\hat{\delta}_n$. The latter are maximum likelihood estimators as a consequence of the invariance property of maximum likelihood.

The null hypothesis of homogeneity encompasses three distinct possibilities:

$$\mathcal{H}_o : \; \pi^* = 1 \text{ and } \alpha_1^* = 1; \alpha_1^* = \alpha_2^* = 1; \text{ or } \pi^* = 0 \text{ and } \alpha_2^* = 1.$$

The heterogeneous alternative is

$$\mathcal{H}_a : \pi^* \in (0,1) \text{ and } \alpha_1^* \neq \alpha_2^*.$$

2.3.1 Asymptotic Null Distribution

Tests of $\mathcal{H}_o$ based on LR_n are non-standard. If $\pi^* = 1$ (resp. $\pi^* = 0$), then α_2^* (resp. α_1^*) is not identified, so that there is a nuisance parameter present only under the alternative (see Davies, 1977; 1987). Further, $\pi^* = 1$ (resp. $\pi^* = 0$) is on the boundary of its parameter space, violating the standard interiority condition. Alternatively, if $\alpha_1^* = \alpha_2^*$, then π^* is not identified. Consequently, the LR statistic does not have the standard chi-square distribution under the null.

Instead, as we show, the LR statistic converges weakly under the null to a function of a Gaussian process. Specifically, under the null

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

where $\mathcal{G}$ denotes a standard mean zero Gaussian process, that is, it has mean zero and variance one for every α, but its covariance structure may differ from case to case. Here, different models imply different covariance structures.

This situation is also discussed by Cho and White (2007a), who test for the presence of two regimes in a regime-switching process against a single regime, using the LR test. They assume that the transition probability matrix of a regime-switching process has identical rows, so that the associated likelihood function can be viewed as a mixture probability. Their null model is nested in the mixture model, implying that their LR test also has the identical structure for $\mathcal{H}_o$ as above. A result similar to (2) follows from their analysis, as well; in this sense, the current paper exploits the methodology in Cho and White (2007a) in the framework of unobserved heterogeneity. For further details, see Cho and White (2007a).

To state our results, we adopt the convention that when $\beta = 0$, then $g(X_t; \beta) = const$, the case of no regressors. To analyze important special cases, we consider certain restricted versions of $\mathcal{M}_o$ and $\mathcal{M}_a$, with their corresponding constrained maximum likelihood estimators. For example, we refer to the models

$$\mathcal{M}_{o|\beta=0,\gamma=1} \equiv \{f(\cdot \mid \cdot; \delta, \beta, \gamma) : (\delta, \beta, \gamma) \in D \times \{0\} \times \{1\}\},$$

$$\mathcal{M}_{a|\beta=0,\gamma=1} \equiv \{f_a(\cdot \mid \cdot; \pi, \alpha_1, \alpha_2, \beta, \gamma) : (\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0,1] \times A \times A \times \{0\} \times \{1\}\}$$

as discrete mixtures of exponentials with no regressors and to

$$\mathcal{M}_{o|\beta=0} \equiv \{f(\cdot \mid \cdot; \delta, \beta, \gamma) : (\delta, \beta, \gamma) \in D \times \{0\} \times \Gamma\}$$

$$\mathcal{M}_{a|\beta=0} \equiv \{f_a(\cdot \mid \cdot\, ; \pi, \alpha_1, \alpha_2, \beta, \gamma) : (\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0,1] \times A \times A \times \{0\} \times \Gamma\}$$

as discrete mixtures of Weibulls with no regressors. When these are specified, the LR statistic is computed with the indicated constraint(s) imposed. When the DGP is an element of a given model, we say that model is *correctly specified*, following White (1994).

We state our first formal results as follows; for conciseness, formal assumptions appear in the Appendix.

THEOREM 1: *Suppose Assumptions A1 to A4 and* $\mathcal{H}_o$ *hold, and* $\inf A > 1/2$. *Further,* (i) *if* $\mathcal{M}_{o|\beta=0,\gamma=1}$ *is correctly specified, then* $LR_n \Rightarrow \mathcal{LR}_1 \equiv \sup_{\alpha \in A} (\max[0, \mathcal{G}_1(\alpha)])^2$, *where* $\mathcal{G}_1$ *is a standard Gaussian process with*

$$E[\mathcal{G}_1(\alpha)\mathcal{G}_1(\alpha')] = \frac{(2\alpha-1)^{1/2}(2\alpha'-1)^{1/2}}{\alpha+\alpha'-1}; \tag{3}$$

(ii) *if* $\mathcal{M}_{o|\gamma=1}$ *is correctly specified, then* $LR_n \Rightarrow \mathcal{LR}_1$;
(iii) *if* $\mathcal{M}_{o|\beta=0}$ *is correctly specified, then* $LR_n \Rightarrow \mathcal{LR}_2 \equiv \sup_{\alpha \in A} (\max[0, \mathcal{G}_2(\alpha)])^2$, *where* $\mathcal{G}_2$ *is a standard Gaussian process with*

$$E[\mathcal{G}_2(\alpha)\mathcal{G}_2(\alpha')] = \frac{(\alpha-1)(\alpha'-1)\{(\alpha-1)(\alpha'-1)/(\alpha+\alpha'-1) - (6/\pi^2)\ln(\alpha)\ln(\alpha')\}/\alpha\alpha'}{\left[\frac{(1-\alpha)^2}{\alpha^2}\left\{\frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{6}{\pi^2}\ln(\alpha)^2\right\}\right]^{1/2}\left[\frac{(1-\alpha')^2}{\alpha'^2}\left\{\frac{(\alpha'-1)^2}{(2\alpha'-1)} - \frac{6}{\pi^2}\ln(\alpha')^2\right\}\right]^{1/2}}; \tag{4}$$

(iv) *if* $\mathcal{M}_o$ *is correctly specified, then* $LR_n \Rightarrow \mathcal{LR}_3 \equiv \sup_{\alpha \in A} (\max[0, \mathcal{G}_3(\alpha)])^2$, *where* $\mathcal{G}_3$ *is a standard Gaussian process with*

$$E[\mathcal{G}_3(\alpha)\mathcal{G}_3(\alpha')] = \frac{(\alpha-1)(\alpha'-1)\{(\alpha-1)(\alpha'-1)/(\alpha+\alpha'-1) - (\xi^*/\gamma^{*2})\ln(\alpha)\ln(\alpha')\}/\alpha\alpha'}{\left[\frac{(1-\alpha)^2}{\alpha^2}\left\{\frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{\xi^*}{\gamma^{*2}}\ln(\alpha)^2\right\}\right]^{1/2}\left[\frac{(1-\alpha')^2}{\alpha'^2}\left\{\frac{(\alpha'-1)^2}{(2\alpha'-1)} - \frac{\xi^*}{\gamma^{*2}}\ln(\alpha')^2\right\}\right]^{1/2}}, \tag{5}$$

where

$$\xi^* \equiv \left[\frac{\pi^2}{6\gamma^{*2}} + \text{var}[\phi_t^*] - \text{cov}[\phi_t^*, d_t^*]'(\text{var}[d_t^*])^{-1}\text{cov}[\phi_t^*, d_t^*]\right]^{-1},$$

$$\phi_t^* \equiv -\{\ln(\zeta^*) + \tilde{\gamma} - 1 + \ln[g(X_t;\beta^*)]\}/\gamma^*,$$

$$d_t^* \equiv \nabla_\beta \ln[g(X_t;\beta^*)],$$

and $\tilde{\gamma} \approx 0.57721$ *is Euler's constant* (*see formula* 1.20 *in Spiegel*, 1968).

Theorem 1 is straightforwardly proved by verifying the regularity conditions given in theorem $6(a)$ of Cho and White (2007a), which builds on work of Andrews (1999, 2001). The weak limits in Theorem 1 involve the 'max' operator, whereas the 'min' operator is used in Cho and White (2007a). The symmetry of Gaussian processes ensures that this does not matter. Also, the results of Theorem 1 are derived under the null $\pi^* = 0$, whereas Cho and White (2007a) impose $\pi^* = 1$. By the symmetry of the mixture, either null yields the same result.

The condition $\inf A > 1/2$ is key to obtaining the given covariance structures. Without this, the given integrations yielding eqs. (3), (4), and (5) cannot be verified. In a related (but non-duration) context, Liu, Pasarica, and Shao (2003) consider a mixture of gamma distributions without explanatory variables and test the same hypothesis as ours (homogeneity vs. two-point mixture), obtaining the same covariance structure as given in Theorem 1(i). Nevertheless, they suppose that $\inf A > 1$, restricting the scope of possible alternatives. Theorem 1 suggests that their covariance structure is still obtainable even when $\inf A > 1/2$, under mild regularity conditions.

The covariance structures in Theorem 1 deserve further comment, as estimating additional nuisance parameters in general yields Gaussian processes with more complicated covariance structures. First, estimating the power coefficient γ^* yields parameter estimation errors that substantially modify the covariance structure. The covariance structure (3) cannot easily be viewed as a special case of (4) and (5). Unless the terms $6\ln(\alpha)^2/\pi^2$ and $\xi^* \ln(\alpha)^2/\gamma^{*2}$ in (4) and (5) respectively disappear, there is no way for (4) and (5) to equal (3); nor can these terms be avoided as long as γ^* is estimated. On the other had, estimating β^* does not modify the covariance structure nearly as much. Note that with the exponential distribution, we obtain the same Gaussian process whether or not we estimate β^*. Also, for the Weibull distribution, estimating β^* only modifies the coefficient of $\ln(\alpha)^2$, from $6/\pi^2$ to ξ^*/γ^{*2}. As detailed in Corollary 1 below, these two values may be identical if some special but not too stringent conditions hold.

Note that the given Gaussian processes are obtained when $\pi^* = 0$ (or, symmetrically, $\pi^* = 1$), which accounts for the α index. These processes are also relevant to the weak limit of the LR statistic obtained when $\alpha_1^* = \alpha_2^*$. By applying Cho and White (2007a), we find that when α approaches 1, the probability limit of $(\max[0, \mathcal{G}_j(\alpha)])^2$, $j = 1, 2, 3$, is the weak limit of the LR statistic obtained when $\alpha_1^* = \alpha_2^*$. Thus, Theorem 1 takes account of all elements of the null hypothesis.

We obtain the given Gaussian processes $\mathcal{G}$ with their particular covariance structures as a consequence of our choice of discrete mixture alternative. If a different heterogeneity distribution is assumed, then different and potentially much more complicated limiting distributions may obtain. The convenient forms for $\mathcal{G}$ found here are a particular benefit of using the two-point discrete mixture for unobserved heterogeneity.

The particular form of $g(X_t;\beta^*)$ does not affect the conclusions of Theorem 1(i)-(iii). For any function satisfying the conditions given in the Appendix, the same covariance structure is obtained. Thus, the same asymptotic null distribution holds for DGPs with different conditional means and different model specifications. Only the conditional distribution of $Y_t \mid X_t$ and the fact that ζ^* appears as the coefficient of $g(X_t;\beta^*)$ determine the covariance structures (3) and (4).

Nevertheless, as pointed out by one of the referees, the condition in A2(i) that $g(X_t;\cdot)$ is four times continuously differentiable might be thought restrictive. As long as the alternative model contains elements near the null model with $\alpha_1^* = \alpha_2^* = 1$, this smoothness is necessary. Cho and White (2007a, section 2.3) explain why: essentially, second-order Taylor series analysis breaks down in this neighborhood; instead, fourth-order analysis is required. Consequently, if $g(X_t;\cdot)$ is only two times continuously differentiable, then testing for unobserved heterogeneity should focus on a modified alternative to the null model. Specifically, one can restrict the parameter space A in Assumption A2(ii) to $A^\circ(\delta) \equiv A \backslash \mathcal{N}_\delta(1)$, where $\mathcal{N}_\delta(1) \equiv \{\alpha \in A : |\alpha - 1| < \delta\}$ and $\delta \in (0, 1/2)$. Subtracting the neighborhood $\mathcal{N}_\delta(1)$ from A to obtain $A^\circ(\delta)$ is equivalent to eliminating from the original alternative model $\mathcal{M}_a$ the alternative PDFs near the null model with $\alpha_1^* = \alpha_2^* = 1$. Liu, Pasarica, and Shao (2003) avoid the identification problem associated with $\alpha_1^* = \alpha_2^* = 1$ by assuming $\inf A > 1$.

We note that in Theorem 1(iv), ϕ_t^* is computed as

$$\phi_t^* = E\left\{ \frac{\partial}{\partial \delta} \ln[f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)] \frac{\partial}{\partial \gamma} \ln[f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)] \mid X_t \right\}.$$

This is one of the off-diagonal elements of the outer product of the gradient (OPG) representation of the information matrix associated with $\mathcal{M}_o$; numerous elements of this matrix contain this factor, as can be seen in the proof of Theorem 1(iv). This plays an important role in the covariance structure of the Gaussian process.

Examining the proof of Theorem 1, we see that the covariance structure is a function of α and α' in which only the ratio between the null and alternative intercepts matters in determining the asymptotic null distribution of the LR statistic. This is the main reason for parameterizing our alternative model as we do in (1). Other noteworthy technical aspects of Theorem 1 are similar to Theorem 6(a) of Cho and White (2007a). We therefore do not repeat these explanations here.

Note that the covariance structures given in eq.(3) for the mixture of exponentials is the same whether or not regressors are present. In contrast, for the mixture of Weibulls, different asymptotic null distributions are generally obtained with and without regressors. The covariance structure (4) is a special case of (5). That is, if $\xi^* = 6\gamma^{*2}/\pi^2$, then the covariance structure (5) simplifies to (4), so that the distributions of $\mathcal{G}_2$ and $\mathcal{G}_3$ coincide, in which case we write $\mathcal{G}_3 \overset{d}{=} \mathcal{G}_2$. A sufficient condition for this is that $\phi_t^* = 0$, implying that $g(X_t; \beta^*) = \exp(1 - \tilde{\gamma})/\zeta^*$ by the definition of ϕ_t^* in Theorem 1(iv). A much weaker sufficient condition holds for several cases popular in the literature, in particular for Cox's (1972) proportional hazards model, as Corollary 1 now verifies.

COROLLARY 1: *Suppose the conditions of Theorem* 1(iv) *hold. If in addition* $g(X_t; \beta^*) = \exp(X_t'\beta^*)$ *then* $\mathcal{G}_3 \overset{d}{=} \mathcal{G}_2$.

This is easily proved by showing that

$$\text{var}[\phi_t^*] - \text{cov}[\phi_t^*, d_t^*]'(\text{var}[d_t^*])^{-1}\text{cov}[\phi_t^*, d_t^*] = 0,$$

which implies $\xi^* = 6\gamma^{*2}/\pi^2$.

We provide two further remarks relevant to Theorem 1 and Corollary 1. First, there is considerable prior literature on discrete mixtures. Hartigan (1985) examines the LR statistic for the discrete mixture of normals without unknown parameters under the null. Chernoff and Lander (1995) analyze the LR statistic for the discrete mixture of binomials and show that the covariance structure of the limiting Gaussian process converges to that given by Hartigan (1985) as the number of points with positive mass tends to infinity. Chen and Chen (2001) examine both discrete mixtures of normals with known variance and discrete mixtures of Poissons. All of these studies consider only unconditional distributions for scalar random variables. Dacunha-Castelle and Gassiat (1999) generalize these analyses by studying mixtures of conditional distributions, and Cho and White (2007a) further extend the scope of mixture models. For example, as Cho and White (2007a) demonstrate, the discrete mixture of normals with unknown means and variances cannot be analyzed in the framework of Dacunha-Castelle and Gassiat (1999), although it falls into the framework of Cho and White (2007a).

Second, the parameter space A affects the asymptotic null distribution, similar to Hartigan (1985). If A is replaced by a larger set, say $\tilde{A} \supset A$, then the asymptotic null distribution associated with $\tilde{A}$ is first-order stochastically dominated by that associated with A. This implies that the relevant parameter space needs to be carefully specified prior to testing for unobserved heterogeneity. If the duration model is implied by a structural equation specifying A, then that A should be used for inference on unobserved heterogeneity. Otherwise, A should be determined by considerations of size and power. We discuss this in detail in Section 3, where we conduct our Monte Carlo experiments.

It is not an easy task to compute the analytical distribution for the maximum of an arbitrary Gaussian process. Davies (1977, 1987) provides tail lower bounds for the distribution of the maximum when the underlying processes are Gaussian and chi-square respectively. Cho and White (2007a) apply Davies's method, as well as the so-called comparison method of Piterbarg (1996) to obtain another lower bound. These methods are effective for obtaining conservative critical values. Hansen (1996) provides a procedure similar to the wild bootstrap that can yield consistent p-values. For the cases of Theorem 1, however, consistent asymptotic critical values can be straightforwardly obtained using orthonormal bases, as in Chernoff and Lander (1995) and Cho and White (2007a, b). We provide these in Theorem 2.

THEOREM 2: *Let* $\{Z_k : k = 0, 1, 2, \cdots\}$ *be an IID sequence of* $N(0,1)$ *random variables.* (i) *Let* $\mathcal{G}_1$ *be as in Theorem* $1(i)$. *Then* $\mathcal{G}_1 \stackrel{d}{=} \bar{\mathcal{G}}_1$, *where for each* $\alpha \in A$, $\bar{\mathcal{G}}_1(\alpha) \equiv \sum_{k=2}^{\infty} a_k(\alpha) Z_k$, *and*

$$a_k(\alpha) \equiv \left[\frac{(\alpha-1)^4}{\alpha^2(2\alpha-1)}\right]^{-1/2} \left(\frac{\alpha-1}{\alpha}\right)^k.$$

(ii) *Let* $\mathcal{G}_2$ *be as in Theorem* $1(iii)$. *When* $\inf A > 1/2$, $\mathcal{G}_2 \stackrel{d}{=} \bar{\mathcal{G}}_2$, *where for each* $\alpha \in A$, $\bar{\mathcal{G}}_2(\alpha) \equiv \sum_{k=1}^{\infty} b_k(\alpha) Z_k$, *and*

$$b_k(\alpha) \equiv \left[\frac{(\alpha-1)^2}{\alpha^2}\left\{\frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{6}{\pi^2}\ln(\alpha)^2\right\}\right]^{-1/2} \left\{\left(\frac{\alpha-1}{\alpha}\right)^k - \frac{6}{\pi^2 k}\ln(\alpha)\right\}\left(\frac{\alpha-1}{\alpha}\right).$$

(iii) (a) *Let* $\mathcal{G}_3$ *be as in Theorem* $1(iv)$. *Let* $\theta^* \equiv (\delta^*, \beta^{*\prime}, \gamma^*)'$. *When* $\inf A > 1/2$, $\mathcal{G}_3 \stackrel{d}{=} \bar{\mathcal{G}}_3^*$, *where for each* $\alpha \in A$, $\bar{\mathcal{G}}_3^*(\alpha) \equiv \sum_{k=0}^{\infty} c_k(\alpha, \theta^*) Z_k$,

$$c_0(\alpha,\theta^*) \equiv \left[\frac{(\alpha-1)^2}{\alpha^2}\left\{\frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{\xi^*}{\gamma^{*2}}\ln(\alpha)^2\right\}\right]^{-1/2}\left\{\frac{6}{\pi^2} - \frac{\xi^*}{\gamma^{*2}}\right\}^{1/2}\left(\frac{\alpha-1}{\alpha}\right)\ln(\alpha),$$

and for $k = 1, 2, ...,$

$$c_k(\alpha, \theta^*) \equiv \left[\frac{(\alpha-1)^2}{\alpha^2} \left\{ \frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{\xi^*}{\gamma^{*2}} \ln(\alpha)^2 \right\} \right]^{-1/2} \left\{ \left(\frac{\alpha-1}{\alpha} \right)^k - \frac{6}{\pi^2 k} \ln(\alpha) \right\} \left(\frac{\alpha-1}{\alpha} \right).$$

(b) *Further, let* $\hat{\theta}_n \equiv (\hat{\delta}_n, \hat{\beta}_n', \hat{\gamma}_n)'$ *be such that* $\hat{\theta}_n = \theta^* + o_p(1)$ *and let* $\hat{\xi}_n \equiv \xi(\hat{\theta}_n)$, *where* $\xi(\cdot)$ *is a continuous function of* θ *such that* $\hat{\xi}_n = \xi^* + o_p(1)$. *Then* $\sup_{\alpha \in A} |\widehat{\mathcal{G}}_{3,n}(\alpha) - \bar{\mathcal{G}}_3^*(\alpha)| = o_p(1)$, *where* $\widehat{\mathcal{G}}_{3,n}(\alpha) \equiv \sum_{k=0}^{\infty} c_k(\alpha, \hat{\theta}_n) Z_k$,

$$c_0(\alpha, \hat{\theta}_n) \equiv \left[\frac{(\alpha-1)^2}{\alpha^2} \left\{ \frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{\hat{\xi}_n}{\hat{\gamma}_n^2} \ln(\alpha)^2 \right\} \right]^{-1/2} \left\{ \frac{6}{\pi^2} - \frac{\hat{\xi}_n}{\hat{\gamma}_n^2} \right\}^{1/2} \left(\frac{\alpha-1}{\alpha} \right) \ln(\alpha),$$

and for $k = 1, 2, ...,$

$$c_k(\alpha, \hat{\theta}_n) \equiv \left[\frac{(\alpha-1)^2}{\alpha^2} \left\{ \frac{(\alpha-1)^2}{(2\alpha-1)} - \frac{\hat{\xi}_n}{\hat{\gamma}_n^2} \ln(\alpha)^2 \right\} \right]^{-1/2} \left\{ \left(\frac{\alpha-1}{\alpha} \right)^k - \frac{6}{\pi^2 k} \ln(\alpha) \right\} \left(\frac{\alpha-1}{\alpha} \right).$$

Several remarks are relevant for Theorem 2. First, we prove Theorem 2 by showing that the covariance structures of $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$ can be represented as $\sum_{k=2}^{\infty} a_k(\alpha) a_k(\alpha')$, $\sum_{k=1}^{\infty} b_k(\alpha) b_k(\alpha')$, and $\sum_{k=0}^{\infty} c_k(\alpha, \theta^*) c_k(\alpha', \theta^*)$ respectively. Second, in Theorem 2 (ii) and (iii), we specifically impose $\inf A > 1/2$. The orthonormal expansion in Theorem 2(ii) can be effectively obtained when[2] $\inf A \geq 1/2$. Nevertheless, $\sum_{k=1}^{\infty} (\alpha-1)^{2k}/\alpha^2$ is not defined when $\alpha = 1/2$, so $\inf A > 1/2$ is necessary. Third, $(6/\pi^2 - \xi^*/\gamma^{*2}) \geq 0$ as shown in the Appendix, so that $c_0(\,\cdot\,, \theta^*)$ is well-defined. Fourth, we do not specify the specific form of $\xi(\,\cdot\,)$, as there are numerous possibilities yielding the same result. For example, $\xi(\,\cdot\,)$ can be specified by replacing cov$[\phi_t^*, d_t^*]$ and var$[d_t^*]$ appearing in ξ^* with their sample analogs based on estimators $(\hat{\gamma}_n, \hat{\beta}_n')'$. Fifth, although the critical values obtained from Theorem 2 $(i$, ii, $iii.a)$ are asymptotically precise, they may be imprecise when the sample size n is too small, as usual. Moreover, these critical values are not sample dependent, in contrast to those generated by Hansen's (1996) procedure. The critical value bounds in Davies (1977) and Piterbarg (1996) are asymptotically valid, but they are conservative and may therefore be imprecise, even asymptotically. Nevertheless, if tests are constructed using mixtures other than those specified here, then orthonormal bases similar to those in Theorem 2 may not be available. In

[2]Formula 20.20 in Spiegel (1968) states that $\sum_{k=1}^{\infty} [(\alpha-1)/\alpha]^k / k = \ln(\alpha)$ if $\alpha \geq 1/2$.

such cases, these alternative procedures can be usefully exploited. Finally, it is well known that the exponential distribution kernel can be expanded as an infinite sum involving the Laguerre polynomials (Kiefer, 1985). We note that the representation $\bar{\mathcal{G}}_1$ can alternatively be obtained by applying a central limit theorem to these Laguerre polynomials. Essentially, the orthogonality property of the Laguerre polynomials yields the independent normal sequence of the Z_k's. Theorem 2 shows how the orthogonality conditions of the Laguerre polynomials are associated with the mixture model under the null hypothesis.

The Gaussian processes $\bar{\mathcal{G}}_1$, $\bar{\mathcal{G}}_2$, and $\bar{\mathcal{G}}_3^*$ are weighted averages of IID standard normals, with weights tending to zero as k tends to infinity. Thus, these can be easily and well approximated by $\bar{\mathcal{G}}_{1,m}$, $\bar{\mathcal{G}}_{2,m}$, and $\bar{\mathcal{G}}_{3,m}^*$, where m is a sufficiently large integer[3] and for each α,

$$\bar{\mathcal{G}}_{1,m}(\alpha) \equiv \sum_{k=2}^{m} a_k(\alpha) Z_k, \qquad \bar{\mathcal{G}}_{2,m}(\alpha) \equiv \sum_{k=1}^{m} b_k(\alpha) Z_k, \qquad \text{and} \qquad \bar{\mathcal{G}}_{3,m}^*(\alpha) \equiv \sum_{k=0}^{m} c_k(\alpha, \theta^*) Z_k.$$

Thus, simulating the empirical distributions of

$$\mathcal{LR}_{1,m} \equiv \sup_{\alpha \in A} \max[0, \bar{\mathcal{G}}_{1,m}(\alpha)]^2, \qquad \mathcal{LR}_{2,m} \equiv \sup_{\alpha \in A} \max[0, \bar{\mathcal{G}}_{2,m}(\alpha)]^2, \qquad \text{and}$$

$$\mathcal{LR}_{3,m} \equiv \sup_{\alpha \in A} \max[0, \bar{\mathcal{G}}_{3,m}^*(\alpha)]^2$$

can also closely approximate the asymptotic null distributions of the LR test statistics. Further, Theorem 2($iii.b$) ensures that

$$\mathcal{LR}_{3,m} = \widehat{\mathcal{LR}}_{3,n,m} + o_p(1),$$

where

$$\widehat{\mathcal{LR}}_{3,n,m} \equiv \sup_{\alpha \in A} \max[0, \hat{\mathcal{G}}_{3,n,m}(\alpha)]^2 \qquad \text{and} \qquad \hat{\mathcal{G}}_{3,n,m}(\alpha) \equiv \sum_{k=0}^{m} c_k(\alpha, \hat{\theta}_n) Z_k,$$

implying that simulating $\widehat{\mathcal{LR}}_{3,n,m}$ can consistently estimate the distribution of $LR_{3,m}$, based on the approximation $\hat{\mathcal{G}}_{3,n,m}$ to $\mathcal{G}_3$.

In the literature on mixture distributions, Monte Carlo simulation of the LR statistic is of-

[3]Our experience suggests that m does not have to be too large, because the coefficients of Z_k decrease geometrically, uniformly in α.

ten used in attempts to obtain an asymptotic null distribution. For this, the EM (expectation-maximization) algorithm is often used. Nevertheless, as Mosler and Seidel (2001), among others, point out, this method is not quite successful. In particular, it is difficult to ensure convergence of the Monte Carlo distributions, even for large sample sizes. We emphasize that the main reason for this is the neglect of the impact of the parameter space A. As shown by Hartigan (1985), the asymptotic distribution of the LR statistic depends crucially on the size of the parameter space. Indeed, the LR statistic becomes unbounded in probability when the parameter space is unbounded. We thus explicitly constrain the parameter space A to be compact and obtain critical values using simulation to compute the empirical distributions of $\mathcal{LR}_{1,m}$, $\mathcal{LR}_{2,m}$, $\mathcal{LR}_{3,m}$, and $\widehat{\mathcal{LR}}_{3,n,m}$. We denote these $\mathcal{LR}_{1,m}(A)$, $\mathcal{LR}_{2,m}(A)$, $\mathcal{LR}_{3,m}(A)$, and $\widehat{\mathcal{LR}}_{3,n,m}(A)$, respectively, to emphasize their dependence on the parameter space.

2.3.2 Asymptotic Power of the Test

As we assume that the model is correctly specified under the alternative, the consistency of the LR test statistic straightforwardly follows. The following theorem states this result.

THEOREM 3: *Suppose Assumptions A1 to A4 and* $\mathcal{H}_a$ *hold, and* $\inf A > 1/2$. *Further, if* $\mathcal{M}_a$ *is correctly specified, then for any sequence* $\{c_n\}$ *such that* $c_n = o(n)$, $P(LR_n \geq c_n) \to 1$ *as* $n \to \infty$.

This follows straightforwardly by applying the Kullback-Leibler information criterion (KLIC).

Nevertheless, there are several caveats to Theorem 3. First, the consistency of the LR test statistic requires careful interpretation. Rejection may be due to the violation of any of the given regularity conditions, not least of which is the correct specification assumption. For example, the correct specification of $\mathcal{M}_a$ is violated if the functional form of $g(X_t;\,\cdot\,)$ in the mixture model is misspecified, leading to the consistent rejection of $\mathcal{H}_o$. Thus, testing for the correct specification of $E[Y_t|X_t]$ may be necessary to draw proper conclusions from the LR test. Specifically, if the functional form of $g(X_t;\,\cdot\,)$ is correctly specified under the alternative, then we have

$$E[Y_t|X_t] = \Gamma(1 + 1/\gamma^*)/\{\psi^* g(X_t;\beta^*)\}^{(1/\gamma^*)}, \tag{6}$$

where $\psi^* \equiv [\pi^*/\delta_1^{*(1/\gamma^*)} + (1-\pi^*)/\delta_2^{*(1/\gamma^*)}]^{-\gamma^*}$. Thus, if one or more specification tests (e.g., Bierens, 1990; Stinchcombe and White, 1998; Cho, Huang, and White, 2008) cannot reject the hypothesis in (6), then rejecting $\mathcal{H}_o$ using the LR test is plausibly due to unobserved heterogeneity, rather than misspecification under $\mathcal{H}_a$. Alternatively, if one fails to reject $\mathcal{H}_o$, and

$$\Gamma(1 + 1/\gamma^*)[1/\{\delta^* g(X_t; \beta^*)\}^{(1/\gamma^*)}],$$

obtained by letting $\pi^* = 1$ and $\delta_1^* = \delta_*$ in (6), turns out to be misspecified for $E[Y_t|X_t]$, then there are model specification problems not signalled by the LR test, part of the "implicit null" hypothesis of the test. There is a variety of other test statistics useful for empirically examining the maintained regularity conditions for the present LR test. We recommend their use as complements to the present LR test, whenever possible.

Provided the mixture decreases entropy, a consistent test generally results even under misspecified alternatives. Nevertheless, forming a mixture with a misspecified alternative need not strictly reduce entropy (increase likelihood) when the scope of the alternative is too narrow. Although increasing the number of components in the mixture will eventually decrease entropy under mild regularity conditions, as implied by Stinchcombe and White (1998) and Ferguson (1983), it need not decrease strictly as the number of components in the mixture increases. In such cases, the LR statistic can even be degenerate under the null. For example, the LR statistic is degenerate if a normal distribution is tested against a mixture of two normals, when data follow a uniform distribution. This happens because the null parameter values are on the boundary of the parameter space. In essence, if the first-order condition on the null parameter space is non-zero when the null and alternative models have the same entropy, then the LR statistic is degenerate. Such cases can arise when models are misspecified and the model scope afforded by the parameter space or the assumed distributional conditions is too narrow to exploit the flexibility otherwise provided by the mixture.

Cho and White (2008) provide conditions for the LR statistic to be non-degenerate when the null and alternative models have the same levels of entropy for misspecified mixture models. These cases provide a class of alternatives against which the LR statistic based on exponential or Weibull mixtures does have power. Otherwise, the LR test need not have power. By verifying these conditions when the mixture model is possibly misspecified, one can gain further support for the indicated inference. We explore this further in Section 3, where we conduct Monte Carlo experiments under a wide range of DGP assumptions, including misspecified mixture models.

On the other hand, more powerful tests for unobserved heterogeneity might be obtained by enlarging the scope of the alternative model. For example, the researcher may increase the size of the parameter space A or compare likelihoods using multiple mixture alternatives. We examine the first option in our Monte Carlo experiments. The second possibility presents a significant challenge, however, as there are more unidentified nuisance parameters under the null and more than two ways to generate the null model from the alternative. To obtain the asymptotic null distribution, one must interrelate the asymptotic null distributions obtained for each case. As there is no theory in the literature addressing this, and as developing this theory is beyond the scope of this paper, we do not pursue this here.

2.3.3 The Censored Duration Case

Many duration datasets involve censoring, mainly due to incomplete sampling designs, especially in cross-section data. We now consider modified LR statistics suitable for testing unobserved heterogeneity with censored data and discuss their properties. As we show, these properties differ considerably from the uncensored case.

There are many possible censoring schemes, but for simplicity and conciseness, we restrict attention here either to (i) Type I (fixed) censoring, where for a given constant $c < \infty$, the observed duration is $Y_t^c = \min[Y_t, c]$ or to (ii) random censoring, in which case $Y_t^c = \min[Y_t, C_t]$, where C_t and (Y_t, X_t) are independent. In either case, we observe $\{(Y_t^c, D_t, X_t')'\}$, where $D_t = 1$ if $Y_t^c = Y_t$, and $D_t = 0$ otherwise.

The simplest situation is Type I censoring. Here, the conditional density of $(Y_t^c, D_t)|X_t$ is defined by

$$f^c(y^c, d \mid X_t; \delta^*, \beta^*, \gamma^*) \equiv f(y^c \mid X_t; \delta^*, \beta^*, \gamma^*)^d [1 - F(c|X_t; \delta^*, \beta^*, \gamma^*)]^{1-d},$$

where $F(c|X_t; \delta^*, \beta^*, \gamma^*)$ is the CDF of Y_t given X_t evaluated at c. If we parameterize the heterogeneous PDF as

$$f_a^c(y^c, d|X_t; \pi, \alpha_1, \alpha_2, \beta, \gamma) = \pi f^c(y^c, d|X_t; \alpha_1\zeta^*, \beta, \gamma) + (1-\pi) f^c(y^c, d|X_t; \alpha_2\zeta^*, \beta, \gamma) \quad (7)$$

as before, then the LR statistic can be constructed as

$$LR_n^c \equiv 2\left\{\sum_{t=1}^{n}\ln[f_a^c(Y_t, D_t \mid X_t; \hat{\pi}_n^c, \hat{\alpha}_{1n}^c, \hat{\alpha}_{2n}^c, \hat{\beta}_{an}^c, \hat{\gamma}_{an}^c)] - \sum_{t=1}^{n}\ln[f^c(Y_t, D_t \mid X_t; \hat{\delta}_n^c, \hat{\beta}_n^c, \hat{\gamma}_n^c)]\right\},$$

where $(\hat{\delta}_n^c, \hat{\beta}_n^c, \hat{\gamma}_n^c)$ and $(\hat{\pi}_n^c, \hat{\alpha}_{1n}^c, \hat{\alpha}_{2n}^c, \hat{\beta}_{an}^c, \hat{\gamma}_{an}^c)$ are the censored MLEs obtained under the null and alternative hypotheses respectively as before, and $\hat{\alpha}_{1n}^c \equiv \hat{\delta}_{1n}^c/\hat{\delta}_n^c, \hat{\alpha}_{2n}^c \equiv \hat{\delta}_{2n}^c/\hat{\delta}_n^c$.

As in the previous case, the asymptotic distribution of the LR statistic converges to a standard Gaussian process under the null of no unobserved heterogeneity. That is,

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

and $\mathcal{G}^c$ denotes a particular standard Gaussian process with covariance kernel depending on the specifics of the case at hand (exponential or Weibull, with or without regressors).

For example, applying theorem 6(a) of Cho and White (2007) to the mixture of censored exponentials with regressors,

$$\mathcal{M}_{a|\gamma=1}^c \equiv \{f_a^c(\cdot \mid \cdot; \pi, \alpha_1, \alpha_2, \beta, \gamma) : (\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0,1] \times A \times A \times B \times \{1\}\},$$

we find that $LR_n^c \Rightarrow \mathcal{LR}^c \equiv \sup_{\alpha \in A}(\max[0, \mathcal{G}_1^c(\alpha)])^2$, where $\mathcal{G}_1^c(\,\cdot\,)$ is a standard Gaussian process with

$$E[\mathcal{G}_1^c(\alpha)\mathcal{G}_1^c(\alpha')] = \frac{\rho^c(\alpha, \alpha')}{\sqrt{\rho^c(\alpha, \alpha)}\sqrt{\rho^c(\alpha', \alpha')}},$$

where

$$\rho^c(\alpha, \alpha') \equiv \kappa(\alpha)\kappa(\alpha')\left[\frac{1 - \tau(\alpha)\tau(\alpha')}{1 - \kappa(\alpha)\kappa(\alpha')} - \frac{\{1 - \tilde{\tau}(\alpha)\}\{1 - \tilde{\tau}(\alpha')\}}{1 - \tilde{\tau}(1)}\right]$$

with $\kappa(\alpha) \equiv (\alpha - 1)/\alpha$,

$$\tau(\alpha) \equiv E[\exp\{-(\alpha - 1/2)c^{\gamma^*}\zeta^* g(X_t; \beta^*)\}], \quad \text{and} \quad \tilde{\tau}(\alpha) \equiv E[\exp\{-\alpha c^{\gamma^*}\zeta^* g(X_t; \beta^*)\}].$$

Observe that if, for given α, c tends to infinity, then $\tau(\alpha)$, $\tilde{\tau}(\alpha)$, and $\tilde{\tau}(1)$ tend to zero, and we obtain the covariance structure in (3) in the limit as censoring vanishes.

Nevertheless, the different covariance structures arising with censoring present new challenges for obtaining consistent critical values. One obvious difference is that, unless $g(X_t; \beta^*) \equiv const$, the functional form of $\rho^c(\,\cdot\,,\,\cdot\,)$ now generally depends on the unconditional distribution of X_t, through τ and $\tilde{\tau}$. Although these functions can be consistently estimated (replace unknown coefficients with consistent estimates and expectations with sample averages), these estimates are required for all grid points α. Even more challenging, however, is the fact that $\rho^c(\,\cdot\,,\,\cdot\,)$ has a form that does not readily lend itself to representation as an infinite sum of orthonormal bases.

Consequently, testing for unobserved heterogeneity with censored data must be conducted in a way that does not require explicit handling of the covariance structure. Fortunately, Hansen's (1996) weighted bootstrap is ideally suited for this purpose. In particular, this method can deliver valid asymptotic p-values without having to compute the associated covariance structure, supporting tests of unobserved heterogeneity for the censored case. Note, however, that use of Hansen's bootstrap is much more computationally demanding than the use of orthogonal bases possible in the uncensored case. Further, each model requires custom programming, and the complexity of the programming task is so great that just producing error-free computer code is a true logistical challenge. Cho, Cheong, and White provide detail.

For the case of random censoring, the joint conditional PDF of $(Y_t^c, D_t)|X_t$ is defined by

$$f^c(y^c, d|X_t; \delta^*, \beta^*, \gamma^*) \equiv \{f(y^c|X_t; \delta^*, \beta^*, \gamma^*)\}^d \{1 - F(y^c|X_t; \delta^*, \beta^*, \gamma^*)\}^{1-d} \{1 - G(y^c)\}^d \{g(y^c)\}^{1-d},$$

where $g(\,\cdot\,)$ and $G(\,\cdot\,)$ are the PDF and CDF of C_t, respectively. We note that $1 - G(\,\cdot\,)$ and $g(\,\cdot\,)$ enter multiplicatively, so that we can drop these terms when parameterizing the model for estimation (see, e.g., White (1994, p. 144)).

We construct the alternative model as a parameterized mixture, as above. Specifically, the alternative model is

$$\mathcal{M}_a^c \equiv \{f_a^c(\,\cdot\,,\,\cdot\,|\,\cdot\,; \pi, \alpha_1, \alpha_2, \beta, \gamma) : (\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0,1] \times A \times A \times B \times \Gamma\},$$

where

$$f_a^c(y^c, d|X_t; \pi, \alpha_1, \alpha_2, \beta, \gamma) \equiv \pi \cdot f^c(y^c, d|X_t; \alpha_1\zeta^*, \beta, \gamma) + (1 - \pi) \cdot f^c(y^c, d|X_t; \alpha_2\zeta^*, \beta, \gamma).$$

The LR statistic LR_n^c is computed analogously to that for the fixed censoring case. Once again, we have that under the null of no unobserved heterogeneity

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

where $\mathcal{G}^c$ denotes a particular standard Gaussian process depending on the specifics of the case at hand. As above, the covariance structures in the various cases depend on the distribution of (X_t, C_t) and do not readily permit orthogonal representations, requiring use of Hansen's (1996) bootstrap. For brevity and because we do not rely on the covariance structure to compute critical values, we omit analysis of these covariance structures here. We do, however, examine the performance of LR_n^c with random censoring using critical values based on Hansen's (1996) bootstrap in our Monte Carlo experiments of Section 3.2.

3 Monte Carlo Experiments

3.1 Uncensored Duration Data

In this section, we conduct Monte Carlo simulations to examine level and power properties of the LR test, comparing these to the properties of Chesher's (1984) IM test and Kiefer's (1985) and Sharma's (1987) LM tests.

To examine level properties, we consider the following DGPs:

- $Y_t \sim$ IID Exp(1);
- $Y_t \sim$ IID Weibull$(1, 1)$;
- $Y_t \mid X_t \sim$ IID Exp$(\exp(X_t))$;
- $Y_t \mid X_t \sim$ IID Weibull$(\exp(X_t), 1)$,

where Exp(δ^*) denotes the exponential distribution with coefficient δ^* and Weibull($\cdot$, $\cdot$) similarly denotes the Weibull distribution with the specified coefficients. For the third and fourth DGPs, $X_t \sim$ IID $N(0, 1)$. Here, we assume the standard normal distribution for X_t because of its familiarity and regularity, ensuring that the conditions of the Appendix hold. The finite sample results given below may be different if X_t follows a different distribution. We obtain parameter estimates from the corresponding correctly specified models, defined using the parameterizations

- $Y_t \sim$ IID Exp(δ);
- $Y_t \sim$ IID Weibull(δ, γ);
- $Y_t \mid X_t \sim$ IID Exp$(\delta \exp(X_t\beta))$;
- $Y_t \mid X_t \sim$ IID Weibull$(\delta \exp(X_t\beta), \gamma)$.

To investigate the effects of specifying different parameter spaces A for α, we consider nine different possibilities: $A_1 \equiv [7/9, 2.0]$, $A_2 \equiv [7/9, 3.0]$, $A_3 \equiv [7/9, 4.0]$, $A_4 \equiv [2/3, 2.0]$, $A_5 \equiv [2/3, 3.0]$, and $A_6 \equiv [2/3, 4.0]$, $A_7 \equiv [5/9, 2.0]$, $A_8 \equiv [5/9, 3.0]$, and $A_9 \equiv [5/9, 4.0]$. The choice of the other parameter spaces has no impact on the null distribution, so we leave these unspecified.

Table 1 provides critical values for the various parameter spaces. These are computed by simulating $\bar{\mathcal{G}}_{1,m}$ and $\bar{\mathcal{G}}_{2,m}$ 100,000 times for $m = 500$. The maxima of these functions are computed by grid search with grid distance 0.01. Corollary 1 implies that it is not necessary to simulate $\hat{\mathcal{G}}_{3,n,m}$ for the present experiments. We see that Table 1 exhibits critical values increasing as the parameter space expands, as Theorem 1 implies. To enforce the infimum condition in Theorem 2, we set $\inf A = 7/9$, $\inf A = 2/3$ and $\inf A = 5/9$ respectively.

The other test statistics are defined as follows. First, we let IM_n denote Lancaster's (1984) IM test statistic. A variety of IM test statistics can be devised by combining off-diagonal elements of the IM equality with diagonal elements. For simplicity in our simulations, we focus solely on the diagonal element associated with δ, so that IM_n converges in distribution to χ_1^2 under the null. Specifically,

$$IM_n = \boldsymbol{\iota}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\boldsymbol{\iota},$$

where $\boldsymbol{\iota}$ is an $n \times 1$ vector of ones, and $\mathbf{W}$ is the $n \times 4$ matrix

$$\begin{bmatrix} \nabla'_{(\delta,\beta,\gamma)} \ln f(Y_1|X_1; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) & \left[\frac{\partial}{\partial\delta} \ln f(Y_1|X_1; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)\right]^2 + \frac{\partial^2}{\partial\delta^2} \ln f(Y_1|X_1; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) \\ \vdots & \vdots \\ \nabla'_{(\delta,\beta,\gamma)} \ln f(Y_n|X_n; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) & \left[\frac{\partial}{\partial\delta} \ln f(Y_n|X_n; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)\right]^2 + \frac{\partial^2}{\partial\delta^2} \ln f(Y_n \mid X_n; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) \end{bmatrix}$$

if γ^* is estimated. When γ^* is known, the first-order derivative with respect to γ is omitted, so that $\mathbf{W}$ is $n \times 3$.

Next, we define Kiefer's (1985) and Sharma's (1987) LM test statistics. These statistics have the form

$$LM_{d,n} \equiv n\bar{L}'_{d,n}\{\hat{V}_{d,n}\}^{-1}\bar{L}_{d,n},$$

where the subscript d indicates the order of the Laguerre polynomials underlying the LM test statistic, $\bar{L}_{d,n} \equiv n^{-1}\sum_{t=1}^{n}\hat{L}_{d,n,t}$,

$$\hat{L}_{d,n,t} \equiv [L_2(\{\hat{\delta}_n g(X_t;\hat{\beta}_n)\}^{1/\hat{\gamma}_n}Y_t;\hat{\gamma}_n),\ \cdots, L_d(\{\hat{\delta}_n g(X_t;\hat{\beta}_n)\}^{1/\hat{\gamma}_n}Y_t;\hat{\gamma}_n)]'$$

$$\hat{V}_{d,n} \equiv n^{-1}\sum_{t=1}^{n}\{\hat{L}_{d,n,t}-\bar{L}_{d,n}\}\{\hat{L}_{d,n,t}-\bar{L}_{d,n}\}',$$

and for $j = 2, 3, ...,$ $L_j(\,\cdot\,;\gamma)$ is jth Laguerre polynomial with parameter γ. The first few Laguerre polynomials are

$$L_1(x;\gamma) \equiv 1 - x^{\gamma}, \quad L_2(x;\gamma) \equiv \frac{1}{2}(x^{2\gamma} - 4x^{\gamma} + 2), \quad \text{and}$$

$$L_3(x,\gamma) \equiv \frac{1}{6}(-x^{3\gamma} + 9x^{2\gamma} - 18x^{\gamma} + 6).$$

When $\gamma = 1$, we have Kiefer's (1985) exponential distribution case. Sharma (1987) considers the Weibull distribution, in which the estimator $\hat{\gamma}_n$ appears.

Prieger (2000) points out that the asymptotic distributions of Weibull-based LM statistics are affected by parameter estimation error, whereas those of the exponential-based LM statistics are not. To accommodate parameter estimation in the Weibull case, we replace $\hat{V}_{d,n}$ with $\hat{W}_{d,n}$, so that

$$LM_n \equiv n\bar{L}'_{d,n}\{\hat{W}_{d,n}\}^{-1}\bar{L}_{d,n},$$

where $\hat{W}_{d,n} \equiv \hat{V}_{d,n} - \hat{H}_{d,n}\hat{B}_n^{-1}\hat{H}'_{d,n}$,

$$\hat{H}_{d,n} \equiv n^{-1}\sum_{t=1}^{n}\hat{L}_{d,n,t}\,\hat{S}'_{n,t}, \quad \text{and} \quad \hat{B}_n \equiv n^{-1}\sum_{t=1}^{n}\hat{S}_{n,t}\,\hat{S}'_{n,t},$$

where for $t = 1, 2, ...,$

$$\hat{S}_{n,t} \equiv \begin{bmatrix} \frac{\partial}{\partial\delta} \ln f(Y_t \mid X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) \\ \frac{\partial}{\partial\beta} \ln f(Y_t \mid X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) \\ \frac{\partial}{\partial\gamma} \ln f(Y_t \mid X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n) \end{bmatrix}.$$

Under the null, the LM statistic is distributed asymptotically as χ^2_{d-1}. For our simulations, we consider two cases, $d = 2$ and $d = 3$.

Table 2 contains our level simulation results. The number of replications is $10,000$. All statistics converge weakly to the claimed null distributions.[4] In particular, the empirical distributions of the LR statistics converge to the null distribution from above, so that rejection rates are smaller than the nominal level when the sample sizes are small. Also, the farther $\inf A$ is from 1/2, the closer empirical rejection rates are to the nominal level for smaller n, as suggested by Theorem 2. Note that for each k, the denominators of $a_k(\alpha)$, $a_k(\alpha)$, and $c_k(\alpha, \theta^*)$ diverge to ∞ as α converges to 1/2 from above. Thus, obtaining critical values by simulating $\bar{\mathcal{G}}_{1,m}(\,\cdot\,)$, $\bar{\mathcal{G}}_{2,m}(\,\cdot\,)$, and $\hat{\mathcal{G}}_{3,n,m}(\,\cdot\,)$ requires m to be very large when $\inf A$ is close to 1/2. Even with $m = 500$, this can cause noticeable level distortions, as is evident in Table 2.

This behavior can also be seen in Figures 1 and 2, where we compare empirical with asymptotic distributions for $n = 100$, 500, and $5,000$, with $A = [7/9, 2]$, $[2/3, 2]$, and $[5/3, 2]$. Figures 1 and 2 show the empirical distributions and estimated density functions for the exponential and Weibull cases, respectively. The empirical distributions and density functions are displayed in the first and second columns, respectively. In particular, as the LR test statistics have probability masses at zero, we estimate their density functions only for positive arguments using the kernel estimation with a standard normal kernel. Note that for given n, when $\inf A$ approaches $1/2$, the empirical distribution of the LR statistic diverges from its asymptotic counterpart. This discrepancy has two possible sources: the selected m could be too small; and/or n is not sufficiently large.

We also note that level distortion is influenced by $\sup A$, but not as much. This is evident from Figures 3 and 4. Note that as the level α increases, the empirical distributions of the LR statistics diverge from their asymptotic counterparts, but not as much as in Figures 1 and 2.

This behavior implies that $\inf A$ is more important for level distortions than $\sup A$. Thus, selecting $\inf A$ to be about $2/3$ appears to be reasonable, whereas $\sup A$ can be selected to be moderately large.

[4] See the footnotes to Table 2.

In contrast, the distributions of the IM and LM tests converge from below, leading to rejection rates greater than the nominal level. For small and moderate size samples, we observe the well-known level distortions of the IM test (e.g. Horowitz, 1994). The LM test levels converge relatively slowly. We note also that convergence with $d = 3$ is slower than that with $d = 2$. Interestingly, the level distortion for the LM statistic with $d = 3$ is more extreme than that of the IM statistic.

For power comparisons, we consider two families of DGPs, respectively generating data as

- $Y_t \mid (\delta_t, X_t) \sim$ IID Exp$(\delta_t \exp(X_t))$;
- $Y_t \mid (\delta_t, X_t) \sim$ IID Weibull$(\delta_t \exp(X_t), 1)$.

As above, $X_t \sim$ IID $N(0, 1)$. Within each family, we investigate the following mixture distributions for δ_t :

- Discrete mixture: $\delta_t \sim$ IID DM$(0.7370, 1.9296; 0.5)$;
- Gamma mixture: $\delta_t \sim$ IID Gamma$(5, 5)$;
- Log-normal mixture: $\delta_t \sim$ IID Log-normal$(-\ln(1.2)/2, \ln(1.2))$;
- Uniform mixture I: $\delta_t \sim$ IID Uniform$[0.30053, 2.3661]$;
- Uniform mixture II: $\delta_t \sim$ IID Uniform$[1, 5/3]$,

where DM$(a, b; p)$ yields a discrete mixture such that $P[\delta_t = a] = p$ and $P[\delta_t = b] = 1 - p$. The first four DGPs generate heterogeneity distributions with variation factor (the ratio of the variance to the square of the population mean) 0.2. This factor is of particular interest in the literature (e.g., Lancaster, 1979). The final DGP is intended to examine how the LR, the IM, and LM tests behave for a smaller variation factor; this value is about 0.02.

Note that only the first mixture distribution above is a two-component discrete mixture. The rest are more general alternatives; as noted above, we expect our statistics to have good power against these. We also note that the mixture models with $A = [2/3, 2]$, $[2/3, 3]$, $[2/3, 4]$, $[5/9, 2]$, $[5/9, 3]$, and $[5/9, 4]$ are correctly specified when unobserved heterogeneity follows the two-component discrete mixture. For the other distributions, the mixture models are misspecified.

For the first family, we estimate parameters using the exponential specification; for the second family, we use the Weibull specification.

Power simulation results are presented in Tables 3 and 4. The number of replications is $2,000$. We apply level-adjusted critical values, obtained from the simulations under the null, to accommodate small-sample level distortions. These level adjustments are typically not feasible in practice, but we use them neverthless, as meaningful power comparisons cannot be performed otherwise.

As we see in Tables 3 and 4, for the DGPs considered here, the LR tests have better power than the IM and LM tests. As expected, the LR test has power against all alternatives, not just the discrete mixture.

Our experience regarding A in terms of power can be summarized as follows. First, if the mixture model is correctly specified, selecting the smallest parameter space yields the most powerful tests. That is, for the two-component discrete mixture case, $LR_n([2/3, 2])$ yields the most powerful results for the correctly specified models. Second, however, if the mixture models are misspecified, then the simulation results are mixed. This behavior can be related to $\inf A$ and $\sup A$, as in our level studies. If $\inf A$ is too close to 1, it may lead to an inconsistent LR test. For example, when the heterogeneity follows the gamma distribution and $\inf A = 8/9$, the LR statistic has no power for $\sup A = 2$, 3, and 4. This is mainly because $\inf A$ is too close to 1 for the MLE to be in the interior of A even asymptotically, as anticipated by Cho and White (2008). Nevertheless, when $\inf A$ is $7/9$, $2/3$, or $5/9$, the LR tests are consistent for every misspecified mixture model. Further, they have better power as $\inf A$ or $\sup A$ gets smaller. These results suggest that choosing $\inf A$ to be about $2/3$ gives reasonable power, at least against the popular alternatives in the literature, whereas $\sup A$ can be selected without too much constraint.

This is the same lesson we learn from our study of level distortion, so this appears to be a reasonable recommendation when the heterogeneity distribution is unknown under the alternative. Accordingly, we apply this choice for our empirical applications.

Observe that in only a few cases does $LM_{2,n}$ have power different from that of the IM tests; they are similar in most cases. This is mainly because $LM_{2,n}$ is based on $L_{n,2}$, which is virtually the same as the second-order derivative of the likelihood function exploited by the IM test. This helps explain the inferior power of the LM tests relative to the LR tests. Nearly identical results obtain when regressors are absent, but we do not report these for brevity.

Despite the lower power of the IM and LM tests, these still have utility. As mentioned above, the asymptotic null distributions of the LR statistics are model dependent, whereas those of the IM and LM tests are not.

3.2 Censored Duration Data

To examine level properties for our LR statistic under censoring, we consider the following DGP with random censoring:

- $Y_t^c \equiv \min[Y_t, C_t]$, $Y_t|X_t \sim$ IID Weibull$(\exp(X_t), 1)$, and $C_t \sim$ IID Exp(1),

where C_t and (Y_t, X_t) are independent. As noted above, we need not model the distribution of C_t in constructing our LR statistic.

We restrict attention to this censored Weibull (C-Weibull) DGP for several reasons. First, we do not consider the censored exponential distribution case, as it is essentially a special case of the censored Weibull case. Second, we do not consider Type I (fixed) censoring as this is a special case of the randomly censored case. As the C-Weibull DGP contains the other cases in this sense, we view the C-Weibull DGP as both representative and most salient. Third, and most significantly, the extreme computational burden associated with simulating the weighted bootstrap forces us to focus attention strictly on the most salient case.

The specific procedure for conducting Hansen's (1996) weighted bootstrap is as follows. First, we compute the score for each grid point $\alpha \in A$ as $\hat{S}_{nt}(\alpha) := \{\hat{D}_{nt}(\alpha)\}^{-\frac{1}{2}}\hat{W}_{nt}(\alpha)$, where

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

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

$$\hat{R}_{nt}(\alpha) \equiv f^c(Y_t, D_t|X_t; \alpha\hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)/f^c(Y_t, D_t|X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n),$$

$$\hat{U}_{nt} \equiv \nabla_{(\delta,\beta,\gamma)}\ln[f^c(Y_t, D_t|X_t; \hat{\delta}_n, \hat{\beta}_n, \hat{\gamma}_n)],$$

and we consider nine parameter spaces for A as before. That is, $A = [7/9, 2]$, $[7/9, 3]$, $[7/9, 4]$, $[2/3, 2]$, $[2/3, 3]$, $[2/3, 4]$, $[5/9, 2]$, $[5/9, 3]$, and $[5/9, 4]$. The grid distance is 0.01, as well.

Second, we generate $\mathcal{Z}_{jt} \sim IID\, N(0,1)$, $t = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, J$, and simulate the asymptotic distribution of the LR statistic by computing the empirical distribution of

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

Due to the immense computational burden involved, we let $J = 500$, following Hansen's (1996) recommendations. Although J is not large enough for highly precise estimates, Hansen's (1996) simulation experience suggests that this will yield solid results at relatively modest cost.

Third, we compare the LR test statistic to this empirical distribution by computing the proportion of simulated outcomes exceeding the LR test statistic value. That is, the empirical level is $\hat{p}_n \equiv J^{-1} \sum_{j=1}^{J} I[LR_n^c(A) < \mathcal{LR}_{jn}(A)]$, where $I[\cdot]$ is the usual indicator function. Finally, we repeat this process $N = 5,000$ times, generating $\hat{p}_n^{(i)}, i = 1, ..., N$, and we compute the proportion of outcomes whose $\hat{p}_n^{(i)}$ is less than the specified level (e.g., $\boldsymbol{\alpha} = 5\%$). That is, we compute $N^{-1} \sum_{i=1}^{N} I[\hat{p}_n^{(i)} < \boldsymbol{\alpha}]$. Under the null, this converges to the significance level corresponding to the specified nominal level, $\boldsymbol{\alpha}$, because $\hat{p}_n$ weakly converges to U[0, 1] in distribution by theorem 2 of Hansen (1996). On the other hand, under the alternative, $\hat{p}_n$ converges to zero in probability because $LR_n^c(A)$ is not bounded in probability by the same argument as in Theorem 3, whereas $\mathcal{LR}_{jn}(A)$ is bounded in probability. This holds because $\mathcal{LR}_{jn}(A)$ is constructed from a zero mean process involving multiples of $\mathcal{Z}_{jt}$.

We present these estimates in Table 5, where we let the level of significance be 5%. We see that the estimated empirical rejection rates are below 5% but approach 5% as the sample size n increases. This is also what Hansen (1996) finds. In contrast to other choices of A, we observe that when A is $[2/3, 2]$, $[2/3, 3]$, $[2/3, 4]$, or $[5/9, 2]$ the empirical rejection rates are relatively stable and are relatively close to 5%, even for modest sample sizes,

For power comparisons, we slightly modify the previous considered alternative DGPs. In particular we consider

- $Y_t^c \equiv \min[Y_t, C_t]$, $Y_t, | \ (\delta_t, X_t) \sim$ IID Weibull$(\delta_t \exp(X_t), 1)$, and $C_t \sim$ IID Exp(1).

As above, $X_t \sim$ IID $N(0, 1)$. We thus investigate the same mixture distributions for δ_t as for the uncensored case. Specifically, we consider the same mixture models for δ_t as in Section 3.1: discrete, gamma, log-normal, uniform I, and uniform II. We call these five DGPs the C-discrete mixture, C-gamma mixture, C-log-normal mixture, C-uniform mixture I, and C-uniform mixture II respectively.

The simulation procedure is identical to the level case. The only difference is that the null DGPs are replaced by these five alternatives.

Simulation results are presented in Table 6. The number of repetitions is now $N = 2,000$, so that the empirical rejection rates are computed by $N^{-1}\sum_{i=1}^{N} I[\hat{p}_n^{(i)} < 0.05]$. Under the alternative, this estimate should converge to 100%. The LR test statistic is indeed consistent against these alternative DGPs, as can be seen in Table 6.

For censored data, our experience regarding selecting A in terms of power can be summarized as follows. First, even if the mixture model is correctly specified, selecting the smallest parameter space does not necessarily yield the most powerful test. This result differs from the uncensored case. Here, all parameter spaces yielded more or less similar power results, and there is no apparent obvious order for the parameter spaces in terms of power. Second, if the mixture models are misspecified, then the most powerful tests are obtained when $\inf A$ and $\sup A$ are smallest; this also differs from the uncensored case. These results suggest that when durations are censored, letting $\inf A$ be close to 1/2 can yield better power and that $\sup A$ needs to be smaller than in the uncensored case. Specifically, for all alternative DGPs considered here, the most powerful results are obtained by letting $A = [5/9, 2]$.

We also conduct IM and LM heterogeneity tests for censored data, analogous to those for the uncensored case, denoted by $IM_n^c, LM_{2,n}^c$, and $LM_{3,n}^c$. The results appear in Tables 5 and 6. As before, we see serious level distortions for these statistics. In comparing power, we therefore again use level-adjusted crticial values. For the most part, the LR tests dominate the IM test and $LM_{3,n}^c$. On the other hand, $LM_{2,n}^c$ outperforms most of the LR tests. Because of the level distortions, this advantage cannot be practically exploited in samples less than 1,000 or 2,000. Nevertheless, in larger samples, $LM_{2,n}^c$ may have a practically useful modest power advantage.

To close this section, we note that Cho, Cheong, and White (2010) perform further simulations comparing LR tests based on asymptotic and weighted bootstrap critical values. They study the same environments as here, considering only the uncensored case; they find that the weighted bootstrap outperforms the asymptotic approach in terms of both level and power. In particular, the weighted bootstrap level is relatively less sensitive to the choice of A. Cho, Cheong, and White (2010) also document the greater computational costs associated with the weighted bootstrap. Comparing their results to those presented here for the censored case, we see that results for the uncensored case are superior to those for the censored case. This should not be surprising, as the uncensored case preserves information absent from the censored case.

4 Applications

4.1 Re-examination of van den Berg and Ridder (1998)

van den Berg and Ridder (1998) estimate reduced form equations for Netherlands unemployment durations and test for unobserved heterogeneity using a LR statistic. The data are the OSA labor supply data collected by the Netherlands Organization for Strategic Labour Market Research. van den Berg and Ridder (1998) estimate three models: the exponential, the mixture of exponentials, and the mixture of Weibulls using 366 observations:

$$Y_t \mid X_t \sim \text{ IID Exp}(\delta \exp(X_t'\beta)), \tag{8}$$

$$Y_t \mid X_t \sim \text{ IID } \pi\text{Exp}(\alpha_1 \zeta^* \exp(X_t'\beta)) + (1-\pi)\text{Exp}(\alpha_2 \zeta^* \exp(X_t'\beta)), \tag{9}$$

$$Y_t \mid X_t \sim \text{ IID } \pi\text{Weibull}(\alpha_1 \zeta^* \exp(X_t'\beta), \gamma) + (1-\pi)\text{Weibull}(\alpha_2 \zeta^* \exp(X_t'\beta), \gamma), \tag{10}$$

respectively. The regressors X_t are 'age,' 'education,' and 'occupation level dummy' variables. To identify these mixture models, they let $\alpha_1 = 0.28$ and 0.13 in the mixture of exponentials and Weibull distributions respectively and estimate the other parameters by maximum likelihood, including ζ^*. This differs somewhat from our approach using $\hat{\delta}_n$. Nevertheless, the LR statistics obtained are numerically identical by the invariance principle, and our estimators can be easily obtained from theirs. Note that each model is recursively nested in the latter, so that comparing the log-likelihoods is equivalent to testing the following hypotheses:

- (8) versus (9): testing for unobserved heterogeneity;
- (8) versus (10): testing for unobserved heterogeneity and $\gamma^* = 1$;
- (9) versus (10): testing $\gamma^* = 1$.

van den Berg and Ridder (1998) identify unobserved heterogeneity by rejecting the first hypothesis, presumably using asymptotic critical values from the chi-square distribution[5]; they do not reject

[5] van den Berg and Ridder (1998) reject the null and infer the presence of significant heterogeneity without explicitly describing their procedure. Also, they (implicitly) assume correct specification for the conditional mean equation.

the exponential distribution in the third hypothesis. Nevertheless, these results are ambiguous, as the standard chi-square critical values for the LR test statistic do not apply, as discussed above. For their first hypothesis, the LR statistic value is about 10.28. If we cast the parameter estimate given in van den Berg and Ridder (1998, Table III) into the present context, the global maxima must have been achieved by $\hat{\delta}_n = 0.2787$ and $(\hat{\zeta}_n, \alpha_1, \hat{\alpha}_{2n}) = (0.0773, 0.2800, 1.630)$ for the null and alternative models respectively. They let $\alpha_1 = 0.2800$ to avoid the identification problem instead of fixing ζ as in our analysis, but this does not matter, as it can be easily rephrased. For this, we fix ζ at $\hat{\delta}_n$ and conversely estimate α_1 and α_2 of the current study by $\alpha_1\hat{\zeta}_n/\hat{\delta}_n$ and $\hat{\alpha}_{2n}\hat{\zeta}_n/\hat{\delta}_n$ respectively. These are respectively 0.7764 and 4.5200, implying that the smallest parameter space yielding this global maximum in terms of A of current paper is $A = [0.7764, 4.5200]$. Although we can compute the associated critical value from this estimated parameter space, we choose a somewhat larger parameter space, $A = [2/3, 5]$, in order to accommodate parameter estimation error residing in $\hat{\delta}_n$ and the lessons of the previous section. The associated p-value is 0.00205, which is quite small. Thus, we affirm van den Berg and Ridder's (1998) original inference.

4.2 Re-examination of Ghysels, Gourieroux, and Jasiak (2004)

Ghysels, Gourieroux, and Jasiak (2004) (GGJ) argue that an accurate liquidity analysis of financial markets must accommodate both the conditional mean and dispersion of interarrival durations of stock transactions. As the conditional mean and variance of intertrade durations are indicators for market liquidity and risk respectively, data analysis simultaneously accommodating these becomes more accurate than analysis focusing only on one of these.

GGJ specify a model for interarrival durations of stock transactions driven by two unobserved factors, $F_t = (F_{1t}, F_{2t})'$. Their stochastic volatility duration (SVD) model is specified by the parameterization

$$Y_t = \frac{1}{a}\frac{H(1, F_{1t})}{H(b, F_{2t})} \qquad \text{where} \tag{11}$$

$$F_t = \sum_{j=1}^{p} \Phi_j F_{t-j} + \varepsilon_t, \tag{12}$$

where Y_t is the interarrival duration of stock transactions; a, b, and Φ_j, $j = 1, ..., p$ are model parameters; $H(b, F_{2t})$ is such that $H(b, F_{2t}) = G(b, \Phi(F_{2t}))$, where $G(b, \cdot)$ is the quantile function of the $\gamma(b,b)$ distribution, with Φ the standard normal CDF; and ε_t is a Gaussian white noise of dimension two.

The SVD model extends the exponential model with gamma heterogeneity in Kalbfleisch and Prentice (1990) and Lancaster (1990). Note that $H(b, F_{2t})$ is a gamma variable with conditional mean 1 and and variance $1/b$ respectively, given $\mathcal{F}_t := \sigma(F_t, F_{t-1}, \ldots)$, so that if $b = 1$, $H(1, F_{1t})$ is a conditionally exponential random variable. Thus, if $\varepsilon_t \sim IID\, N(0, I_2)$ and $\Phi_j = 0$ for $j = 1, 2, \ldots, p$, then the model in (11) reduces to the exponential model with gamma heterogeneity popularly specified for duration data in labor economics. Here, F_{1t} and F_{2t} jointly determine $E[Y_t|\mathcal{F}_t]$ and var$[Y_t|\mathcal{F}_t]$. In particular, unobserved heterogeneity is determined by $H(b, F_{2t})$. If the variance of $H(b, F_{2t})$ is zero, then the conditional distribution of Y_t given $\mathcal{F}$ must be the exponential distribution, and unobserved heterogeneity is absent. The SVD model has the same motivation as do stochastic volatility models, in the sense that the two unobserved Gaussian components (F_{1t}, F_{2t}) determine the conditional mean and variance of Y_t.

GGJ (2004) estimate their SVD model using the duration between stock trades of Alcatel in July 1996. Their data represent $5,000$ observations extracted from the Paris Stock Exchange, and they specify the following linear ACD (LACD) model for $E[Y_t|\mathcal{F}_t]$:

$$\Psi_t = \omega + \alpha Y_{t-1} + \rho_1 \Psi_{t-1} + \rho_2 \Psi_{t-2}, \tag{13}$$

where $\Psi_t := E[Y_t|\mathcal{F}_t]$. They estimate their SVD model by the simulated method of moments, and they graphically compare the implied properties of this model with the ACD model without unobserved heterogeneity. They conclude that the SVD model provides a better description of the data than the ACD model without unobserved heterogeneity based on an evaluation of their graphical results. See GGJ for more information on their data and their empirical analysis.

Here, we revisit their data and apply the methodology of the current study. The goal of this review is twofold. First, this illustrates application of our LR statistic in a time series context, with particular attention paid to the correct specification of the conditional mean. Second, this illustrates the empirical motivation for a model of the conditional variance of duration, based on the apparent presence or absence of unobserved heterogeneity. For this, we apply the LR test based upon the mixture of conditional exponential distributions nested within the SVD model. We view this application of our tests as complementing GGJ's graphical testing procedures for the SVD.

Our testing procedure is as follows. First, we test whether or not the linear ACD (LACD) model given in (13) is correctly specified for $E[Y_t|\mathcal{F}_t]$. Thus, the hypotheses of interest can be stated as follows:

$$\mathcal{H}_o^{LACD}: \text{ LACD is correct for } E[Y_t|\mathcal{F}_t]; \text{ vs. } \mathcal{H}_a^{LACD}: \text{ LACD is misspecified for } E[Y_t|\mathcal{F}_t].$$

For this, we exploit a specification test given by Cho, Huang, and White (2008). Their test extends Bierens's (1990) specification test; this involves running a functional regression of

$$\hat{G}_t(\gamma) := \left[\frac{1}{\hat{\Psi}_t}\left(\frac{Y_t}{\hat{\Psi}_t} - 1\right)\right]\psi(Y_{t-1}\gamma),$$

on deterministic functions of γ, where $\psi(x)$ is the logistic function, so that $\psi(x) := 1/[1 + \exp(-x)]$;

$$\hat{\Psi}_t = \hat{\omega}_n + \hat{\alpha}_n Y_{t-1} + \hat{\rho}_{1n}\hat{\Psi}_{t-1} + \hat{\rho}_{2n}\hat{\Psi}_{t-2}; \tag{14}$$

and $(\hat{\omega}_n, \hat{\alpha}_n, \hat{\rho}_{1n}, \hat{\rho}_{2n})$ is the quasi-MLE (QMLE) based on the exponential distribution. Note that $\hat{G}_t(\gamma)$ is constructed by multiplying the logistic function by the score with respect to ω. If the conditional mean equation is correctly specified, then the population mean of $\hat{G}_t(\gamma)$ has to be zero for every γ. Cho, Huang, and White (2008) test this property using functional regression. Their theoretical results and simulations show that this test can consistently detect misspecified models by selecting functional regressors appropriately. (See Cho, Huang, and White (2008) for further details.)

We apply their test, choosing the regressors to be the constant, γ, and γ^2. That is, after minimizing

$$\frac{1}{n}\sum_{t=1}^{n}\int_{-0.5}^{0.5}\{\hat{G}_t(\gamma) - \xi_0 - \xi_1\gamma - \xi_2\gamma^2\}^2 d\gamma$$

with respect to (ξ_0, ξ_1, ξ_2), we test whether or not the probability limits of the estimated coefficients $(\xi_0^*, \xi_1^*, \xi_2^*)$ are equal to zero using Wald statistics. If the conditional mean equation is correctly specified, then the limits have to be zero, and the Wald test follows the chi-square distribution under the null. Here, selecting the functional regressors to be the constant, γ, and γ^2, corresponds to selecting particular alternatives to the correct specification hypothesis.

Our inference results are given in Table 7. The first panel of Table 7 shows the specification testing results. $\mathcal{W}_n^{(c)}$ and $\mathcal{W}_n^{(j)}$ test $(\xi_0^*, \xi_1^*, \xi_2^*) = (0,0,0)$ and $\xi_j^* = 0$ for $j = 0$, 1, and 2, respectively. Using the Hochberg-Bonferroni (HB) bound (Hochberg, 1988) yields a p-value upper bound[6] valid for the four multiple hypothesis tests conducted here. We report these in the rows labeled "HB Bound" in Table 7. In the present instance, the HB bound signals a highly significant result. We thus conclude that the LACD model is misspecified for $E[Y_t|\mathcal{F}_t]$.

In the same panel, we also present the LR statistics for testing unobserved heterogeneity using the mixture of exponentials. In particular, our mixture model specifies

$$\frac{\alpha_j}{\hat{\omega}_{on} + \alpha Y_{t-1} + \rho_1 \Psi_{t-1} + \rho_2 \Psi_{t-2}}$$

for $\alpha_j \delta^* g(X_t; \beta^*)$, where $j = 1$ and 2, and $\hat{\omega}_{on}$ is the estimate for ω^* under the null. $\hat{\omega}_{on}$ is used for model identification. We also obtain $\hat{\alpha}_{1n} = 1.1402$ and $\hat{\alpha}_{2n} = 0.9650$, and $\hat{\pi}_n \approx 0$ without any parameter space condition under the alternative, and the LR statistic value is 0.48884. We consider three parameter spaces for the LR statistics. That is, the parameter space A for α_1 and α_2 is set to $[2/3, 2]$, $[2/3, 3]$, and $[2/3, 4]$. These are selected based on our experience with the previously reported Monte Carlo experiments. We note that the obtained $\hat{\alpha}_{1n}$ and $\hat{\alpha}_{2n}$ are feasible for A chosen to be $[2/3, 2]$, $[2/3, 3]$, and $[2/3, 4]$. When A is set to $[2/3, 4]$, $\hat{\alpha}_{1n} = 1.14$ and $\hat{\alpha}_{2n} = 0.96$. We also present in Table 7 the associated p-values using the asymptotic distributions given by Theorem 2(i). Note that the associated p-values are much greater than 0.05, so that unobserved heterogeneity appears to be absent.

Nevertheless, this appearance is deceptive, due to the model misspecification for conditional mean. With the conditional mean misspecified, no valid conclusions can be drawn as to unobserved heterogeneity.

We therefore extend the LACD model. Fernandes and Grammig (2006) consider alternative models to LACD, obtained by applying the Box-Cox transformation. Following their motivation, we extend their augmented ACD models from the most restricted case. First, we consider a power ACD (PACD) model specified as

$$\Psi_t^\lambda = \omega + \alpha Y_{t-1}^\lambda + \beta_1 \Psi_{t-1}^\lambda + \beta_2 \Psi_{t-2}^\lambda. \tag{15}$$

[6]Hochberg's method involves ordering the $p-$values from testing r hypotheses as $p_{(1)}, ..., p_{(r)}$ and computing the bound as $HB = \min_{j=1,...,r}(r - j + 1)p(j)$

We view (15) as a modest extension of the LACD model, as it reduces to the linear ACD model by letting $\lambda = 1$. We apply the same specification tests as before for the following hypotheses:

$$\mathcal{H}_o^{PACD}: \text{ PACD is correct for } E[Y_t|\mathcal{F}_t]; \quad \text{vs.} \quad \mathcal{H}_a^{PACD}: \text{ PACD is misspecified for } E[Y_t|\mathcal{F}_t],$$

and present the results in the second panel of Table 7. The HB bound signals a moderately significant result, leading us again to reject the correct specification assumption.

Accordingly, we extend the PACD to the following flexible power ACD (FPACD) model:

$$\Psi_t^{\lambda} = \omega + \alpha Y_{t-1}^{\nu} + \beta_1 \Psi_{t-1}^{\lambda} + \beta_2 \Psi_{t-2}^{\lambda}. \tag{16}$$

Note that the power coefficient of Y_{t-1} now differs from that of Ψ_t; the power and linear ACD models are thus special cases of this model. We now test the hypotheses:

$$\mathcal{H}_o^{FPACD}: \text{ FPACD is correct for } E[Y_t|\mathcal{F}_t]; \quad \text{vs.} \quad \mathcal{H}_a^{FPACD}: \text{ FPACD is misspecified for } E[Y_t|\mathcal{F}_t],$$

and present the results in the third panel of Table 7. Here, the HB bound gives a $p-$value far from standard significance levels, suggesting that the FPACD model is not misspecified. If so, we can conduct valid tests for unobserved heterogeneity using the LR test. As before, we specify our mixture model as

$$\frac{\alpha_j}{(\hat{\omega}_{on} + \alpha Y_{t-1}^{\nu} + \rho_1 \Psi_{t-1}^{\lambda} + \rho_2 \Psi_{t-2}^{\lambda})^{(1/\lambda)}}$$

for $\alpha_j \delta^* g(X_t; \beta^*)$, where $j = 1$ and 2, and $\hat{\omega}_{on}$ is estimated using the null model. Using LR tests based upon each of the three parameter spaces $[2/3, 2]$, $[2/3, 3]$, and $[2/3, 4]$, we reject the null of no unobserved heterogeneity. The HB bound also strongly signals rejection.

This rejection motivates the need for an adequate model for the conditional dispersion of duration. Although the conditional mean equation we arrive at differs from that in GGJ, our analysis affirms that the exponential distribution misspecifies the conditional distribution, so that the SVD model accommodating conditional dispersion of duration becomes more appropriate than the exponential model without heterogeneity.

5 Conclusion

Unobserved heterogeneity is an important concern in many areas of economics. In particular, unobserved heterogeneity in duration models, especially in labor economics and financial econometrics, can create serious difficulties for inference.

We provide tests for unobserved heterogeneity in duration models using the LR statistic, based on mixtures of exponential or Weibull distributions. We treat both the uncensored and the censored duration cases. The asymptotic null distributions of the LR statistics are not standard chi-square, due to non-standard features of the null hypothesis: parameters on the boundary of the parameter space and nuisance parameters identified only under the alternative. Nevertheless, the asymptotic distributions of the LR statistic under the null can be represented as functions of particular Gaussian processes. In the uncensored case, these processes can be conveniently represented using orthonormal bases. In turn, these representations deliver straightforward consistent estimates of asymptotic critical values. In the censored case, orthonormal representations are not readily available. Nevertheless, a weighted bootstrap procedure of Hansen (1996) can be applied to deliver consistent estimates of asymptotic critical values.

We conduct a range of Monte Carlo simulation experiments that provide insight into finite sample level and power properties of the LR statistics. We find that, for the DGPs considered in this paper, these properties are reasonable. In particular, for the uncensored case, the LR test outperforms the IM test of Chesher (1984) and the LM tests of Kiefer (1985) and Sharma (1987). For the censored case, we also find reasonable level and power properties. We thus recommend using the LR tests developed here, whenever the underlying regularity conditions can be plausibly assumed.

Finally, we apply our results to revisit the unemployment duration analysis of van den Berg and Ridder (1998) and the stock interarrival duration analysis of Ghysels, Gourieroux, and Jasiak (2004). As it turns out, our analysis affirms the original inferences drawn by these authors.

6 Appendix

6.1 Assumptions

A1: (i) $\{(Y_t, X_t')'\}$ *is a strictly stationary geometric* β*-mixing process with* β*-mixing coefficients* $\beta_\tau \leq c\rho^\tau$ *for some* $c > 0$ *and* $\rho \in [0, 1)$, *where* Y_t *is* $\mathbb{R}^+$*-valued,* X_t *is* $\mathbb{R}^k$*−valued,* $k \in \mathbb{N}$, *and* X_t *does not contain a constant term.*

(ii) *For* $t = 1, 2, ...$, *conditional on* X_t, Y_t *has the following conditional density:*

$$m(y \mid X_t; \pi^*, \delta_1^*, \delta_2^*, \beta^*, \gamma^*) \equiv \pi^* f(y \mid X_t; \delta_1^*, \beta^*, \gamma^*) + (1 - \pi^*) f(y \mid X_t; \delta_2^*, \beta^*, \gamma^*)$$

for some $(\pi^*, \delta_1^*, \delta_2^*, \beta^{*\prime}, \gamma^*) \in [0, 1] \times D \times D \times B \times \Gamma$, *where* $D \times D \times B \times \Gamma$ *is a convex compact subset of* $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^+, d \in \mathbb{N}$;

$$f(y \mid X_t; \delta^*, \beta^*, \gamma^*) = \delta^* \gamma^* g(X_t; \beta^*) y^{\gamma^* - 1} \exp(-\delta^* g(X_t; \beta^*) y^{\gamma^*});$$

for each $\beta \in B$, $g(\,\cdot\,; \beta) : \mathbb{R}^k \to \mathbb{R}$ *is a Borel measurable function; and*

$$m(\cdot \mid X_t; \pi^*, \delta_1^*, \delta_2^*, \beta^*, \gamma^*) = p(\cdot \mid X_t, Y_{t-1}, X_{t-1}, ...),$$

where $p(\cdot \mid X_t, Y_{t-1}, X_{t-1}, ...)$ *is the conditional probability density function of* Y_t *given* X_t, Y_{t-1}, $X_{t-1}, ...$.

A2: (i) $g(X_t; \cdot\,)$ *is four times continuously differentiable almost surely.*

(ii) *Suppose* $(\zeta^*, \beta_{o,a}^{*\prime}, \gamma_{o,a}^*) \equiv \arg\max_{\delta,\beta,\gamma \in D \times B \times \Gamma} E[\ln f(Y_t \mid X_t; \delta, \beta, \gamma)]$ *exists and is unique, and that for each* $(\pi, \alpha_1, \alpha_2, \beta, \gamma) \in [0, 1] \times A \times A \times B \times \Gamma$, $E[\ell_t(\pi, \alpha_1, \alpha_2, \beta, \gamma)]$ *exists and is finite*, *where* $A \equiv \{\alpha : \alpha\zeta^* \in D\}$ *and*

$$\ell_t(\pi, \alpha_1, \alpha_2, \beta, \gamma) \equiv \ln[\pi f(Y_t \mid X_t; \alpha_1 \zeta^*, \beta, \gamma) + (1 - \pi) f(Y_t \mid X_t; \alpha_2 \zeta^*, \beta, \gamma)].$$

A3: *There exists a sequence of positive*, *strictly stationary*, *and ergodic random variables* $\{M_t\}$ *such that for some* $\epsilon > 0$,

1. $E[M_t^{1+\epsilon}] < \Delta < \infty$;

2. $\sup_{(\pi,\alpha_1,\alpha_2,\beta,\gamma)} |\nabla_j \ell_t(\pi,\alpha_1,\alpha_2,\beta,\gamma)\nabla_k \ell_t(\pi,\alpha_1,\alpha_2,\beta,\gamma)| \leq M_t$;

3. $\sup_{(\pi,\alpha_1,\alpha_2,\beta,\gamma)} |\nabla_{j,k} \ell_t(\pi,\alpha_1,\alpha_2,\beta,\gamma)| \leq M_t$;

4. $|\nabla_{i_1} f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)|^4 \leq M_t$;

5. $|\nabla_{i_1}\nabla_{i_2} f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)|^2 \leq M_t$;

6. $|\nabla_{i_1}\nabla_{i_2}\nabla_{i_3} f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)|^2 \leq M_t$; *and*

7. $\sup_{(\delta,\beta,\gamma)} |\nabla_{i_1}\nabla_{i_2}\nabla_{i_3}\nabla_{i_4} f(Y_t \mid X_t;\delta,\beta,\gamma)/f(Y_t \mid X_t;\delta,\beta,\gamma)| \leq M_t$,

where $j,k \in \{\pi,\alpha_1,\alpha_2,\beta_1,\cdots,\beta_d,\gamma\}$ *and* $i_1,\cdots,i_4 \in \{\delta,\beta_1,\cdots,\beta_d,\gamma\}$.

For each α and α' in A, we define

$$\mathbf{A}(\alpha,\alpha') \equiv \begin{bmatrix} E[r_t(\alpha)r_t(\alpha')]-1 & E[r_t(\alpha)u_t] & E[r_t(\alpha)s_t'] \\ E[u_t r_t(\alpha')] & E[u_t^2] & E[u_t s_t'] \\ E[s_t r_t(\alpha')] & E[s_t u_t] & E[s_t s_t'] \end{bmatrix},$$

where

$$u_t \equiv \nabla_\delta^2 f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*),$$

$$r_t(\alpha) \equiv f(Y_t \mid X_t;\alpha\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*),$$

and $s_t \equiv \nabla_{(\delta,\beta,\gamma)} f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)/f(Y_t \mid X_t;\delta^*,\beta^*,\gamma^*)$. Also, let

$$\mathbf{B}(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*) \equiv E[\nabla_{(\pi,\alpha_1,\alpha_2,\beta,\gamma)}\ell_t(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)\nabla_{(\pi,\alpha_1,\alpha_2,\beta,\gamma)}\ell_t(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)'].$$

We let $\lambda_{\min}$ and $\lambda_{\max}$ be the minimum and the maximum eigenvalues of a given matrix respectively.

A4: *For every* $(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)$, $\lambda_{\min}\{\mathbf{B}(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)\} \geq 0$ *such that*

(*a*) *if* $\lambda_{\min}\{\mathbf{B}(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)\} > 0$ *then* $\lambda_{\max}\{\mathbf{B}(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)\} < \infty$; *or*

(*b*) *if* $\lambda_{\min}\{\mathbf{B}(\pi^*,\alpha_1^*,\alpha_2^*,\beta^*,\gamma^*)\} = 0$ *then for every* $\epsilon > 0$, $\lambda_{\min}\{\mathbf{A}(\alpha,\alpha)\} > 0$ *and* $\lambda_{\max}\{\mathbf{A}(\alpha,\alpha)\} < \infty$, *uniformly in* $\alpha \in A(\epsilon) \equiv \{\alpha \in A : |\alpha - 1| \geq \epsilon\}$.

Remarks: Assumptions A.3 and A.4 should be interpreted as holding under $\mathcal{H}_o$. Next, although $g(\,\cdot\,;\,\cdot\,)$ is not specified, parts of $\mathbf{A}(\alpha,\alpha')$ can be analytically derived. For example, $E[u_t^2] = 8$ and $E[u_t s_t] = [-2,-2,(2E\{\ln[\delta^* g(X_t;\beta^*)]\} + 2\tilde{\gamma} - 4)/\gamma^*]'$.

6.2 Proofs

To conserve space, the proofs are omitted. They can be found at the following website:

`http://web.yonsei.ac.kr/jinseocho`

References

ABBRING, J. AND VAN DEN BERG, G. (2007): "The Unobserved Heterogeneity Distribution in Duration Analysis," *Biometrika*, 94, 1–13.

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.

BIERENS, H. (1990): "A Consistent Conditional Moment Test of Functional Form," *Econometrica*, 58, 1443–1458.

CARRASCO, M. AND CHEN, X. (2002): "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models," *Econometric Theory*, 18, 17–39.

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.

CHESHER, A. (1984): "Testing for Neglected Heterogeneity," *Econometrica*, 52, 865–872.

CHO, J. AND WHITE, H. (2007a): "Testing for Regime Switching," *Econometrica*, 75, 1671–1720.

CHO, J. AND WHITE, H. (2007b): "Generalized Runs Tests for the IID Hypothesis," UCSD Department of Economics Discussion Paper.

CHO, J. AND WHITE, H. (2008): "Testing for the Mixture Hypothesis with Misspecified Distributions in the Exponential Family," Discussion Paper, Department of Economics, Korea University.

CHO, J., HUANG, M., AND WHITE, H. (2008): "Testing for a Constant Mean Function Using Functional Regression," Discussion Paper, Department of Economics, Korea University.

CHO, J., CHEONG, T., AND WHITE, H. (2010): "Experience with the Weighted Bootstrap in Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models," Discussion Paper, School of Economics, Yonsei University.

COX, D. (1972): "Regression Models and Life Tables," *Journal of the Royal Statistical Society: Series B*, 34, 187–220.

DACUNHA-CASTELLE, D. AND GASSIAT, E. (1999): "Testing the Order of a Model Using Locally Conic Parametrization: Population Mixtures and Stationary ARMA Processes," *Annals of Statistics*, 27, 1178–1209.

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.

EASLEY, D. AND O'HARA, M. (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 905–927.

ENGLE, R. AND RUSSELL, J. (1998): "Autoregressive Conditional Duration: a New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.

ENGLE, R. (2000): "The Econometrics of High Frequency Data," *Econometrica*, 68, 1–22.

FERGUSON, T. (1983): "Bayesian Density Estimation by Mixtures of Normal Distributions," *Recent Advances in Statistics*, edited by M. Rizvi, J. Rustagi, and D. Siegmund. New York: Academic Press. 287–302.

FERNANDES, M. AND GRAMMIG, J. (2006): “A Family of Autoregressive Conditional Duration Models,” *Journal of Econometrics*, 130, 1–23.

GHYSELS, E., GOURIEROUX, C. AND JASIAK, J. (2004): “Stochastic Volatility Duration Models,” *Journal of Econometrics*, 119, 413–433.

HANSEN, B. (1996): “Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis,” *Econometrica*, 413-440.

HARTIGAN, J. (1985): “Failure of Log-Likelihood Ratio Test,” *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*, edited by Lucien M. Le Cam and Richard A. Olshen, 2. Berkeley: University of California Press, 807–810.

HECKMAN, J. AND SINGER, B. (1984): “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models of Duration Data,” *Econometrica*, 52, 271–320.

HONG, Y.-M. AND LIU, J. (2007): ”Generalized Residual-Based Specification Testing for Duration Models with Censoring,” Cornell University Department of Economics Working Paper.

HONORÉ, B. (1990): “Simple Estimation of a Duration Model with Unobserved Heterogeneity,” *Econometrica*, 58, 453–473.

HOROWITZ, J. (1994): “Bootstrap-based Critical Values for the Information Matrix Test,” *Journal of Econometrics*, 61, 395–411.

HOROWITZ, J. (1999): “Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity,” *Econometrica*, 67, 1001–1028.

KALBFLEISCH, J. AND PRENTICE, R. (1980): *The Statistical Analysis of Failure Time Data*. Wiley: New York.

KIEFER, N. (1985): “Specification Diagnostics Based on Laguerre Alternatives for Econometric Models of Duration,” *Journal of Econometrics*, 28, 135–154.

KIEFER, N. (1988): “Economic Duration Data and Hazard Functions,” *Journal of Economic Literature*, 26, 646–679.

LANCASTER, T. (1979): "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47, 939–956.

LANCASTER, T. (1984): "The Covariance Matrix of the Information Matrix Test," *Econometrica*, 52, 1051–1054.

LANCASTER, T. (1985): "Generalized Residuals and Heterogenous Duration Models with Applications to the Weibull Model," *Journal of Econometrics*, 28, 155–169.

LANCASTER, T. (1992): *The Econometric Analysis of Transition Data*. New York: Cambridge University Press.

LIU, X., PASARICA, C. AND SHAO, Y. (2003): "Testing Homogeneity in Gamma Mixture Models," *Scandinavian Journal of Statistics*, 30, 227–239.

MEYER, B. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58, 757–782.

MOSLER, K. AND SEIDEL, W. (2001): "Testing for Homogeneity in an Exponential Mixture Model," *Australian & New Zealand Journal of Statistics*, 43, 231–247.

NICKELL, S. (1979): "Estimating the Probability of Leaving Unemployment," *Econometrica*, 47, 1249–1266.

O'HARA, M. (1995): *Market Microstructure Theory*. Oxford: Basil Blackwell.

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

PRIEGER, J. (2000): "Conditional Moment Tests for Parametric Duration Models," UC Davis Department of Economics Discussion Paper.

PRIEGER, J. (2003): "Bootstrapping the Conditional Moment Test for Parametric Duration Models," *Applied Economics Letters*, 10, 597–600.

RUDIN, W. (1976): *Principles of Mathematical Analysis*. New York: McGraw-Hill.

SHARMA, S. (1987): "Specification Diagnostics for Econometric Models of Durations," UCLA Department of Economics Discussion Paper.

SPIEGEL, M. (1968): *Mathematical Handbook of Formulas and Tables*. New York: McGraw-Hill.

STINCHCOMBE, M. AND WHITE, H. (1998): "Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative," *Econometric Theory*, 14, 295-324.

VAN DEN BERG, G. AND RIDDER, G. (1998): "An Empirical Equilibrium Search Model of the Labor Market," *Econometrica*, 66, 1183–1221.

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.

WHITE, H. (1994): *Estimation, Inference and Specification Analysis*. New York: Cambridge University Press.

YOON, B. (1981): "A Model of Unemployment Duration with Variable Search Intensity," *Review of Economics and Statistics*, 63, 599–609.

Table 1: Critical Values for the LR Test Statistics

Model: $Y_t \mid X_t \sim \pi$Exp$(\delta_1 \exp(X_t\beta)) + (1-\pi)Exp(\delta_2 \exp(X_t\beta)))$

Nominal Level $\setminus A$	$[7/9, 2.0]$	$[7/9, 3.0]$	$[7/9, 4.0]$
1.00 %	6.6987	6.9369	6.9381
2.50 %	4.9000	5.1769	5.2959
5.00 %	3.6341	3.8242	3.9784
7.50 %	2.9129	3.0955	3.2422
10.0 %	2.4033	2.5880	2.7178
Nominal Level $\setminus A$	$[2/3, 2.0]$	$[2/3, 3.0]$	$[2/3, 4.0]$
1.00 %	6.8474	7.0800	7.2518
2.50 %	5.1236	5.3180	5.4671
5.00 %	3.8559	4.0564	4.1675
7.50 %	3.1246	3.2973	3.4215
10.0 %	2.6094	2.7731	2.8934
Nominal Level $\setminus A$	$[5/9, 2.0]$	$[5/9, 3.0]$	$[5/9, 4.0]$
1.00 %	7.2722	7.4972	7.6755
2.50 %	5.4961	5.7758	5.9029
5.00 %	4.2338	4.4468	4.5362
7.50 %	3.4803	3.6934	3.7536
10.0 %	2.9425	3.1359	3.2113

Model: $Y_t \mid X_t \sim \pi$Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)$Weibull$(\delta_2 \exp(X_t\beta), \gamma)$

Nominal Level $\setminus A$	$[7/9, 2.0]$	$[7/9, 3.0]$	$[7/9, 4.0]$
1.00 %	7.0850	7.7021	8.0447
2.50 %	5.3053	5.8698	6.2237
5.00 %	4.0103	4.5422	4.8646
7.50 %	3.2729	3.7725	4.0982
10.0 %	2.7606	3.2382	3.5459
Nominal Level $\setminus A$	$[2/3, 2.0]$	$[2/3, 3.0]$	$[2/3, 4.0]$
1.00 %	7.2559	7.7330	8.0846
2.50 %	5.5245	5.9643	6.2995
5.00 %	4.2281	4.6775	4.9720
7.50 %	3.4693	3.9220	4.2107
10.0 %	2.9284	3.3940	3.6845
Nominal Level $\setminus A$	$[5/9, 2.0]$	$[5/9, 3.0]$	$[5/9, 4.0]$
1.00 %	7.6512	8.1806	8.3686
2.50 %	5.8213	6.3973	6.6212
5.00 %	4.5385	5.0368	5.2406
7.50 %	3.7698	4.2434	4.4599
10.0 %	3.2341	3.6631	3.8993

Notes: The figures provide the critical values for the LR statistic, obtained by applying Theorem 2. These values are obtained by generating the Gauss processes in Theorem 2 100,000 times. A grid search method is used to obtain the maximum of the Gaussian process. The grid distance is 0.01, and m is set to 500.

Table 2: Levels of the Tests (Nominal Level: 5%)
Number of Repetitions: 10,000

Statistics \ Sample Size	50	100	500	1,000	2,000	5,000
DGP: $Y_t \sim$ IID Exp(1)						
Model: $Y_t \sim \pi$Exp$(\delta_1) + (1-\pi)$Exp(δ_2)						
$LR_n([7/9,\ 2])$	1.07	1.74	3.35	3.55	4.00	3.88
$LR_n([7/9,\ 3])$	1.79	2.59	3.46	3.92	3.76	4.11
$LR_n([7/9,\ 4])$	1.77	2.32	3.43	3.62	3.78	4.17
$LR_n([2/3,\ 2])$	1.45	2.17	3.28	3.58	3.73	3.75
$LR_n([2/3,\ 3])$	1.97	2.26	3.46	3.73	4.11	4.00
$LR_n([2/3,\ 4])$	2.14	2.63	3.44	3.48	3.84	4.25
$LR_n([5/9,\ 2])$	1.30	1.89	2.79	2.74	3.48	3.73
$LR_n([5/9,\ 3])$	1.97	2.38	2.97	2.92	2.98	3.28
$LR_n([5/9,\ 4])$	2.12	2.21	2.80	2.98	3.58	3.36
IM_n	18.34	14.79	8.07	6.84	6.58	5.28
$LM_{2,n}$	8.63	8.48	6.96	6.78	5.62	5.44
$LM_{3,n}$	4.46	6.50	15.72	15.53	12.69	10.10[1]
DGP: $Y_t \sim$ IID Weibull$(1, 1)$						
Model: $Y_t \sim \pi$Weibull$(\delta_1, \gamma) + (1-\pi)$Weibull$(\delta_2, \gamma)$						
Statistics \ Sample Size	50	100	500	1,000	2,000	5,000
$LR_n([7/9,\ 2])$	0.00	0.13	1.43	2.48	3.54	4.21
$LR_n([7/9,\ 3])$	0.01	0.17	1.35	2.91	4.02	4.41
$LR_n([7/9,\ 4])$	0.10	0.17	1.57	3.39	3.96	4.02
$LR_n([2/3,\ 2])$	0.13	0.29	2.92	3.46	3.74	4.54
$LR_n([2/3,\ 3])$	0.26	0.53	3.52	3.92	3.91	4.10
$LR_n([2/3,\ 4])$	0.29	0.70	3.43	3.85	4.15	4.13
$LR_n([5/9,\ 2])$	0.43	1.17	2.91	3.26	3.18	3.31
$LR_n([5/9,\ 3])$	0.50	1.36	3.44	3.41	3.89	3.82
$LR_n([5/9,\ 4])$	0.84	1.72	3.48	2.96	4.12	4.04
IM_n	19.19	14.57	8.53	7.33	6.56	5.94
$LM_{2,n}$	20.65	15.37	9.50	7.58	6.18	5.66
$LM_{3,n}$	59.20	46.40	25.27	20.13	15.18	11.72[2]

Table 2: Levels of the Tests (Nominal Level: 5%, Continued)
Number of Repetitions: 10,000

DGP: $Y_t \mid X_t \sim$ IID Exp$(\exp(X_t))$						
Model: $Y_t \mid X_t \sim \pi$Exp$(\delta_1 \exp(X_t\beta)) + (1-\pi)Exp(\delta_2 \exp(X_t\beta))$						
Statistics \ Sample Size	50	100	500	1,000	2,000	5,000
$LR_n([7/9,\ 2])$	0.68	1.62	2.93	3.32	3.83	4.35
$LR_n([7/9,\ 3])$	1.27	2.15	3.10	3.61	3.85	4.17
$LR_n([7/9,\ 4])$	1.35	1.71	2.97	3.44	3.89	4.55
$LR_n([2/3,\ 2])$	1.14	1.84	2.83	3.24	3.69	3.93
$LR_n([2/3,\ 3])$	1.53	2.33	3.33	3.53	3.64	3.74
$LR_n([2/3,\ 4])$	1.62	2.26	3.35	3.41	3.71	3.89
$LR_n([5/9,\ 2])$	1.13	1.66	2.65	2.89	3.19	3.53
$LR_n([5/9,\ 3])$	1.45	1.92	2.31	3.02	3.31	3.32
$LR_n([5/9,\ 4])$	1.44	1.99	2.31	3.20	3.31	3.48
IM_n	23.01	16.86	9.40	7.60	6.52	5.83
$LM_{2,n}$	11.00	10.55	7.21	6.81	5.77	5.22
$LM_{3,n}$	5.29	7.77	17.24	15.04	12.50	9.81[3]
DGP: $Y_t \mid X_t \sim$ IID Weibull$(\exp(X_t), 1)$						
Model: $Y_t \mid X_t \sim \pi$Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)$Weibull$(\delta_2 \exp(X_t\beta), \gamma)$						
Statistics \ Sample Size	50	100	500	1,000	2,000	5,000
$LR_n([7/9,\ 2])$	0.04	0.03	1.25	2.60	3.79	3.80
$LR_n([7/9,\ 3])$	0.03	0.12	1.21	3.03	3.74	4.41
$LR_n([7/9,\ 4])$	0.07	0.17	1.68	2.96	4.06	4.34
$LR_n([2/3,\ 2])$	0.15	0.22	2.72	3.22	4.13	4.05
$LR_n([2/3,\ 3])$	0.23	0.46	2.86	3.88	4.20	4.29
$LR_n([2/3,\ 4])$	0.33	0.76	3.25	3.70	4.18	4.25
$LR_n([5/9,\ 2])$	0.36	1.10	2.76	3.21	3.53	3.39
$LR_n([5/9,\ 3])$	0.56	1.51	3.11	3.55	3.38	3.82
$LR_n([5/9,\ 4])$	0.86	1.57	3.44	3.50	3.78	3.68
IM_n	22.36	17.36	9.30	7.59	6.68	5.76
$LM_{2,n}$	25.05	17.69	9.72	7.93	6.37	5.81
$LM_{3,n}$	64.26	49.28	26.70	21.38	16.28	11.68[4]

Notes: The figures are the empirical rejection rates of the LR, IM_n, $LM_{2,n}$, and $LM_{3,n}$ statistics under the null hypothesis. For the LR statistics, nine parameter spaces are examined for α, specifically $[7/9,\ 2]$, $[7/9,\ 3]$, $[7/9,\ 4]$, $[2/3,\ 2]$, $[2/3,\ 3]$, $[2/3,\ 4]$, $[5/9,\ 2]$, $[5/9,\ 3]$, and $[5/9,\ 4]$, respectively. The LR statistics are indexed by these parameter spaces.

[1]: The empirical level of $LM_{3,n}$ is 5.62 when $n = 100,000$ and repetitions = 5,000.

[2]: The empirical level of $LM_{3,n}$ is 5.88 when $n = 100,000$ and repetitions = 5,000.

[3]: The empirical level of $LM_{3,n}$ is 6.14 when $n = 100,000$ and repetitions = 5,000.

[4]: The empirical level of $LM_{3,n}$ is 5.82 when $n = 100,000$ and repetitions = 5,000.

Table 3: Power of the Tests (Level Distortion Adjusted to 5%)
Number of Repetitions: 2,000
Model: $Y_t \mid X_t \sim \text{Exp}(\delta \exp(X_t \beta))$

$LR_n([7/9,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	32.15	50.30	97.45	100.0	100.0	100.0
Gamma Mixture	17.35	26.40	87.95	98.80	100.0	100.0
Log-normal Mixture	12.90	23.75	82.70	98.80	100.0	100.0
Uniform Mixture I	47.05	73.50	100.0	100.0	100.0	100.0
Uniform Mixture II	6.90	6.85	13.15	16.75	22.50	41.95

$LR_n([7/9,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	33.05	50.20	97.55	100.0	100.0	100.0
Gamma Mixture	17.40	27.30	88.85	98.95	100.0	100.0
Log-normal Mixture	14.10	24.40	82.60	98.70	100.0	100.0
Uniform Mixture I	46.75	72.45	100.0	100.0	100.0	100.0
Uniform Mixture II	7.20	7.15	13.15	16.20	22.20	41.70

$LR_n([7/9,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	32.25	51.10	97.50	100.0	100.0	100.0
Gamma Mixture	17.30	28.85	88.00	98.80	100.0	100.0
Log-normal Mixture	14.45	24.95	81.90	98.50	100.0	100.0
Uniform Mixture I	44.90	72.25	100.0	100.0	100.0	100.0
Uniform Mixture II	7.00	7.75	12.85	15.50	21.35	38.50

$LR_n([2/3,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	30.70	49.45	98.10	100.0	100.0	100.0
Gamma Mixture	28.70	48.40	98.35	100.0	100.0	100.0
Log-normal Mixture	21.00	37.40	92.70	100.0	100.0	100.0
Uniform Mixture I	48.90	74.70	99.95	100.0	100.0	100.0
Uniform Mixture II	6.35	5.60	12.30	16.65	23.95	42.25

$LR_n([2/3,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	29.80	48.95	97.75	100.0	100.0	100.0
Gamma Mixture	27.30	45.30	98.20	100.0	100.0	100.0
Log-normal Mixture	19.05	35.40	91.85	100.0	100.0	100.0
Uniform Mixture I	46.40	73.20	99.95	100.0	100.0	100.0
Uniform Mixture II	6.30	6.10	12.10	15.40	25.00	43.10

$LR_n([2/3,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	29.55	49.85	97.50	100.0	100.0	100.0
Gamma Mixture	26.80	45.70	98.10	100.0	100.0	100.0
Log-normal Mixture	19.15	35.75	91.50	100.0	100.0	100.0
Uniform Mixture I	45.80	73.20	99.95	100.0	100.0	100.0
Uniform Mixture II	6.35	6.70	12.05	15.35	24.95	42.90

Table 3: Power of the Tests (Level Distortion Adjusted to 5%, Continued)
Number of Repetitions: 2,000
Model: $Y_t \mid X_t \sim \text{Exp}(\delta \exp(X_t\beta))$

$LR_n([5/9,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	30.95	49.70	97.05	99.95	100.0	100.0
Gamma Mixture	36.20	56.50	98.90	100.0	100.0	100.0
Log-normal Mixture	28.00	41.75	93.90	99.65	100.0	100.0
Uniform Mixture I	48.70	74.60	99.90	100.0	100.0	100.0
Uniform Mixture II	7.00	5.45	10.90	15.90	21.85	41.75

$LR_n([5/9,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	30.60	49.45	97.05	99.95	100.0	100.0
Gamma Mixture	33.10	54.15	98.35	99.95	100.0	100.0
Log-normal Mixture	24.80	39.55	93.20	99.50	100.0	100.0
Uniform Mixture I	47.95	74.20	99.90	100.0	100.0	100.0
Uniform Mixture II	7.45	5.60	11.40	15.10	21.40	42.70

$LR_n([5/9,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	30.10	48.40	96.95	99.95	100.0	100.0
Gamma Mixture	31.95	52.85	98.10	99.90	100.0	100.0
Log-normal Mixture	24.40	39.05	92.25	99.35	100.0	100.0
Uniform Mixture I	46.60	73.70	99.85	100.0	100.0	100.0
Uniform Mixture II	7.40	5.95	10.80	14.20	21.00	40.50

IM_n						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	0.55	0.70	76.70	99.00	100.0	100.0
Gamma Mixture	0.80	2.25	76.25	98.00	100.0	100.0
Log-normal Mixture	1.35	0.90	55.80	98.00	100.0	100.0
Uniform Mixture I	1.10	5.90	96.05	100.0	100.0	100.0
Uniform Mixture II	4.90	2.95	2.40	4.30	8.50	22.50

LM_{2n}						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	1.65	3.05	77.45	99.25	100.0	100.0
Gamma Mixture	1.65	2.05	65.05	90.20	97.90	99.45
Log-normal Mixture	1.60	1.70	49.90	90.70	99.60	99.95
Uniform Mixture I	3.00	5.60	88.40	99.20	100.0	100.0
Uniform Mixture II	4.45	2.95	3.45	3.35	8.25	24.70

LM_{3n}						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	5.25	12.30	56.40	95.65	100.0	100.0
Gamma Mixture	4.70	10.50	49.85	92.35	99.95	100.0
Log-normal Mixture	4.35	7.10	26.75	70.70	99.95	100.0
Uniform Mixture I	9.45	21.30	86.10	99.95	100.0	100.0
Uniform Mixture II	4.50	3.55	2.75	2.75	2.40	18.55

Notes: The figures are the empirical rejection rates of the LR, IM_n, $LM_{2,n}$, and $LM_{3,n}$ statistics under the five alternative hypothesis: discrete mixture, gamma mixture, log-normal mixture, uniform mixture I, and uniform mixture II. For the LR statistics, nine parameter spaces are examined for α, specifically $[7/9,\ 2]$, $[7/9,\ 3]$, $[7/9,\ 4]$, $[2/3,\ 2]$, $[2/3,\ 3]$, $[2/3,\ 4]$, $[5/9,\ 2]$, $[5/9,\ 3]$, and $[5/9,\ 4]$, respectively. The LR statistics are indexed by these parameter spaces.

Table 4: Power of the Tests (Level Distortion Adjusted to 5%)
Number of Repetitions: 2,000
Model: $Y_t \mid X_t \sim$ Weibull$(\delta \exp(X_t\beta), \gamma)$

$LR_n([7/9,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	16.05	27.30	63.55	85.05	97.90	100.0
Gamma Mixture	5.35	7.10	27.40	63.65	96.75	100.0
Log-normal Mixture	5.40	5.90	17.30	41.70	85.35	100.0
Uniform Mixture I	24.50	44.50	94.30	99.80	100.0	100.0
Uniform Mixture II	11.30	17.60	21.15	18.35	20.75	26.05
$LR_n([7/9,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	16.10	26.45	62.05	82.15	97.35	100.0
Gamma Mixture	3.90	4.25	16.30	47.85	94.45	99.90
Log-normal Mixture	3.80	3.60	10.80	28.10	78.00	99.95
Uniform Mixture I	22.95	35.85	92.25	99.60	100.0	100.0
Uniform Mixture II	13.65	18.90	19.55	18.00	17.40	21.65
$LR_n([7/9,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	16.65	25.30	56.75	80.40	96.75	100.0
Gamma Mixture	3.10	3.50	10.75	41.45	92.50	99.90
Log-normal Mixture	3.80	3.40	6.50	22.50	72.75	99.95
Uniform Mixture I	20.55	31.90	89.25	99.50	100.0	100.0
Uniform Mixture II	15.10	17.15	16.75	15.25	15.20	19.25
$LR_n([2/3,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	17.55	24.65	56.15	83.85	97.65	100.0
Gamma Mixture	8.30	16.40	62.40	93.50	99.85	100.0
Log-normal Mixture	7.25	12.05	41.30	75.60	95.10	100.0
Uniform Mixture I	24.90	44.95	94.30	99.60	100.0	100.0
Uniform Mixture II	7.45	10.55	8.15	10.10	12.20	21.55
$LR_n([2/3,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	17.50	21.65	55.55	80.80	97.35	100.0
Gamma Mixture	5.80	8.90	53.90	90.05	99.70	100.0
Log-normal Mixture	6.25	6.80	33.40	68.05	93.75	100.0
Uniform Mixture I	21.60	36.45	92.25	99.50	100.0	100.0
Uniform Mixture II	8.15	10.40	8.65	8.70	11.70	19.40
$LR_n([2/3,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	15.85	18.85	51.20	79.00	96.80	100.0
Gamma Mixture	4.60	5.75	45.70	88.55	99.65	100.0
Log-normal Mixture	5.05	4.60	27.85	64.75	93.15	100.0
Uniform Mixture I	19.35	31.35	90.70	99.50	100.0	100.0
Uniform Mixture II	9.25	10.00	8.00	8.65	10.25	17.90

Table 4: Power of the Tests (Level Distortion Adjusted to 5%, Continued)
Number of Repetitions: 2,000

$LR_n([5/9,\ 2])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	14.45	20.75	54.20	82.90	97.10	100.0
Gamma Mixture	10.70	16.40	73.55	94.75	99.70	99.95
Log-normal Mixture	7.80	10.70	50.45	77.35	95.60	100.0
Uniform Mixture I	24.65	37.30	93.50	99.70	100.0	100.0
Uniform Mixture II	4.90	6.50	7.70	9.30	12.25	19.50

$LR_n([5/9,\ 3])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	14.90	21.05	52.35	79.70	96.60	100.0
Gamma Mixture	7.50	12.35	64.55	84.95	94.25	98.15
Log-normal Mixture	6.00	8.55	42.55	67.65	89.75	99.30
Uniform Mixture I	22.10	29.50	91.80	99.45	100.0	100.0
Uniform Mixture II	5.75	6.40	7.05	8.30	11.05	16.40

$LR_n([5/9,\ 4])$						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	14.95	19.90	48.70	78.85	96.10	100.0
Gamma Mixture	5.80	10.55	60.65	84.15	94.05	98.15
Log-normal Mixture	5.10	7.75	39.10	66.25	89.00	99.30
Uniform Mixture I	18.30	30.20	90.80	99.30	100.0	100.0
Uniform Mixture II	5.15	6.65	5.45	7.80	9.90	14.05

IM_n						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	2.60	1.95	18.10	49.70	86.95	100.0
Gamma Mixture	3.30	2.25	10.15	81.10	99.20	100.0
Log-normal Mixture	3.10	2.50	19.75	50.30	88.80	100.0
Uniform Mixture I	3.90	5.90	71.35	96.65	100.0	100.0
Uniform Mixture II	4.60	4.30	3.95	3.75	4.65	10.15

LM_{2n}						
Discrete Mixture	1.75	1.65	17.05	51.20	87.80	99.90
Gamma Mixture	3.15	5.60	41.50	81.70	99.35	100.0
Log-normal Mixture	2.95	2.85	17.80	53.25	88.55	100.0
Uniform Mixture I	2.90	5.90	68.50	97.60	100.0	100.0
Uniform Mixture II	4.60	4.25	3.50	3.60	4.95	11.60

LM_{3n}						
DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	2.10	1.95	2.40	12.80	69.00	99.90
Gamma Mixture	2.70	1.00	1.25	21.75	90.60	100.0
Log-normal Mixture	2.00	1.95	0.70	4.55	47.05	99.95
Uniform Mixture I	1.35	0.50	5.75	61.80	99.75	100.0
Uniform Mixture II	4.90	4.05	3.25	2.85	2.25	10.85

Notes: The figures are the empirical rejection rates of the LR, IM_n, $LM_{2,n}$, and $LM_{3,n}$ statistics under the five alternative hypothesis: discrete mixture, gamma mixture, log-normal mixture, uniform mixture I, and uniform mixture II. For the LR statistics, nine parameter spaces are examined for α, specifically $[7/9,\ 2]$, $[7/9,\ 3]$, $[7/9,\ 4]$, $[2/3,\ 2]$, $[2/3,\ 3]$, $[2/3,\ 4]$, $[5/9,\ 2]$, $[5/9,\ 3]$, and $[5/9,\ 4]$, respectively. The LR statistics are indexed by these parameter spaces.

Table 5: Bootstrapped Levels of the LR Tests and Other Tests (Nominal Level: 5%)
Number of Repetitions: 5,000, Number of Bootstrapping: 500

DGP: $Y_t^c := \min[Y_t, C_t]$; $Y_t \mid X_t \sim$ IID Weibull$(\exp(X_t), 1)$; and $C_t \sim$ IID Exp(1)					
Model: $(Y_t^c, D_t) \mid X_t \sim \pi$ C-Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)$ C-Weibull$(\delta_2 \exp(X_t\beta), \gamma)$					
Statistics \ Sample Size	50	100	500	1,000	2,000
$LR_n^c([7/9,\ 2])$	0.00	0.02	0.52	1.50	2.78
$LR_n^c([7/9,\ 3])$	0.06	0.14	1.08	2.00	3.46
$LR_n^c([7/9,\ 4])$	0.04	0.10	0.92	1.98	3.42
$LR_n^c([2/3,\ 2])$	0.00	0.20	1.54	2.86	4.16
$LR_n^c([2/3,\ 3])$	0.16	0.54	2.72	3.36	4.06
$LR_n^c([2/3,\ 4])$	0.22	0.36	2.76	3.54	3.44
$LR_n^c([5/9,\ 2])$	0.20	0.78	3.24	3.82	4.10
$LR_n^c([5/9,\ 3])$	0.30	0.94	3.24	4.00	3.78
$LR_n^c([5/9,\ 4])$	0.40	1.26	3.36	3.52	3.80
IM_n^c	16.33	13.11	8.45	7.05	5.93
$LM_{2,n}^c$	23.77	17.32	9.39	7.70	6.78
$LM_{3,n}^c$	63.21	47.92	26.24	20.40	15.94

Notes: The figures are the empirical rejection rates of the LR statistics under the null hypothesis. For the LR statistics, nine parameter spaces are examined for α, specifically $[7/9,\ 2]$, $[7/9,\ 3]$, $[7/9,\ 4]$, $[2/3,\ 2]$, $[2/3,\ 3]$, $[2/3,\ 4]$, $[5/9,\ 2]$, $[5/9,\ 3]$, and $[5/9,\ 4]$, respectively. The LR statistics are indexed by these parameter spaces. The censored Weibull (C-Weibull) variable has the following joint conditional density function:

$$f^c(y^c, d \mid X_t; \delta^*, \beta^*, \gamma^*) \equiv f(y^c \mid X_t; \delta^*, \beta^*, \gamma^*)^d [1 - F(y^c | X_t; \delta^*, \beta^*, \gamma^*)]^{1-d},$$

where $f(\cdot \mid X_t; \delta^*, \beta^*, \gamma^*)$ and $F(\cdot \mid X_t; \delta^*, \beta^*, \gamma^*)$ are the PDF and CDF of conditional Weibull random variable given X_t respectively. The number of repetitions for the information matrix and the Lagrange multiplier test is 10,000.

Table 6: Bootstrapped Power of the LR Tests and Other Tests (Level: 5%)
Number of Repetitions: 2,000
Model: $(Y_t^c, D_t) \mid X_t \sim \pi\,$C-Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)\,$C-Weibull$(\delta_2 \exp(X_t\beta), \gamma)$

$LR_n^c([7/9,\ 2])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.00	0.45	15.70	39.15	72.25	97.95
C-Gamma Mixture	0.00	0.05	1.25	5.50	25.70	84.85
C-Log-normal Mixture	0.00	0.00	0.95	5.05	18.65	69.15
C-Uniform Mixture I	0.00	0.65	25.80	62.05	92.15	99.95
C-Uniform Mixture II	0.00	0.15	2.20	3.65	6.80	10.00

$LR_n^c([7/9,\ 3])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.30	0.85	18.90	38.80	70.35	97.45
C-Gamma Mixture	0.00	0.10	1.10	4.00	18.20	80.00
C-Log-normal Mixture	0.00	0.15	0.95	4.35	14.85	62.30
C-Uniform Mixture I	0.20	1.30	26.35	58.95	91.30	99.90
C-Uniform Mixture II	0.10	0.30	2.75	4.30	5.60	10.55

$LR_n^c([7/9,\ 4])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.20	1.50	15.55	36.80	67.90	96.85
C-Gamma Mixture	0.15	0.10	0.70	2.75	17.20	73.30
C-Log-normal Mixture	0.05	0.20	1.00	3.15	12.40	56.85
C-Uniform Mixture I	0.35	1.15	21.60	51.20	88.25	99.90
C-Uniform Mixture II	0.20	1.20	3.80	4.70	5.55	10.00

$LR_n^c([2/3,\ 2])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.15	0.90	22.65	45.95	71.95	97.80
C-Gamma Mixture	0.15	0.35	7.65	25.50	56.70	95.50
C-Log-normal Mixture	0.05	0.05	6.65	18.80	41.40	81.50
C-Uniform Mixture I	0.30	1.70	38.55	68.80	92.45	100.0
C-Uniform Mixture II	0.15	0.85	3.75	5.60	6.55	10.45

$LR_n^c([2/3,\ 3])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.40	2.20	23.10	43.00	73.65	97.25
C-Gamma Mixture	0.10	0.30	6.70	21.90	51.20	92.10
C-Log-normal Mixture	0.10	0.45	5.60	16.90	37.30	77.30
C-Uniform Mixture I	0.45	3.15	33.80	64.70	90.65	99.95
C-Uniform Mixture II	0.70	1.15	4.75	5.05	6.15	9.35

$LR_n^c([2/3,\ 4])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.60	2.40	22.35	41.95	70.40	96.15
C-Gamma Mixture	0.10	0.50	5.50	19.95	48.15	91.65
C-Log-normal Mixture	0.25	0.35	5.45	14.45	33.30	79.20
C-Uniform Mixture I	0.75	2.85	32.25	62.35	90.20	99.90
C-Uniform Mixture II	0.55	1.50	4.25	4.80	6.70	9.00

Table 6: Bootstrapped Power of the LR Tests and Other Tests (Level: 5%, Continued)
Number of Repetitions: 2,000

$LR_n^c([5/9,\ 2])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	0.60	3.15	25.65	42.40	70.75	97.75
C-Gamma Mixture	0.30	1.40	18.85	37.40	64.40	94.85
C-Log-normal Mixture	0.15	1.10	13.60	28.55	47.45	81.60
C-Uniform Mixture I	0.85	5.05	42.10	69.35	93.65	99.85
C-Uniform Mixture II	0.25	0.85	4.55	4.60	6.95	9.55

$LR_n^c([5/9,\ 3])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	1.45	4.25	26.35	43.20	84.00	97.10
C-Gamma Mixture	0.55	1.40	16.65	34.80	60.85	93.25
C-Log-normal Mixture	0.70	1.10	13.60	25.15	45.85	79.25
C-Uniform Mixture I	1.80	5.75	39.15	66.00	92.15	100.0
C-Uniform Mixture II	0.80	2.20	4.55	5.60	6.85	9.05

$LR_n^c([5/9,\ 4])$

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
C-Discrete Mixture	1.60	4.30	23.50	41.90	68.40	96.50
C-Gamma Mixture	0.10	0.90	14.95	32.75	56.85	92.90
C-Log-normal Mixture	0.40	1.40	10.60	21.55	43.80	78.05
C-Uniform Mixture I	1.95	5.20	36.60	63.50	89.60	100.0
C-Uniform Mixture II	0.95	2.40	4.00	5.65	5.80	8.55

IM_n^c

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	4.80	4.40	8.45	22.40	56.15	93.45
Gamma Mixture	4.45	3.70	8.30	18.85	46.20	88.85
Log-normal Mixture	5.10	4.05	5.05	11.75	28.90	69.00
Uniform Mixture I	5.10	4.05	19.20	46.80	84.40	99.80
Uniform Mixture II	5.75	4.75	4.60	3.70	4.65	5.80

LM_{2n}^c

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	2.25	1.75	16.70	49.15	84.65	99.90
Gamma Mixture	3.00	4.45	43.60	83.45	99.25	100.0
Log-normal Mixture	2.60	2.65	19.00	48.65	87.30	99.95
Uniform Mixture I	3.50	5.35	68.75	97.05	100.0	100.0
Uniform Mixture II	4.35	3.85	3.00	3.60	5.50	9.65

LM_{3n}^c

DGP \ Sample Size	50	100	500	1,000	2,000	5,000
Discrete Mixture	2.15	2.25	1.65	12.90	60.70	99.85
Gamma Mixture	2.00	1.05	1.35	19.50	87.45	100.0
Log-normal Mixture	2.40	1.50	0.45	4.00	42.55	99.15
Uniform Mixture I	1.35	0.50	4.60	57.00	99.60	100.0
Uniform Mixture II	4.40	3.65	3.05	3.15	2.35	3.40

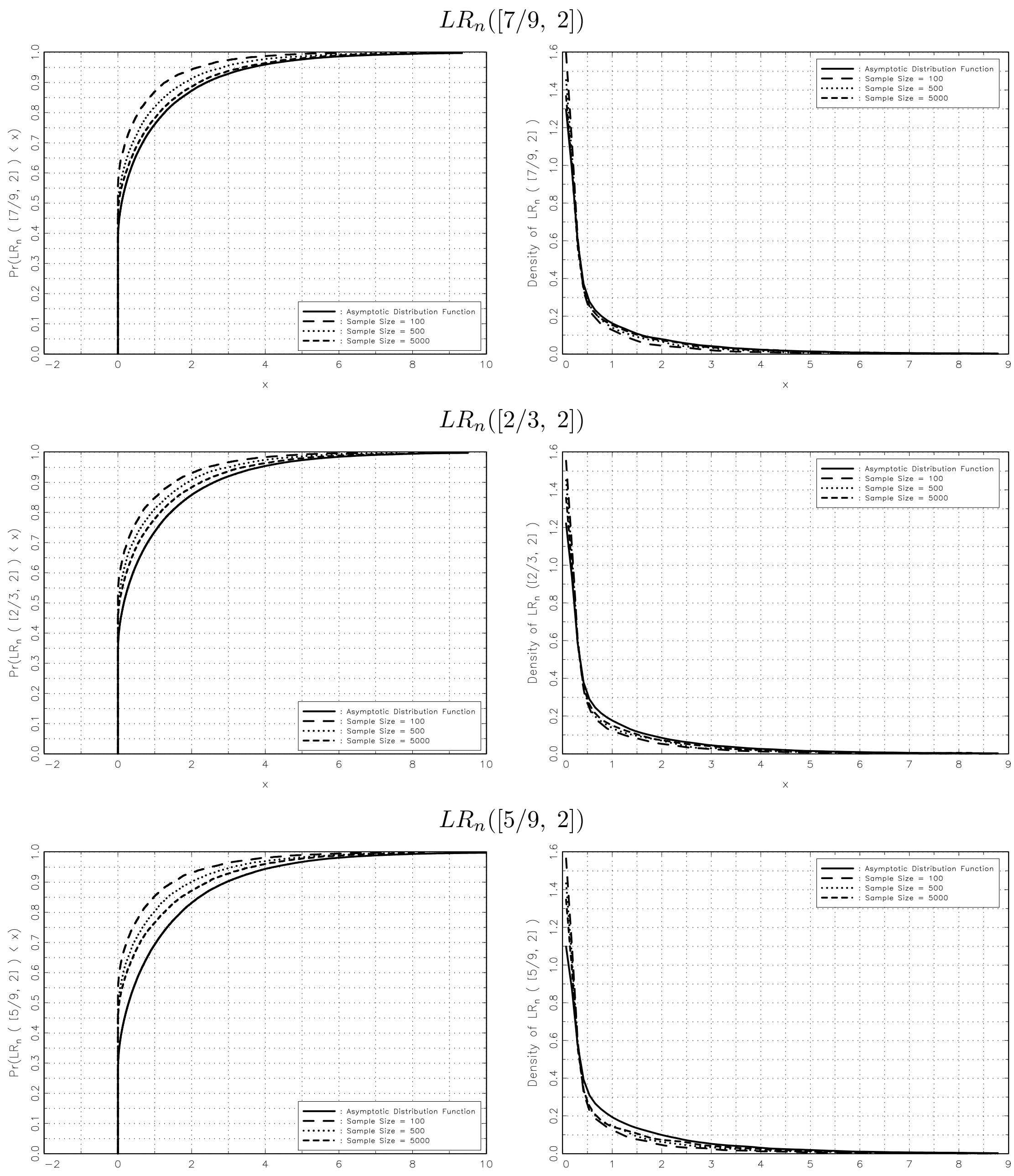
Notes: The figures are the empirical rejection rates of the LR statistics under the alternative hypothesis: censored discrete mixture, censored gamma mixture, censored log-normal mixture, censored uniform mixture I, and censored uniform mixture II. For the LR statistics, nine parameter spaces are examined for α, specifically $[7/9,\ 2]$, $[7/9,\ 3]$, $[7/9,\ 4]$, $[2/3,\ 2]$, $[2/3,\ 3]$, $[2/3,\ 4]$, $[5/9,\ 2]$, $[5/9,\ 3]$, and $[5/9,\ 4]$, respectively. The LR statistics are indexed by these parameter spaces. Information matrix and Lagrange multiplier tests are level distortion adjusted p-values.

Table 7: Duration Data Analysis of Alcatel
Number of Observations: 5,000
Sample period: Duration data in July, 1996

Linear ACD model		
Test Statistics	test values	p–values
$\mathcal{W}_n^{(c)}$	32.7268	0.0000
$\mathcal{W}_n^{(1)}$	15.7176	0.0007
$\mathcal{W}_n^{(2)}$	1.4743	0.2246
$\mathcal{W}_n^{(3)}$	22.2580	0.0000
H-B bound		0.0000
$LR_n([2/3, 2])$	0.4884[1]	0.3764
$LR_n([2/3, 3])$	0.4884[1]	0.3984
$LR_n([2/3, 4])$	0.4884[1]	0.4202
H-B bound		0.4202
Power ACD model		
Test Statistics	test values	p–values
$\mathcal{W}_n^{(c)}$	6.3663	0.0950
$\mathcal{W}_n^{(1)}$	2.2991	0.1294
$\mathcal{W}_n^{(2)}$	4.1223	0.0423
$\mathcal{W}_n^{(3)}$	6.0097	0.0142
H-B bound		0.0568
$LR_n([2/3, 2])$	249.1592[2]	0.0000
$LR_n([2/3, 3])$	256.7723[3]	0.0000
$LR_n([2/3, 4])$	256.7723[3]	0.0000
H-B bound		0.0000
Flexible Power ACD model		
Test Statistics	test values	p–values
$\mathcal{W}_n^{(c)}$	4.5029	0.2120
$\mathcal{W}_n^{(1)}$	1.9705	0.1603
$\mathcal{W}_n^{(2)}$	0.0090	0.9243
$\mathcal{W}_n^{(3)}$	0.5610	0.4538
H-B bound		0.6360
$LR_n([2/3, 2])$	242.1964[2]	0.0000
$LR_n([2/3, 3])$	248.4484[4]	0.0000
$LR_n([2/3, 4])$	248.4484[4]	0.0000
H-B bound		0.0000

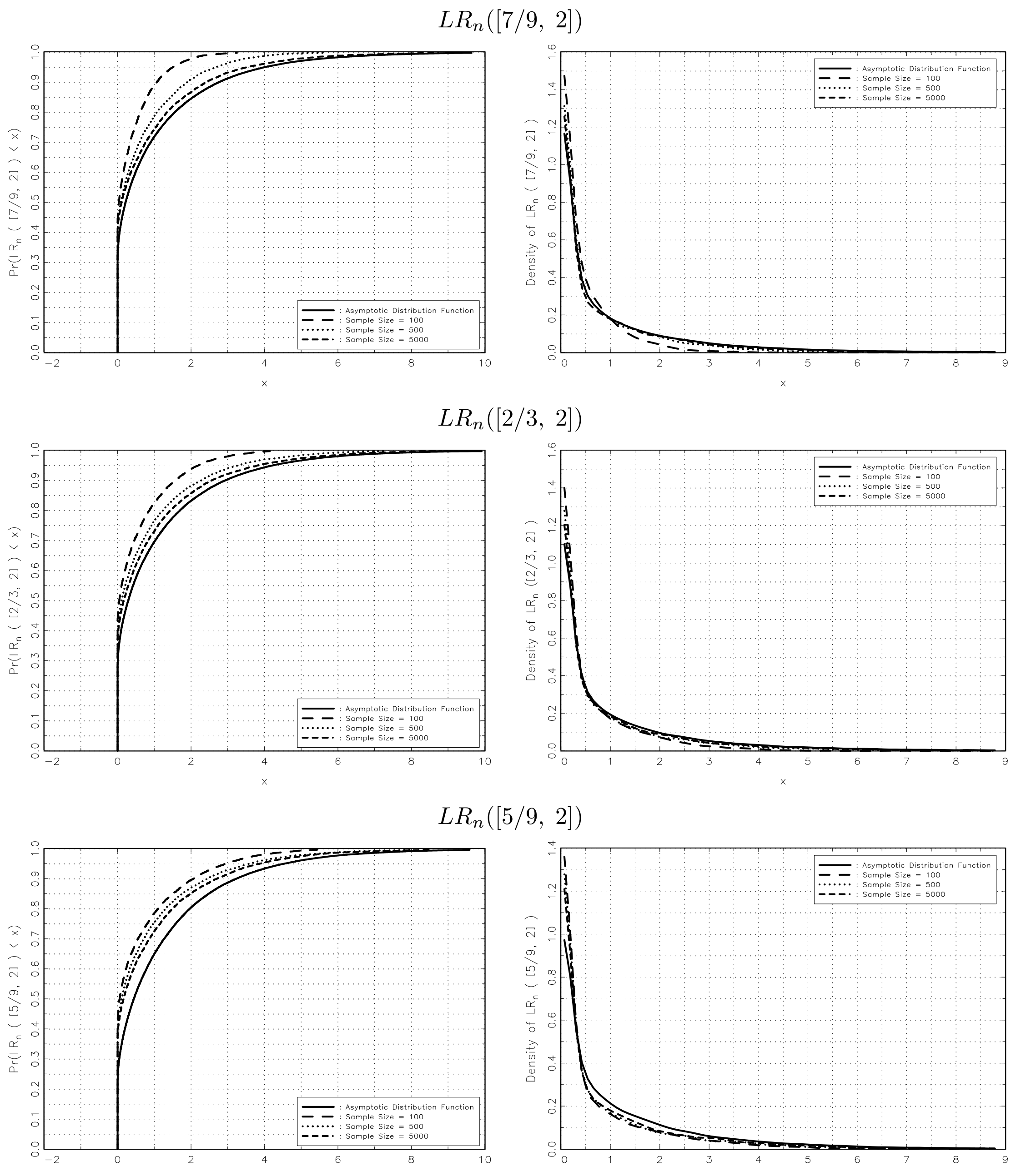
Notes: The figures test correct specification assumption and unobserved heterogeneity. The Wald tests denoted by $\mathcal{W}_n^{(j)}$ $(j = c, 1, 2,$ and $3)$ are the correct specification tests in Cho, Huang, and White (2008). These are applied to the linear ACD, power ACD, and flexible power ACD models. LR statistics test unobserved heterogeneity under various parameter space assumptions.
[1]: $\hat{\alpha}_{1n} = 1.14$ and $\hat{\alpha}_{2n} = 0.96$. This yields the same LR test values for every parameter space.
[2]: $\hat{\alpha}_{1n} = 2.00$ and $\hat{\alpha}_{2n} = 2/3$.
[3]: $\hat{\alpha}_{1n} = 2.41$ and $\hat{\alpha}_{2n} = 2/3$.
[4]: $\hat{\alpha}_{1n} = 2.38$ and $\hat{\alpha}_{2n} = 2/3$.

Figure 1: Empirical Null Distributions and Density Functions of the LR Statistics
Number of Repetitions: 10,000
DGP: $Y_t \mid X_t \sim$ IID Exp$(\exp(X_t))$
Model: $Y_t \mid X_t \sim \pi$Exp$(\delta_1 \exp(X_t\beta)) + (1-\pi)Exp(\delta_2 \exp(X_t\beta))$



Notes: The empirical distributions and density functions of the LR statistics are shown for $A = [7/9, 2]$, $A = [2/3, 2]$, and $A = [5/9, 2]$. The sample sizes are 100, 500, and 5,000. The asymptotic distribution is that given by Theorem 2. For the density functions, the kernel density estimation method based upon the standard normal probability density function is employed.

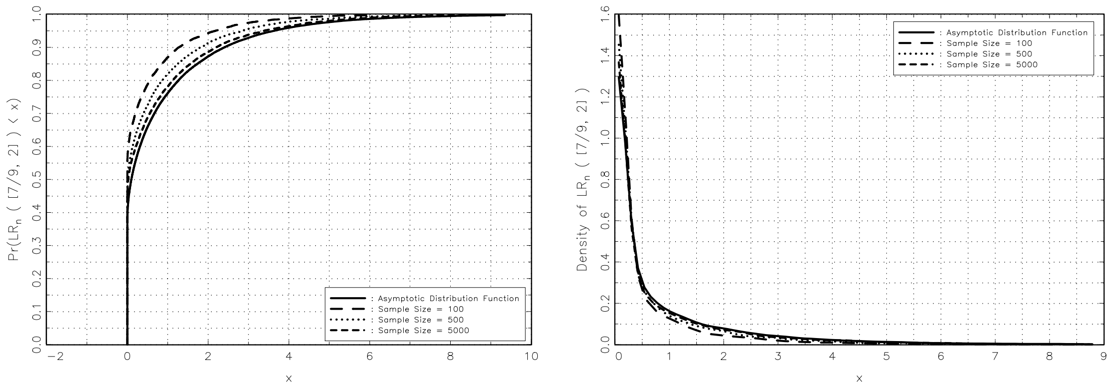
Figure 2: Empirical Null Distributions and Density Functions of the LR Statistics
Number of Repetitions: 10,000
DGP: $Y_t \mid X_t \sim$ IID Weibull$(\exp(X_t), 1)$
Model: $Y_t \mid X_t \sim \pi$Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)$Weibull$(\delta_2 \exp(X_t\beta), \gamma)$



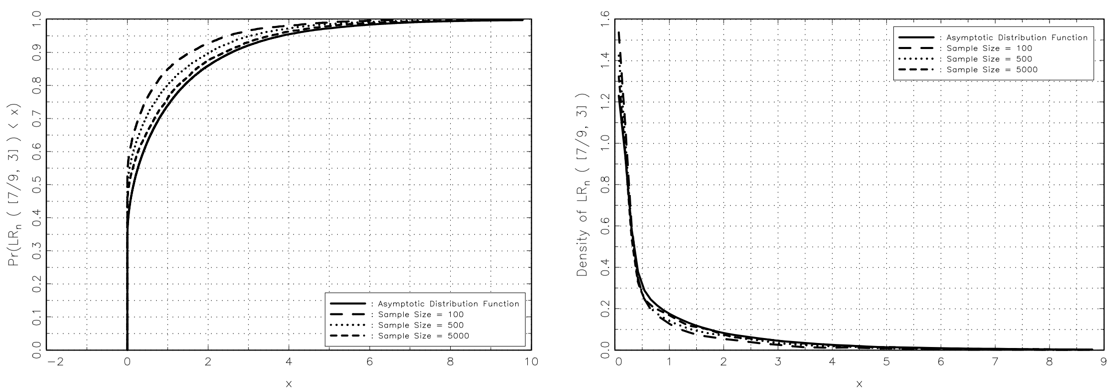
Notes: The empirical distributions and density functions of the LR statistics are shown for $A = [7/9, 2]$, $A = [2/3, 2]$, and $A = [5/9, 2]$. The sample sizes are 100, 500, and 5,000. The asymptotic distribution is that given by Theorem 2. For the density functions, the kernel density estimation method based upon the standard normal probability density function is employed.

Figure 3: Empirical Null Distributions and Density Functions of the LR Statistics
Number of Repetitions: 10,000
DGP: $Y_t \mid X_t \sim$ IID Exp$(\exp(X_t))$
Model: $Y_t \mid X_t \sim \pi$Exp$(\delta_1 \exp(X_t\beta)) + (1-\pi)Exp(\delta_2 \exp(X_t\beta))$

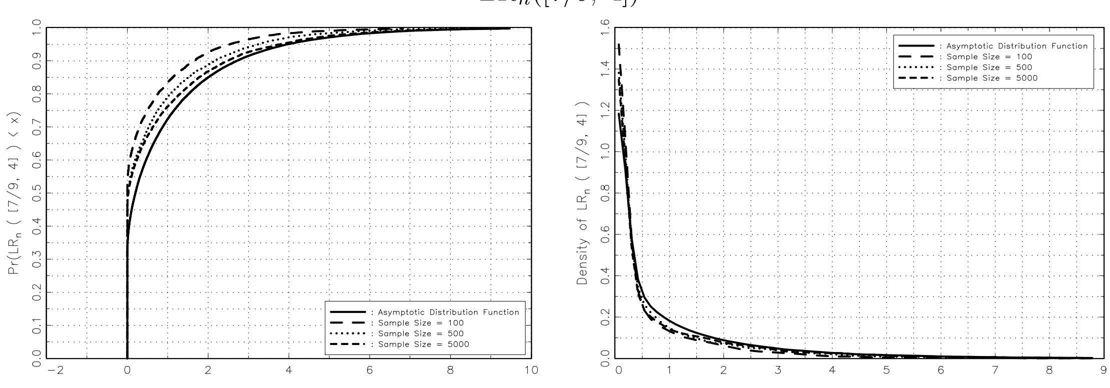
$LR_n([7/9,\ 2])$



$LR_n([7/9,\ 3])$



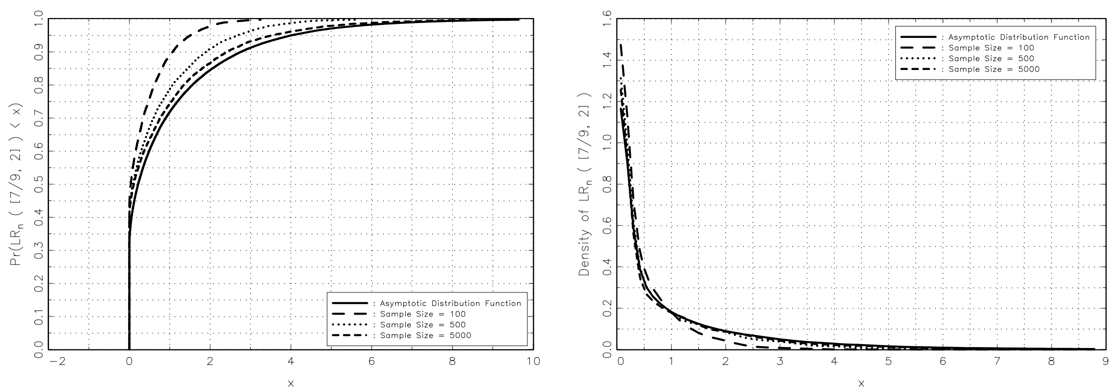
$LR_n([7/9,\ 4])$



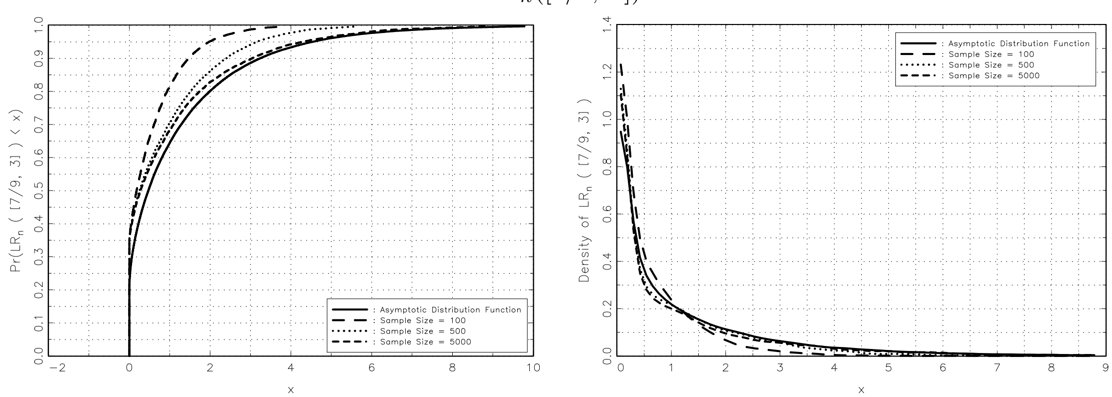
Notes: The empirical distributions and density functions of the LR statistics are shown for $A = [7/9, 2]$, $A = [7/9, 3]$, and $A = [7/9, 4]$. The sample sizes are 100, 500, and 5,000. The asymptotic distribution is that given by Theorem 2. For the density functions, the kernel density estimation method based upon the standard normal probability density function is employed.

Figure 4: Empirical Null Distributions and Density Functions of the LR Statistics
Number of Repetitions: 10,000
DGP: $Y_t \mid X_t \sim$ IID Weibull$(\exp(X_t), 1)$
Model: $Y_t \mid X_t \sim \pi$Weibull$(\delta_1 \exp(X_t\beta), \gamma) + (1-\pi)$Weibull$(\delta_2 \exp(X_t\beta), \gamma)$

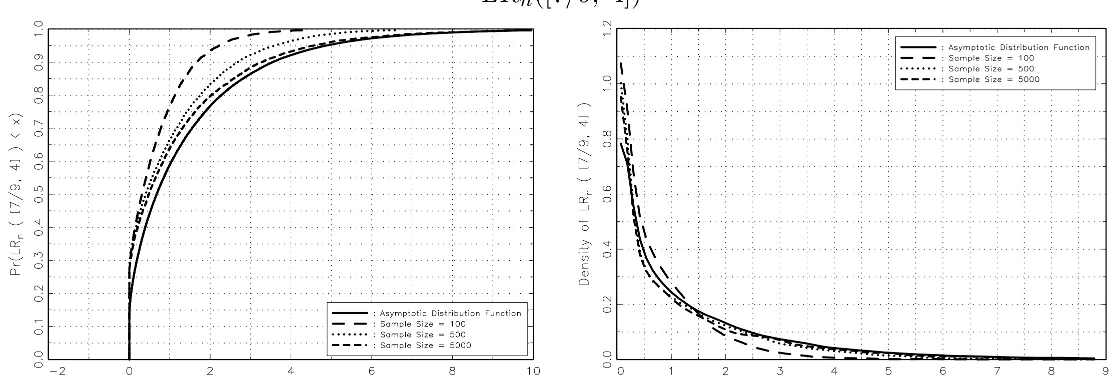
$LR_n([7/9,\ 2])$



$LR_n([7/9,\ 3])$



$LR_n([7/9,\ 4])$



Notes: The empirical distributions and density functions of the LR statistics are shown for $A = [7/9, 2]$, $A = [7/9, 3]$, and $A = [7/9, 4]$. The sample sizes are 100, 500, and 5,000. The asymptotic distribution is that given by Theorem 2. For the density functions, the kernel density estimation method based upon the standard normal probability density function is employed.