

Conditional Moment Tests for Parametric Duration Models

James E. Prieger*
Department of Economics
University of California
One Shields Avenue
Davis, CA 95616-8578
jeprieger@ucdavis.edu

November 7, 2000

Abstract

This paper develops and compares specification tests for parametric duration models estimated with censored data. The tests are based on generalized residuals (the integrated hazard), which is exponentially distributed if the model is correctly specified. I present several conditional moment tests based on the generalized residuals: a raw moments test, a test based on Laguerre polynomials, and a Lagrange multiplier (LM) test. The LM test extends Lancaster's (1985) test by allowing an arbitrarily precise approximation of the likelihood under the alternative. The raw moments test implemented via an auxiliary regression is examined using both asymptotic and bootstrap critical values. Monte Carlo evidence indicates that no one test dominates the others in all situations in terms of size, power, and ease of use. When the data are not censored, the Laguerre test appears to be the best choice. When there is censoring in the data, the Laguerre test is still at least as powerful as the other tests, but the raw moment test may be more convenient to perform. For the convenience of the practitioner the explicit forms of the tests for exponential and Weibull duration models are presented.

*This paper is based on chapter 3 of my dissertation (Prieger, 1999). I thank seminar participants at UC Davis, and especially Colin Cameron, for helpful comments. Earlier versions of the paper benefitted from comments from Alan Marco, Enrico Moretti, Roger Studley, and seminar participants at UC Berkeley. The inevitably remaining mistakes are inevitably mine. PRELIMINARY: PLEASE DO NOT CITE WITHOUT PERMISSION.

Keywords: right censoring, type I censoring, duration analysis, exponential distribution, Weibull distribution, specification test, power curve, bootstrap bias.

1 Introduction

In the linear regression model, correct specification of the distribution of the error term is not necessary for consistent maximum likelihood estimation of regression coefficients as long as the conditional mean is correctly specified as $x_i'\beta$. This robustness to distributional misspecification, however, does not extend to most duration models, especially those involving censoring. Specification testing is therefore important in censored duration models. The leading examples of misspecification in duration models are neglected heterogeneity and duration dependence. One response to the non-robustness of duration models to misspecification has been to develop semi-parametric regression methods. The commonly-used Cox (1972) proportional hazards model makes no assumptions about duration dependence, but assumes there is no heterogeneity not captured by observed covariates.¹ An alternative response to the problem of misspecification is to use parametric models, but to subject them to specification tests.² This paper develops and compares specification tests for parametric duration models estimated with right-censored data.

Many specification tests for duration models have sprung up (Lancaster, 1985; Horowitz and Neumann, 1989; Jaggia, 1991; Sharma, 1992; Jaggia, 1997). A natural building block for specification testing in duration models is the estimated integrated hazard, which can be viewed as a generalized residual. The integrated hazard is unit exponentially distributed if the likelihood is correctly specified, no matter what the duration distribution is. Departures from the assumed duration distribution show up as departures of the generalized residual from the exponential distribution, which is the basis for graphical (Jaggia and Thosar, 1995) and statistical (Lancaster, 1985; Sharma, 1992)

¹See Horowitz (1999) and the citations therein for estimation methods for the proportional hazards model with unobserved heterogeneity.

²Methods such as “semi-nonparametric” series expansion (Gallant and Nychka, 1987) blur the distinction between parametric and semiparametric estimation and lend an arbitrary amount of flexibility to maximum likelihood estimation.

tests based on generalized residuals.³ There are few applications of such tests for censored data, however, a gap this paper attempts to fill.

Testing based on residuals often fits into the framework of the conditional moment test of Newey (1985) and Tauchen (1985). Conditional moment tests have been studied and performed mostly for complete (i.e., uncensored) observations. When some observations are censored in the data, as they often are, the tests must use censored moments (and their expectations) for those observations. Pagan and Vella (1989) provide such a way to incorporate censored observations into conditional moment tests in the context of tobit models.⁴ I present three conditional moment tests modified for censored duration data in the spirit of Pagan and Vella (1989): a raw moments test, a test based on Laguerre polynomials, and a Lagrange multiplier test. The contribution of the paper is to expand the arsenal of available tests for censored duration data, to examine the finite-sample performance of the tests, and to discuss implementation issues for the practitioner.

The conditional moment test is appealing to practitioners for two reasons. First, one need estimate only the single model one wishes to test. Unlike Wald or likelihood ratio tests, one does not need to estimate a more general model that nests the model of interest. Even Lagrange multiplier (LM) tests must be constructed with reference to an alternative hypothesis that requires a more general model, and the derivation of the variance of the statistic can be quite involved. Conditional moment tests require only that one generate some form of residuals from the estimated model and combine them into various statistics for the test. Second, conditional moment tests are appealing because they can be quickly implemented via an auxiliary regression that obviates calculation of the variance matrix for the test statistic.⁵ I show that the auxiliary regression method leads to highly inaccurate inference in small samples, however, which the bootstrap does not remedy when the null hypothesis is false. In particular, although the bootstrap corrects the size problem of the

³Given that generalized residuals can be constructed for any parametric model, not just duration models, the tests presented here can be used for any parametric model. I concentrate on the application to duration models.

⁴In the same paper, Pagan and Vella (1989) discuss tests for duration data, but only for uncensored observations.

⁵As appealing as the conditional moment test is, it is not an omnibus test. There will be some alternatives against which the test has no power (Newey, 1985).

auxiliary regression test, the power curve exhibits severe bias. I develop the Laguerre test as a more accurate and powerful alternative that is not difficult to implement.

The results of the paper indicates that the performance of the Laguerre test at least weakly dominates the others in terms of finite-sample accuracy and power. When the data are not censored, the Laguerre test is as powerful as the LM test that is optimal against the alternative considered, and the Laguerre test as easy or easier to implement as any of the tests. When there is censoring in the data, the Laguerre test is still at least as powerful as the other tests, but the raw moment test may be more convenient to perform. Adding higher moment conditions to the tests does not improve power against heterogeneity when the baseline hazard rate is correctly specified but does improve power against duration dependence in the baseline hazard.

I proceed in the next section by introducing the notation and framework for conditional moment testing of duration models, following Pagan and Vella (1989). In section 3, I present a test based on raw moments of the generalized residuals, which illustrates how the moment conditions are modified for censored data. Jaggia (1991) discusses similar moment-based tests, but only for uncensored data. In section 4, I develop a test based on Laguerre polynomials similar to that of Sharma (1992), but extended for censored data. The Laguerre polynomial test has the computationally convenient property that for some distributions it is asymptotically orthogonal.

Section 5 extends Lancaster's (1985) LM test for unobserved heterogeneity. Jaggia (1997) extends Lancaster's (1985) LM test to include censored data; I further extend it to allow an arbitrarily precise approximation of the likelihood under the alternative of heterogeneity. Higher-order approximation of the likelihood in the LM test turns out to be analogous to adding higher-order moments in moment-based tests. In section 5, I also discuss the kinship between the three tests. The Monte Carlo results in section 6 show that the Laguerre and LM tests generally fare well against the raw moment/auxiliary regression tests. For the convenience of the practitioner, an appendix provides the explicit forms of the tests for exponential and Weibull duration models.

2 The Conditional Moment Approach to Testing

2.1 The Moment Conditions

Let the hazard function of the duration random variable $Y > 0$ be

$$h(y, x, \theta_0) \equiv \lim_{\Delta y \rightarrow 0} \frac{\Pr(y \leq Y < y + \Delta y | Y \geq y, \theta_0, x)}{\Delta y}, \quad (1)$$

the probability of a spell of length y ending the next instant, conditional on lasting at least as long as y , on parameter vector θ_0 , and on vector of ℓ explanatory variables x . The parameter vector θ_0 , with $\dim(\theta_0) = k < \infty$, may comprise ℓ coefficients β_0 and additional $k - \ell$ nuisance parameters (such as the Weibull shape parameter). Define

$$\varepsilon(y, x, \theta_0) \equiv \int_0^y h(t, x, \theta_0) dt, \quad (2)$$

the integrated hazard, to be the generalized error in the sense of Cox and Snell (1968).⁶ The PDF and CDF of the duration process can be stated in terms of the hazard function:

$$f(y|x, \theta_0) = h(y, x, \theta_0) \exp(-\varepsilon(y, x, \theta_0)) \quad (3)$$

$$F(y|x, \theta_0) = 1 - \exp(-\varepsilon(y, x, \theta_0)) \quad (4)$$

(see, e.g., Lancaster, 1985). If h is continuous in y , then h may be replaced with $\partial\varepsilon/\partial y$ in (3) and the PDF may be expressed entirely in terms of ε .

Consider an independent latent sample $\{y_i^*\}$ from Y , $i = 1, \dots, N$. Let the observed sample $\{y_i\}$ be censored, with fixed right censoring points $\{c_i\}$ and censoring indicators $\{d_i\}$ such that $y_i = \min\{y_i^*, c_i\}$ and $d_i = 1\{y_i = c_i\}$, where $1\{\cdot\}$ is the indicator function.⁷ For the rest of the paper, asterisks will denote latent, uncensored quantities, so that $\varepsilon_i^* \equiv \varepsilon(y_i^*, x_i, \theta_0)$ and $\varepsilon_i \equiv \varepsilon(y_i, x_i, \theta_0)$. Then the contribution to the likelihood of a censored observation y_i is $\Pr(Y >$

⁶Note that ε does not have mean zero; for that reason some authors (Pagan and Vella, 1989) prefer to define the generalized error as $\varepsilon - E(\varepsilon)$. I do not follow this convention here. See Gourieroux, Monfort, Renault and Trognon (1987) for another definition of generalized errors.

⁷The tests in this paper also apply if the censoring point C is random but independent of Y (conditional on x).

$y_i|x_i, \theta_0) = 1 - F(y_i|x_i, \theta_0)$, and the log likelihood of the observed sample is

$$l(\theta) = \sum_{i=1}^N (1 - d_i) \log h(y_i, x_i, \theta) - \varepsilon(y_i, x_i, \theta). \quad (5)$$

The conditional moment approach to specification testing exploits the fact that if the model is correctly specified, the sample average moments (evaluated at the estimated parameters and the observed explanatory variables) should be close to the population moment expectations. In a Gaussian model, for example, one might examine the residual vector $e_i \equiv y_i - E(y_i)$ for heteroskedasticity and non-normal kurtosis by looking at the quantities $N^{-1} \sum x_i (e_i^2 - \hat{\sigma}^2)$ and $N^{-1} \sum (e_i^4 - 3\hat{\sigma}^4)$, respectively.

Although one could develop tests for duration models based on residuals $\{e_i\}$, it is convenient to specify moment conditions in terms of ε^* , because ε^* is exponentially distributed for **any** hazard function (Crowley and Hu, 1977). Therefore moment conditions based on ε^* will be the same no matter what the duration distribution is, implying that the conditional moment tests will have general applicability. I assume the moment conditions of interest can be written in terms of ε_i^* or ε_i . Let $m^*: \mathbb{R}_{++} \rightarrow \mathbb{R}^q$ be a vector of conditional moments. Denote $m_i^{0*} \equiv m^*(\varepsilon_i^*)$ and construct m^* so that

$$E(m_i^{0*}|x_i) = \mathbf{0}, \quad i = 1, \dots, N. \quad (6)$$

If the uncensored sample $\{y_i^*\}$ were observed, one could base a specification test on the sample analog of (6). Let $\hat{\theta}$ be the maximum likelihood (ML) estimate of θ_0 , and define $\hat{\varepsilon}_i^* \equiv \varepsilon(y_i^*, x_i, \hat{\theta})$, and $\hat{m}_i^* \equiv m^*(\hat{\varepsilon}_i^*)$.⁸ Then the q dimensional vector

$$\hat{\tau}^* \equiv \frac{1}{N} \sum_{i=1}^N \hat{m}_i^* \quad (7)$$

will be close to zero in the latent sample if the moment restrictions are true in the population.

⁸See Pagan and Vella (1989) for conditional moment testing in the GMM framework.

Censoring complicates matters slightly. Pagan and Vella (1989), in the context of the tobit model, suggest taking the expectation of (6) conditional on the censoring. Letting $w_i = (x_i, c_i, d_i)$ and $m_i^0 \equiv m^0(\varepsilon_i, w_i) \equiv E[m^*(\varepsilon_i^*) | \varepsilon_i, w_i]$, by the law of iterated expectations and (6) we have:

$$E\{m_i^0 | x_i, c_i\} = 0, \quad i = 1, \dots, N. \quad (8)$$

In (8), expectation is taken over (ε_i, d_i) . Thus specification tests for censored samples may be based on the sample analog of (8). Letting $\hat{\varepsilon}_i \equiv \varepsilon(y_i, x_i, \hat{\theta})$ and $\hat{m}_i \equiv m(\hat{\varepsilon}_i, w_i)$, the statistic

$$\hat{\tau} \equiv \frac{1}{N} \sum_{i=1}^N \hat{m}_i \quad (9)$$

will be close to zero in the censored sample if the moment restrictions are true in the population.

The rest of this section presents two asymptotically equivalent test statistics based on (9). The trade-off between the two, as we will see, is one of convenience versus power.

2.2 The Test Statistic

To find the asymptotic distribution of $\hat{\tau}$ under the null hypothesis that the model is correctly specified and that (6) holds, assume that the following probability limits exist:

$$\begin{aligned} M_0 &\equiv \text{plim}_{q \times k} N^{-1} \sum \nabla_{\theta'} m(\varepsilon(y_i, x_i, \theta), w_i) |_{\theta=\theta_0} \\ \mathcal{J}_0 &\equiv - \text{plim}_{k \times k} \frac{1}{N} \nabla_{\theta\theta'} l(\theta) |_{\theta=\theta_0} \\ N^{-\frac{1}{2}} \sum \begin{bmatrix} m_i^{0'} & g_i^{0'} \end{bmatrix} &\xrightarrow{d} \mathcal{N} \left(\begin{matrix} 0 \\ V_0 \end{matrix}, \begin{matrix} 0 & V_0 \\ (q+k) & (q+k) \times (q+k) \end{matrix} \right) \end{aligned} \quad (10)$$

where the subscripts denote the dimensions of the matrices, summations run from 1 to N , and $g_i^0 \equiv g(y_i, w_i, \theta_0)$ is the $(k \times 1)$ derivative of the i th contribution to the log likelihood l : $g_i^0 = \nabla_{\theta} l_i |_{\theta=\theta_0}$.

It will be useful below to partition the variance matrix as:

$$V_0 = \begin{bmatrix} V_{mm'} & V_{mg'} \\ V_{gm'} & V_{gg'} \end{bmatrix} \quad (11)$$

Then it is straightforward to show (Pagan and Vella, 1989) that when $\hat{\theta}$ is the ML estimator we have

$$\sqrt{N}\hat{\tau} \xrightarrow{d} \mathcal{N}(0, \Sigma_0), \quad (12)$$

where

$$\Sigma_0 = V_{mm'} + M_0 \mathcal{J}_0^{-1} V_{gm'} + V_{mg'} \mathcal{J}_0^{-1} M_0' + M_0 \mathcal{J}_0^{-1} V_{gg'} \mathcal{J}_0^{-1} M_0'. \quad (13)$$

By the information equality, for an independent sample $\mathcal{J}_0 = V_{gg'}$. Furthermore, by the generalized information equality (Tauchen, 1985), $E(\nabla_{\theta'} m_i^0) = -E(m_i^0 g_i^{0'})$, and it follows that $M_0 = -V_{mg'}$. These results combine to simplify the expression for the variance to two equivalent forms:

$$\Sigma_0 = V_{mm'} - V_{mg'} V_{gg'}^{-1} V_{gm'} \quad (14)$$

$$= V_{mm'} - M_0 \mathcal{J}_0^{-1} M_0' \quad (15)$$

A convenient form for the test is

$$N \hat{\tau}' \Sigma_0^{-1} \hat{\tau} \xrightarrow{d} \chi^2(q), \quad (16)$$

which requires an estimate of Σ_0 to be feasible.

Under independent sampling, the terms on the right-hand side of (14) will be the following expectations:

$$V_0 \equiv E \left(E \left\{ \begin{bmatrix} m_i^{0'} & g_i^{0'} \end{bmatrix}' \begin{bmatrix} m_i^{0'} & g_i^{0'} \end{bmatrix} \middle| x_i \right\} \right). \quad (17)$$

Therefore V_0 can be consistently estimated with \hat{V} , the usual sample average analog of (17) evaluated at the estimated parameter vector $\hat{\theta}$. Define $\hat{\Sigma}$ to be the estimate of Σ_0 based on \hat{V} and (14). The test statistic (16) using $\hat{\Sigma}$ may be written as

$$N^2 \hat{\tau}' \left\{ S' \left[I - G (G' G)^{-1} G' \right] S \right\}^{-1} \hat{\tau} \xrightarrow{d} \chi^2(q), \quad (18)$$

where G is the $N \times k$ matrix with $\hat{g}_i' = g(y_i, w_i, \hat{\theta})$ as the i th row, and S is the $N \times q$ matrix with \hat{m}_i' as the i th row. To interpret (18), note that the term in the middle has the form of the “residual

maker” for a linear regression. That is, the expression in braces is composed of the squared, summed residuals from regressing \hat{m}_i on \hat{g}_i . Tauchen (1985) shows that one can therefore implement the test via an auxiliary regression: regress \hat{m}_i on \hat{g}_i and a constant (this will be a seemingly unrelated regression [SUR] if $q > 1$) and test the constant(s) for significance. The resulting test statistic is $N^2 \hat{\tau}' \left\{ S' \left[I - G(G'G)^{-1} G' \right] S - \hat{\tau} \hat{\tau}' \right\}^{-1} \hat{\tau}$, which differs from (18) only by the presence of $\hat{\tau} \hat{\tau}'$ in the braced expression, which converges in probability to 0 under the null hypothesis. Thus testing via the auxiliary regression is asymptotically equivalent to testing based (16) using $\hat{\Sigma}$, and the two are nearly equivalent in finite samples. The auxiliary regression method is convenient, in that it can be implemented with any regression software.

While the auxiliary regression method is asymptotically equivalent to (16) and is easy to implement, it has notorious slow convergence to its limiting distribution (Pagan and Vella, 1989). In the simulations performed in Section 6, I find that the actual small sample size of the raw moment test may be over six times its nominal level. As an alternative to the auxiliary regression, when Y follows the exponential distribution one can easily calculate the inner expectation in (17) analytically, and then average over the x_i in the sample to approximate the outer expectation. This analytical estimate generally converges to its probability limit faster than \hat{V} in practice, and performs better than \hat{V} in Monte Carlo power studies (Jaggia, 1997). The analytical estimate is

$$\check{V} = N^{-1} \sum E \left\{ \left[\begin{array}{cc} m_i^{0'} & g_i^{0'} \end{array} \right]' \left[\begin{array}{cc} m_i^{0'} & g_i^{0'} \end{array} \right] \middle| x_i \right\} \bigg|_{\theta=\hat{\theta}} \quad (19)$$

and is used in the Monte Carlo exercise below for the Laguerre and LM tests. Define $\check{\Sigma}$ to be the estimate of Σ_0 based on \check{V} . Although $\check{V}_{mm'}$, the upper left partition of (19), can be calculated for any distribution for the moment tests considered in this paper, $\check{V}_{mg'}$ and $\check{V}_{gg'}$ may not be available for models other than the exponential.

A final possible estimate of V_0 can be based on the sample average analogs of $V_{mm'}$, M_0 , and \mathcal{J}_0 in (15) when $\check{V}_{mg'}$ and $\check{V}_{gg'}$ are not available. This estimate, $\tilde{\Sigma}$, replaces the appropriate elements of (15) with $\hat{V}_{mm'} \equiv N^{-1} \sum \hat{m}_i \hat{m}_i'$, $\hat{M} \equiv N^{-1} \sum \nabla_{\theta'} \hat{m}_i$, and $\hat{\mathcal{J}} \equiv N^{-1} \sum \nabla_{\theta'} \hat{g}_i$. Given the poor

performance of the auxiliary regression method, testing based on (16) using $\tilde{\Sigma}$ may be preferred for models for which $\tilde{\Sigma}$ cannot be easily calculated (e.g., the Weibull model).

Section A.1 in the appendix contains the explicit form of $\tilde{\Sigma}$ for the exponential model and of $\tilde{\Sigma}$ for the Weibull model when the second through fourth moment conditions are used in the three versions of the test presented in the next three sections. Section A.1 also presents the gradient for these models needed to implement the auxiliary regression.

3 Tests Based on Raw Moments

Which moments should one use for testing? Given a fully specified distribution under the alternative hypothesis, an LM test defines the optimal set of conditional moments (Newey, 1985). In practice, one often chooses a test based not only on its asymptotic power but on its ease of implementation.⁹ The practitioner may choose among the infinite number of moments satisfying (6). In this section and the next two, I explore three alternative sets of moments based on the generalized residuals.

Most tests using generalized residuals in the literature are based on raw moments. Here I present the raw moment conditions for censored samples. Because ε^* is distributed unit exponential, the q population raw moment conditions corresponding to (6) are:

$$m_i^{0*} = \begin{bmatrix} \varepsilon_i^{*2} - 2! \\ \vdots \\ \varepsilon_i^{*q+1} - (q+1)! \end{bmatrix}. \quad (20)$$

If there is no censoring in the sample, test statistics can be based directly on the sample analogs of the theoretical moments. Let the element of $\hat{\tau}^*$ in (7) corresponding to the p th power of $\hat{\varepsilon}_i^*$ be:

$$r_p^* = \frac{1}{N} \sum_{i=1}^N (\hat{\varepsilon}_i^{*p} - p!). \quad (21)$$

⁹It has been humorously noted that the actual power of a test is the theoretical power multiplied by the probability that the test is actually used.

In the censored sample, instead of (20) one calculates m_i^0 , the expectation of m_i^{0*} conditional on the censoring. To do this, note that when the error is not censored, we have

$$E(\varepsilon_i^{*p} | \varepsilon_i, d_i = 0) = \varepsilon_i^p, \quad (22)$$

and that when the error is censored, we have

$$E(\varepsilon_i^{*p} | \varepsilon_i, d_i = 1) = \sum_{j=0}^p \frac{p!}{j!} \varepsilon_i^j; \quad (23)$$

see section A.2.2 from the appendix.

From (22) and (23), we find the appropriate element of $\hat{\tau}$ in (9) (the counterpart to r_p^* for censored samples) to be:

$$r_p = \frac{1}{N} \sum_{i=1}^N \left(\varepsilon_i^p - p! + d_i \sum_{j=0}^{p-1} \frac{p!}{j!} \varepsilon_i^j \right) \quad (24)$$

Note that r_p reduces to the usual raw moment conditions for uncensored observations when $d_i = 0$ for all i . Typically one takes $\hat{\tau} = (r_2, \dots, r_{q+1})'$; one cannot test r_1 (refer to discussion of equation (49) below). Of particular interest are the first few moments:

$$r_2 = N^{-1} \sum \hat{\varepsilon}_i^2 - 2 + 2d_i(\hat{\varepsilon}_i + 1) \quad (25)$$

$$r_3 = N^{-1} \sum \hat{\varepsilon}_i^3 - 6 + 3d_i(\hat{\varepsilon}_i^2 + 2\hat{\varepsilon}_i + 2) \quad (26)$$

$$r_4 = N^{-1} \sum \hat{\varepsilon}_i^4 - 24 + 4d_i(\hat{\varepsilon}_i^3 + 3\hat{\varepsilon}_i^2 + 6\hat{\varepsilon}_i + 6) \quad (27)$$

where all summations are over 1 to N . Many practitioners use raw moments and the auxiliary regression form of the test to avoid the matrix calculation of the variance matrix.

4 Tests Based on Laguerre Polynomials

This section develops an alternative to raw moment tests, in which the moments are chosen to be orthonormal polynomials in the generalized residual. Such tests, under certain circumstances, are particularly easy to implement, requiring no matrix computation of \hat{V} . Furthermore, the Monte

Carlo exercises in section 6 show that the Laguerre tests avoid the slow asymptotic convergence of the auxiliary regression method.

Let $f(Z|x, \theta)$ denote the conditional density of a random variable Z . Assume that the moments $\omega_p = E(Z^p|x, \theta)$ exist and are finite for all $p \in \mathbb{N}$. A family of polynomials $\{P_p(Z, x, \theta)\}_{p=0}^{\infty}$, where p is the order of the polynomial, is said to be orthonormal with respect to density f if

$$E(P_n(Z, x, \theta)P_p(Z, x, \theta)|x, \theta) = 1\{n = p\}. \quad (28)$$

When the even moments dominate the others, orthonormal polynomial families always exist and are unique.¹⁰ The orthonormal polynomial family for the (uncensored) exponential distribution is the family of Laguerre polynomials. The p th Laguerre polynomial in Z is

$$L_p(Z) = \sum_{j=0}^p (-1)^j \frac{p!}{(j!)^2 (p-j)!} Z^j, \quad (29)$$

where the usual convention $0! = 1$ holds.¹¹ Because L_0 is 1, it follows from (28) that $E(L_p) = 0$. Note that because the generalized error is exponentially distributed for any duration distribution, the Laguerre polynomials in ε are orthonormal for any data generating process.

Sample moments for censored observations based on Laguerre polynomials take the form

$$\lambda_p = \frac{1}{N} \sum_{i=1}^N (1 - d_i) L_p(\hat{\varepsilon}_i) + d_i \tilde{L}_p(\hat{\varepsilon}_i) \quad (30)$$

for the appropriate element of $\hat{\tau}$ in (9), where L_p is as in (29), and

$$\tilde{L}_p(y) = \sum_{j=0}^p (-1)^j \binom{p}{j} \sum_{m=0}^j \frac{y^m}{m!}$$

is the expectation of the Laguerre polynomial when the duration is censored. The first few moment

¹⁰In particular, let Ω be the matrix with ij element ω_{i+j-2} . Then if Ω is positive definite, there exists a unique orthonormal polynomial family with respect to f (Cramér, 1946). The converse also holds.

¹¹One can show by direct calculation that (29) satisfies the recurrence relation defining the Laguerre polynomials: $(p+1)L_{p+1}(Z) = (2p+1-Z)L_p(Z) - pL_{p-1}(Z)$ (Abramowitz and Stegun, 1964, p.782).

conditions from Laguerre polynomials are

$$\lambda_2 = N^{-1} \sum \frac{1}{2} [\hat{\varepsilon}_i^2 - 4\hat{\varepsilon}_i + 2 + 2d_i (\hat{\varepsilon}_i - 1)] \quad (31)$$

$$\lambda_3 = N^{-1} \sum \frac{1}{6} [-\hat{\varepsilon}_i^3 + 9\hat{\varepsilon}_i^2 - 18\hat{\varepsilon}_i + 6 + 3d_i (-\hat{\varepsilon}_i^2 + 4\hat{\varepsilon}_i - 2)] \quad (32)$$

$$\lambda_4 = N^{-1} \sum \frac{1}{24} [\hat{\varepsilon}_i^4 - 16\hat{\varepsilon}_i^3 + 72\hat{\varepsilon}_i^2 - 96\hat{\varepsilon}_i + 24 + 4d_i (\hat{\varepsilon}_i^3 - 9\hat{\varepsilon}_i^2 + 18\hat{\varepsilon}_i - 6)] \quad (33)$$

When there is no censoring, $d_i = 0$ for all i and λ_p reduces to $\lambda_p^* = N^{-1} \sum L_p(\hat{\varepsilon}_i^*)$.

Sharma (1992), building on work by Kiefer (1985), showed (for uncensored data) that conditional moment tests based on λ_p may be derived as an LM test for unexplained duration dependence. In that case the nesting model is an expansion of (3) by means of Laguerre polynomials in ε_i , and the restricted model with no unexplained duration dependence is (3).

Although the Laguerre polynomials (29) are orthonormal with respect to the uncensored exponential distribution, the modified Laguerre polynomials (30) are not orthonormal—or even orthogonal—with respect to the censored exponential distribution. Although one can construct orthonormal polynomials for the censored exponential distribution by the Gram-Schmidt method, the coefficients of the resulting polynomials are tedious to compute. Given that the advantage of the orthonormal polynomials—their ease of computation (explained below)—is lost in the censored case, I do not present the orthonormal polynomials for the censored distribution.

In some cases orthonormal polynomials lead to test statistics that are particularly easy to compute. The test statistic $\hat{\tau}$ is asymptotically orthogonal if the asymptotic variance matrix of $\hat{\tau}$ is diagonal. Asymptotically orthogonal tests are desirable because the variance matrix is easy to compute, sequential tests are invariant to the order in which they are performed, and joint test statistics are the sum of individual test statistics. For an example of the latter, the test statistic for the joint test that $(\lambda_2, \lambda_3, \lambda_4) = 0$ would be the sum of the test statistics for each individual test of $\lambda_i = 0$, $i = 2, 3, 4$.

Laguerre-based tests are orthogonal only under certain conditions. Tests using (30) are not orthogonal when some data are censored, because then $\{\lambda_p\}$ is not an orthogonal set of polynomials.

Even for uncensored data, orthonormal polynomials do not necessarily lead to orthogonal tests. Orthonormality of the moments ensures that $V_{mm'}$ is diagonal, but the second term in (15) may not be diagonal. The variance matrix Σ_0 is diagonal, and hence the test is asymptotically orthogonal, if $M_0 \equiv \text{plim } N^{-1} \sum \nabla_{\theta'} m_i^0 = 0$. The Laguerre polynomial moments have this property for the exponential model, but they do not for many other commonly used duration models. Although Sharma (1992) conjectures that asymptotic orthogonality is unlikely to extend out of the exponential case, the following proposition (proved in the appendix) shows that asymptotic orthogonality can be extended to some other distributions.

Proposition 1 For any uncensored duration random variable y_i with generalized error of the form $\varepsilon_i = \kappa_i^\alpha \Lambda(y_i)$, with $\kappa_i = \exp(-\beta' x_i)$, $\alpha \neq 0$ a known constant, and Λ a known function, under independent sampling we have $M_0 = \text{plim } N^{-1} \sum \nabla_{\theta'} m_i^0 = 0$, where m_i^0 are the Laguerre moment conditions.

For such processes with no censoring, then, (14) implies that the asymptotic variance is merely $V_{mm'}$. The orthonormality of the Laguerre polynomials furthermore means that $V_{mm'} = I$; i.e. the Laguerre tests are asymptotically orthogonal. Thus calculating the joint test statistic in the uncensored case requires no variance computations or matrix inversions, and in the censored case only $V_{mm'}$ need be calculated and inverted. The class of processes with $\varepsilon_i = \kappa_i^\alpha \Lambda(y_i)$ is a subset of the proportional hazards class; the baseline hazard Λ must be known and contain no elements of θ . The most important member of this class is the exponential duration model. For the exponential model, $\Lambda(y_i) = y_i$. The class also includes the Rayleigh distribution, commonly used in life testing of electronic components. The class does not contain the lognormal or log-logistic models, and includes the Weibull model only if the shape parameter is known. Table 1 characterizes these models.

(Table 1 about here)

5 An LM Test for Unobserved Heterogeneity

A third set of moments may be derived from an LM test for unobserved heterogeneity. This section extends Lancaster's (1985) LM test for neglected heterogeneity in the hazard rate in two directions. Lancaster's (1985) test is for uncensored samples, and is a true LM test only up to a second order approximation of the likelihood. I first modify the test to include censored data, as did Jaggia (1997). I then use higher-order approximations of the likelihood function in the construction of the test statistic, which leads to a test with higher power.

Let the hazard function of the duration process, (1), take the form $h(y|v, x, \theta_0) = vb(y, x, \theta_0)$, $v > 0$. Here v is a multiplicative heterogeneity term satisfying $E(v) = 1$, and b is a baseline hazard rate. A leading example of such a hazard function has $b(y, x, \theta_0) = e^{-x'\beta_0}$ and heterogeneity parameterized as $v = \exp(-u)$, so that the hazard is $h(y) = \exp(-[x'\beta_0 + u])$. This is the form used in the Monte Carlo exercises, but results here will be developed for the general form here.

Define ε_b to be the integrated hazard from (2) when $v = 1$. In terms of ε_b the conditional survival function $S(y)$ —the fraction of durations lasting longer than y —is

$$S(y|v, x, \theta_0) \equiv 1 - F(y|v, x, \theta_0) = \exp(-v\varepsilon_b), \quad (34)$$

where the dependence of ε_b on (y, x, θ_0) is suppressed in the notation. Equation (34) follows directly from (4). The unconditional (on v) survival function is

$$S(y|x, \theta_0) \equiv 1 - F(y|x, \theta_0) = E_v[\exp(-v\varepsilon_b)]. \quad (35)$$

Finally, let $E(v - 1)^p = \mu_p$, the p th central moment. Recall $\mu_1 = 0$; fixing the mean allows identification of the intercept in β_0 when $b = e^{-x'\beta_0}$.

5.1 Approximating the Unrestricted Likelihood

To explore heterogeneity in the duration model, approximate (34) with \tilde{S}_{q+1} , an expansion of order $q+1$ ($q \geq 1$) of S as a function of v about $v = 1$:

$$\tilde{S}_{q+1}(y|v, x, \theta_0) = S(y|v = 1, x, \theta_0) + \sum_{p=1}^{q+1} \frac{d^p}{dv^p} S(y|v, x, \theta_0) \Big|_{v=1} \frac{(v-1)^p}{p!} \quad (36)$$

$$= e^{-\varepsilon_b} \left[1 + \sum_{p=1}^{q+1} a_p(\varepsilon_b) (v-1)^p \right], \quad (37)$$

where (37) follows from (34) and $a_p(\varepsilon) \equiv (-\varepsilon)^p / p!$. Taking expectations with respect to v , we find:

$$\tilde{S}_{q+1}(y|x, \theta_0) = S(y|v = 1, x, \theta_0) \left[1 + \sum_{p=2}^{q+1} a_p(\varepsilon_b) \mu_p \right]. \quad (38)$$

Given that the PDF f equals $-S'$, it follows that f may be approximated by $-\tilde{S}'_{q+1}$:

$$\tilde{f}_{q+1}(y|x, \theta_0) = f(y|v = 1, x, \theta_0) \left\{ 1 + \sum_{p=2}^{q+1} [a_p(\varepsilon_b) + a_{p-1}(\varepsilon_b)] \mu_p \right\} \quad (39)$$

Equation (39) follows from (38) and the fact that $\varepsilon'_b(y) \equiv b(y)$. Therefore f is approximated by (39) when μ_2 and the higher moments (i.e., the heterogeneity) are small.

Now the approximate likelihood of a sample including censored and uncensored observations follows directly. As before, let the indicator variable d_i be 1 if duration y_i is censored and 0 otherwise. Then (5), (38), and (39) lead to $l(\theta, \mu_2, \dots, \mu_{q+1})$, an approximation of the true log likelihood of the sample:

$$\begin{aligned} l(\theta, \mu_2, \dots, \mu_{q+1}) &= \sum_{i=1}^N (1 - d_i) \left\{ \log b_i - \varepsilon_{bi} + \log \left(1 + \sum_{p=2}^{q+1} [a_p(\varepsilon_{bi}) + a_{p-1}(\varepsilon_{bi})] \mu_p \right) \right\} \\ &\quad + d_i \left\{ \log \left[1 + \sum_{p=2}^{q+1} a_p(\varepsilon_{bi}) \mu_p \right] - \varepsilon_{bi} \right\}, \end{aligned} \quad (40)$$

where $b_i \equiv b(y_i, x_i, \theta)$ and $\varepsilon_{bi} \equiv \varepsilon_b(y_i, x_i, \theta)$.

5.2 The LM Test Statistic

Recall the idea of the LM test is that the score (expected first derivatives of the log likelihood function) of the unrestricted model will be close to zero when evaluated at the restricted estimates, if indeed the restrictions are true. A general form of the test is

$$\left[\nabla_{\theta'} \log L_U(\hat{\theta}_R) \right] \left[\mathcal{I}_U(\hat{\theta}_R) \right]^{-1} \left[\nabla_{\theta} \log L_U(\hat{\theta}_R) \right] \xrightarrow{d} \chi^2(q), \quad (41)$$

where L_U is the likelihood of the unrestricted model, \mathcal{I}_U is the information matrix from the unrestricted likelihood, $\hat{\theta}_R$ is the vector of estimates from the restricted model, and q is the number of restrictions. In the present context, our parameter vector of interest is $(\theta, \mu_2, \dots, \mu_{q+1}) \equiv (\theta, \mu)$. The restriction we wish to test is $\mu = 0$.

The average gradient of (40) may be found from

$$N^{-1} \nabla_{\theta} l(\theta, \mu) = N^{-1} \sum_{i=1}^N \left[(1 - d_i) \left(\frac{\nabla_{\theta} b_i}{b_i} - \nabla_{\theta} \varepsilon_{bi} \right) + d_i (-\nabla_{\theta} \varepsilon_{bi}) + \sum_{p=2}^{q+1} o(\mu_p) \right] \quad (42)$$

$$N^{-1} \nabla_{\mu_p} l(\theta, \mu) = N^{-1} \sum_{i=1}^N \left[(1 - d_i) \nabla_{\mu_p} \log \left\{ 1 + \sum_{j=2}^{q+1} [a_j(\varepsilon_{bi}) + a_{j-1}(\varepsilon_{bi})] \mu_j \right\} \right. \quad (43)$$

$$\left. + d_i \nabla_{\mu_p} \log \left\{ 1 + \sum_{j=2}^{q+1} a_j(\varepsilon_{bi}) \mu_j \right\} \right] \quad (44)$$

for $p = 2, \dots, q+1$. Because the restricted estimation will be the same as the ML performed in section 2, we can denote $\hat{\theta}_R$ by $\hat{\theta}$ without ambiguity. Now, notice that because $\hat{\theta}$ maximizes the restricted likelihood, (42) will be zero when evaluated at $(\hat{\theta}, 0)$. Thus, in the statistic (41), all the elements of the outer vectors are zero except for the derivative with respect to μ . The gradient with respect to μ , evaluated at the restricted parameter estimates, are

$$s_p(\hat{\theta}, x) = \nabla_{\mu_p} l(\hat{\theta}, 0) = \sum_{i=1}^N [a_p(\hat{\varepsilon}_i) - (1 - d_i) a_{p-1}(\hat{\varepsilon}_i)]. \quad (45)$$

Of particular interest are the first few s_p :

$$s_2 = N^{-1} \sum \frac{1}{2} [\hat{\varepsilon}_i^2 - 2\hat{\varepsilon}_i + 2d_i\hat{\varepsilon}_i] \quad (46)$$

$$s_3 = N^{-1} \sum -\frac{1}{6} [\hat{\varepsilon}_i^3 - 3\hat{\varepsilon}_i^2 + 3d_i\hat{\varepsilon}_i^2] \quad (47)$$

$$s_4 = N^{-1} \sum \frac{1}{24} [\hat{\varepsilon}_i^4 - 4\hat{\varepsilon}_i^3 + 4d_i\hat{\varepsilon}_i^3]. \quad (48)$$

Rather than calculating the variance matrix from (41) directly, we instead note that s_p is a conditional moment and that the test may be implemented by the methods in sections ?? and ??.

5.3 The Kinship Between the Tests

The moment conditions from the raw moments (r_p), Laguerre polynomials (λ_p), and LM test for heterogeneity (s_p) are closely related. Since the LM test uses the theoretically optimal weighting of the moments against the alternative hypothesis of multiplicative heterogeneity,¹² other moment-based tests can be viewed as sub-optimal weightings of the moments.

When the coefficients β_0 enter the hazard through $\exp(-x'_i\beta_0)$ (the usual parameterization) and x contains a constant, then λ_p and s_p are linear combinations of (r_1, \dots, r_p) for all p , where

$$r_1 = N^{-1} \sum \hat{\varepsilon}_i - 1 + d_i. \quad (49)$$

Under these conditions, equation r_1 is numerically set to zero when evaluated at the ML estimate; it is the first-order condition for the constant from the maximization of the likelihood. In these cases the relationships among the moment conditions are

$$\lambda_p = \sum_{j=1}^p \xi_{pj} r_j \quad (50)$$

$$s_p = \frac{1}{p!} (-1)^p (r_p - p r_{p-1}) \quad (51)$$

where ξ_{pj} is the coefficient on Z^j in $L_p(Z)$; see (29). For $p = 2$, all three conditional moments are numerically equal when evaluated at the ML estimate of θ . For the uncensored case, the

¹² An LM test is an asymptotically locally most powerful invariant test of the null hypothesis vs. the alternative against which it is constructed. Furthermore, in finite samples the LM test is a locally most powerful invariant test if the correct critical value is used (Engle, 1984).

equivalence between s_2 and λ_2 was noted by Sharma (1992) and the equivalence between r_2 and s_2 was noted by Pagan and Vella (1989).¹³ The result here extends this equivalence to the censored case and shows that the higher moment conditions ($p > 2$) are not equivalent. Therefore, in general the performance of the tests will differ in finite samples when moments higher than the second are included.

An interesting equivalence between the LM and Laguerre tests holds for p even larger than two when the distribution of the data are exponential and $\tilde{\Sigma}$ based on (19) is used to form the test statistics. In that case the LM and Laguerre test statistics are numerically equivalent, even though the moment conditions differ. This equivalence holds whether the data are censored or not, but it is unknown if it would extend to distributions other than the exponential.¹⁴ The equivalence does not hold if Σ_0 is estimated with $\hat{\Sigma}$ or $\tilde{\Sigma}$.

Recall that the Laguerre tests may be derived as an LM test for unexplained duration dependence. The equivalence of the Laguerre moments to the LM moments for unobserved heterogeneity when $p = 2$ follows from the kinship between duration dependence and heterogeneity. For example, it is well known that neglected heterogeneity induces apparent duration dependence into the sample and that duration dependence causes over- or under-dispersion (see, e.g., Barlow and Proschan, 1965). The kinship between duration dependence and heterogeneity is also a warning against attaching a structural interpretation to one or the other in any particular application; structural duration dependence will appear in the data as heterogeneity, and vice versa.

6 Monte Carlo Results

In this section I examine the small sample performance of four versions of the above tests applied to an exponential regression model: 1) the raw moments test performed via the auxiliary regression

¹³Prieger (1999) shows that a test based on centered moments matches the LM test even for higher moments.

¹⁴When the second and third moments are used and there is no censoring, the difference between the test statistics can be shown to be proportional to r_1 , and therefore zero when evaluated at the ML estimate. For other cases, I have verified the equivalence numerically for tests including up to fourth moments.

method with asymptotic critical values, 2) the raw moments test performed via the auxiliary regression method with bootstrap critical values, 3) the Laguerre polynomial test using $\check{\Sigma}$ [based on (19)], and 4) the LM test using $\check{\Sigma}$. It is worth emphasizing the the auxiliary regression method could be used with any conditional moment test, including the Laguerre and LM tests, and that a $\check{\Sigma}$ version could also be calculated for the raw moment test. I choose these four test versions for the following reasons. Version 1 appears to be the most commonly advocated test. For example, it is the specification test and method presented for duration data in Greene (2000), a standard graduate-level econometrics text. Given the known performance problems of version 1 and the increasing popularity of the bootstrap in econometrics, version 2 is a plausible next step after version 1. Version 3 is convenient to calculate in many cases, as explained in section 4. Version 4 is the optimal test for the alternative hypotheses in the Monte Carlo design, and so is the benchmark for the other tests.

The Monte Carlo exercises had the following design:

- The duration model is exponential, with PDF given by (52) in the appendix. The explicit moment conditions, gradient, and variance estimates are in section A.1 of the appendix.
- x is composed of a constant and a standard normal random variable. The regressors and $\beta'_0 = (1, 2)$ are fixed throughout all simulations.
- Heterogeneity takes the form $v = e^{-u}$, where v is as in (34) and $u \sim \mathcal{N}\left(\frac{\sigma^2}{2}, \sigma^2\right)$, implying that v has a lognormal distribution with $E(v) = 1$.¹⁵ This is a special case of multiplicative heterogeneity, the alternative hypothesis against which the LM test is optimal.
- The data are right-censored, with fixed censoring point c chosen to achieve a desired percentage of censoring in the data.

Consider first the performance of the tests when only the second moments are used. As is well-

¹⁵A separate set of simulations with Pareto heterogeneity led to qualitatively similar results.

known in the literature (e.g., Chesher and Spady, 1991), test statistics from auxiliary regressions converge very slowly to their asymptotic distribution. The asymptotics rely on the outer product of the gradient of the log likelihood to estimate the variance, which is known to have bad small-sample properties. The problem is exacerbated because the statistics rely on sample averages of high powers of ε , which can be poor estimates of the true expectations. Table 2 presents the actual size of the tests based on second moments, for various sample sizes and levels of censoring. The first column shows that the size of the raw moments test is far from the nominal 5% level when the asymptotic critical value is used, unless the sample sizes are large. The actual test size is about 11% when the sample size is 250 and about 7.5% when the sample size is 1,000. Censoring does not appear to make the distortion worse. When sample sizes increase to 10,000, the size drops to near the correct level, although the bias is still significant for the no-censoring case. Thus although the auxiliary regression method is convenient, it may lead to incorrect inference unless sample sizes are large. The use of the bootstrap (column two) clears up the size distortion quite well for all levels of censoring and sample sizes. None of the bootstrap sizes shows significant bias.

(Table 2 about here)

The sizes of the Laguerre and LM tests are in the final column of Table 2 (recall that the test statistics are identical for the exponential null hypothesis). The sizes of the Laguerre and LM tests tend to be on the low side, more so for smaller sample sizes, although the distortion is small compared with the auxiliary regression method. At the cost of additional computation, the bootstrap could be used to improve the sizes of these tests.

When the second and third moments are used together, the size distortion of the auxiliary regression method is greater than when the second moment alone is used (see Table 3). The levels with the second and third moments are about three times the levels with only the second moment when the sample size is 1,000 or smaller. The test statistic including the third moment contains higher powers of ε than the second-moment-only version, and the standard error of the sample

moments of ε rises with the order. The additional randomness apparently adversely affects the size of the test. Once again, the bootstrap (second column) removes most of the size distortion of the auxiliary regression tests. The sizes of the Laguerre and LM tests are again on the low side for smaller sample sizes.

(Table 3 about here)

The power of the tests against the alternative of multiplicative heterogeneity as in (34) is depicted in figures 1, 2 and 3 for various levels of censoring (none, 25%, and 50% of the sample). In these figures the second and third moment conditions are used and the sample size is 250.¹⁶ When there is no additional variance from heterogeneity (i.e. $\sigma^2 = 0$), the null hypothesis is true, and the plotted point is the size of the test. The amount of heterogeneity increases along the horizontal axis, which is scaled in the graphs to be the percentage increase in the variance of the latent duration variable due to the heterogeneity.

(figure 1 about here)

(figure 2 about here)

(figure 3 about here)

The power curves reveal the following points. First, the size distortion of the auxiliary regression raw moments test with asymptotic critical values contrasts markedly with the relatively accurate bootstrap, LM, and Laguerre tests. Second, as one would expect, the power of the tests decreases as the amount of censoring in the sample increases. As the censoring becomes more severe, there is less information in the sample.¹⁷ Third, both auxiliary regression tests (raw moments with asymptotic and bootstrap critical values) are biased: for small amounts of heterogeneity there is a smaller

¹⁶The bootstrap sample size is 99 and 100,000 iterations are performed. The power is evaluated at 8 to 16 points and curves are smoothed for plotting.

¹⁷Horowitz and Neumann (1989) and Jaggia (1997) report a similar result.

chance of rejecting the null hypothesis when false than when true.¹⁸ The bias persists over a large range of alternatives in the bootstrapped test.¹⁹ The bootstrap test is consistent²⁰ because the auxiliary regression with the true critical value is consistent in this case (Horowitz, 1997, sec. 4.6), so the bias is purely a small sample phenomenon. These results are unfortunate, however, given the convenience of auxiliary regression tests. Fourth, the Laguerre and LM tests have identical power curves (figure 1). The exponential durations, along with use of (19) to estimate the variance, ensures that the LM and Laguerre tests are numerically indistinguishable, as explained in the previous section. In such cases the Laguerre test should be used because it is easier to calculate. Finally, to the right of the region of bias the asymptotic version of the raw moments test has lower power than the Laguerre and LM tests.

Figure 4 shows the power curves with a larger sample size of 1,000 observations. The power of all the tests is higher, and the range of bias of the bootstrap test is smaller. By the time the sample size increases to 10,000 observations (Figure 5), the range and magnitude of the bootstrap bias is quite small, although the test is still less powerful than the Laguerre and LM tests.

(figure 4 about here)

(figure 5 about here)

A final question concerns the number of moments to use in the tests. Theoretically, the more moments used the higher the power of the test. This is most easily seen for the LM test: the more moments, the more accurate is the approximation of the likelihood in (39). Practically, however, the advantage of using higher moments is mitigated by the difficulties stemming from included higher powers of ε . As noted above, higher powers of ε take ever longer to converge to their theoretical averages, and may degrade the performance of the test. The trade-off appears

¹⁸This odd finding of bootstrap bias has been found in at least one other setting. In one of the bootstrapped information matrix tests for the tobit model Horowitz (1994) examines, the power against the examined alternative is less than the size.

¹⁹The power curve of the bootstrapped test does eventually approaches 100% (off the scale of the graph in figures 1–3).

²⁰A test is consistent against an alternative hypothesis if its power goes to one asymptotically.

to go against adding higher moments for the alternative hypotheses of correct baseline model specification but neglected heterogeneity, as in the Monte Carlo design above. Figure 6 shows the power of the LM and Laguerre tests (for $N=250$ and 25% censoring) as higher moments are added. The power of the test generally falls a bit when the third moment is added to the test compared to the second moment only, and falls further when the fourth moment is added to the test. Adding higher moments is likely to be of most use when the baseline duration model is incorrectly specified (Jaggia, 1991). To explore this, I run the LM and Laguerre tests when the true data generating process is lognormal and there is no heterogeneity. The lognormal distribution exhibits duration dependence, while the exponential distribution does not, so this example considers power against omitted duration dependence. The results are in Table 4. Adding higher moments did increase the power of the test: a 40% rejection rate with the second moment, a 49% rejection rate with second and third moments, and a 63% rejection rate with second, third, and fourth moments.

(Figure 6 about here)

(Table 4 about here)

The comparison among the tests is summarized informally in figures 7 and 8. In these graphs, the “ease of use” of the tests is plotted against their “accuracy”. “Ease of use” refers generally to the amount of effort required to perform the test. The auxiliary regression test without bootstrapping places the fewest demands on the econometrician and on computer time, and thus is easiest to use. When there is no censoring, the Laguerre test is about as easy to implement because no variance or matrix calculations are required. The other tests are “less easy” because they require more computer time (e.g. bootstrapping) or more effort manipulating matrices. The “accuracy” of the test is an informal amalgamation of size and power. The auxiliary regression tests score low in this dimension due to bad size (the asymptotic version) or bias in the power curve (the bootstrap version). When the data are not censored, figure 7 shows that the Laguerre test is highest in both dimensions, making it the logical choice. When there is censoring, the bootstrap test is dominated

but which of the other tests is chosen depends on the taste of the econometrician.²¹ Note that all these comparisons are based on the exponential simulations, to which proposition 1 applies. If the duration distribution does not satisfy the conditions of proposition 1, then the comparison among the tests would look like figure 8 for both censored and uncensored data.

(figure 7 about here)

(figure 8 about here)

7 Conclusion

The raw moment specification test performed via auxiliary regression is probably the most commonly used specification test for duration data. Despite the size problems with the test, Pagan and Vella (1989, p.S34) “suspect that the fact that the procedure is so easy to compute will make it attractive to many investigators....” As the bootstrap becomes more commonly used, it is natural to expect its application for these tests to clear up the size distortion. The simulations in this paper show that although the bootstrap does correct the size of the auxiliary regression test, it does so at the cost of bias and low power in general, so that this procedure cannot be recommended. I propose the Laguerre test as an alternative that is just as easy to perform as an auxiliary regression when the data are not censored and the model is exponential (or another model satisfying the conditions of proposition 1). Furthermore, the Laguerre test has higher power than either form of the raw moment test for most alternatives studied, and the same power as the optimal LM test. Adding higher moment conditions to the tests does not improve power vs. heterogeneity when the baseline hazard rate is correctly specified but does improve power when the baseline hazard itself is misspecified due to omitted duration dependence. The Laguerre and LM tests are useful new tools to assess the specification of models for censored duration data.

²¹The reader can place his indifference curves on figure 8 to find his preferred test.

References

- Abramowitz, Milton and Stegun, Irene (1964), *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*, Vol. 55 of United States, National Bureau of Standards, Applied Mathematics Series, Washington, D.C.
- Barlow, Richard E. and Proschan, Frank (1965), *Mathematical Theory of Reliability*, New York: John Wiley & Sons.
- Chesher, Andrew and Spady, Richard (1991), ‘Asymptotic Expansions of the Information Matrix Test Statistic’, *Econometrica* 59(3), 787–815.
- Cox, David R. (1972), ‘Regression Models and Life-Tables’, *Journal of the Royal Statistical Society, Series B* 34, 187–202.
- Cox, David R. and Snell, E.J. (1968), ‘A General Definition of Residuals’, *Journal of the Royal Statistical Society B* 30(2), 248–275.
- Cramér, Harald (1946), *Mathematical Methods of Statistics*, number 9 in ‘Princeton Mathematical Series’, Princeton University Press, Princeton.
- Crowley, J. and Hu, M. (1977), ‘Covariance Analysis of Heart Transplant Survival Data’, *Journal of the American Statistical Association* 72(357), 27–36.
- Engle, Robert F. (1984), Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics, in Z. Griliches and M. Intriligator, eds, ‘*Handbook of Econometrics. Volume II*’, *Handbooks in Economics*, book 2, Amsterdam; New York: North-Holland Pub. Co., chapter 13, pp. 775–826.
- Gallant, A. Ronald and Nychka, Douglas W. (1987), ‘Semi-Nonparametric Maximum Likelihood Estimation’, *Econometrica* 55(2), 363–390.

- Gourieroux, Christian, Monfort, Alain, Renault, Eric and Trognon, Alain (1987), ‘Generalised Residuals’, *Journal of Econometrics* 34, 5–32.
- Greene, William H. (2000), *Econometric Analysis*, Upper Saddle River, N.J.: Prentice Hall.
- Horowitz, Joel L. (1994), ‘Bootstrap-Based Critical Values for the Information Matrix Test’, *Journal of Econometrics* 61, 395–411.
- Horowitz, Joel L. (1997), *Bootstrap Methods in Econometrics*, in D. Kreps and K. Wallis, eds, ‘Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress. Volume 3.’, *Econometric Society (Monographs, no.28)*, Cambridge: Cambridge U. Press.
- Horowitz, Joel L. (1999), ‘Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity’, *Econometrica* 67(5), 1001–1028.
- Horowitz, Joel L. and Neumann, George R. (1989), ‘Specification Testing in Censored Regression Models: Parametric and Semiparametric Methods’, *Journal of Applied Econometrics* 4(0), S61–S86, Supplement.
- Jaggia, Sanjiv (1991), ‘Tests of Moment Restrictions in Parametric Duration Models’, *Economics Letters* 37(1), 35–38.
- Jaggia, Sanjiv (1997), ‘Alternative Forms of the Score Test for Heterogeneity in a Censored Exponential Model’, *The Review of Economics and Statistics* 79(2), 340–343.
- Jaggia, Sanjiv and Thosar, Satish (1995), ‘Contested Tender Offers: An Estimate of the Hazard Function’, *Journal of Business and Economic Statistics* 13(1), 113–119.
- Kiefer, Nicholas M. (1985), ‘Specification Diagnostics Based on Laguerre Alternatives for Econometric Models of Duration’, *Journal of Econometrics* 28(1), 135–154.
- Lancaster, Tony (1985), ‘Generalised Residuals and Heterogeneous Duration Models: With Applications to the Weibull Model’, *Journal of Econometrics* 28(1), 155–169.

- Newey, Whitney (1985), ‘Maximum Likelihood Specification Testing and Conditional Moment Tests’, *Econometrica* 53(5), 1047–1070.
- Pagan, Adrian and Vella, Frank (1989), ‘Diagnostic Tests for Models Based on Individual Data: A Survey’, *Journal of Applied Econometrics* 4(0), S29–S59.
- Prieger, James E. (1999), *Regulation, Innovation, and the Introduction of New Telecommunications Services*, PhD thesis, University of California, Berkeley.
- Sharma, Sunil (1992), ‘On Specification Diagnostics for Econometric Models of Durations’, *Journal of Quantitative Economics* 8(2), 285–307.
- Tauchien, George E. (1985), ‘Diagnostic Testing and Evaluation of Maximum Likelihood Models’, *Journal of Econometrics* 30(1).
- Tsiatis, Anastasios A. (1981), ‘A Large Sample Study of Cox’s Regression Model’, *The Annals of Statistics* 9(1), 93–108.

A Appendix

A.1 Application to the Exponential and Weibull Models

This section presents the specific form of the three tests for the exponential and Weibull duration models for the convenience of the practitioner. The tests work for any distribution; see section 2 for the general form of the test for other applications.

A.1.1 Exponential Model

For the exponential duration model with mean $\kappa_i^{-1} = \exp(x_i' \beta_0)$, the PDF is

$$f(y_i | x_i, \beta_0) = \kappa_i \exp(-\kappa_i y_i) = \exp\left(-x_i' \beta_0 - y_i e^{-x_i' \beta_0}\right). \quad (52)$$

There are no nuisance parameters ($k = \ell$). The moment conditions for the various tests are found by substituting

$$\hat{\varepsilon}_i = y_i \exp(-x_i' \hat{\beta}) \quad (53)$$

in (25)–(27), (31)–(33), and in (46)–(48).

Raw moments test The raw moments test from section 3 is typically implemented via the auxiliary regression method (see section 2.2). The ℓ -vector of scores for the auxiliary regression are

$$\hat{g}_i = [\hat{\varepsilon}_i - (1 - d_i)] x_i \quad (54)$$

for the i th observation for the exponential case. The moment conditions (25)–(27) are regressed via SUR on the scores and a constant and the constants are tested for significance.

Laguerre moments test For the Laguerre test from section 4, when Y is exponential and there is no censoring, proposition 1 applies and the asymptotic variance Σ_0 of the test statistic (12) is I . When there is censoring, Σ_0 may be estimated by $\check{\Sigma}$ based on (14) and (19). The elements of $\check{V}_{mm'}$ as defined in (19) resulting from the second through fourth moment conditions, (31)–(33), are:

$$\check{v}_{11} = 1 - N^{-1} \sum (\hat{\varepsilon}_{c_i}^2 + 1) \exp(-\hat{\varepsilon}_{c_i}) \quad (55)$$

$$\check{v}_{12} = N^{-1} \sum \left(\frac{1}{2} \hat{\varepsilon}_{c_i}^3 - \hat{\varepsilon}_{c_i}^2 + \hat{\varepsilon}_{c_i} \right) \exp(-\hat{\varepsilon}_{c_i}) \quad (56)$$

$$\check{v}_{22} = 1 - N^{-1} \sum \left(\frac{1}{4} \hat{\varepsilon}_{c_i}^4 - \hat{\varepsilon}_{c_i}^3 + 2\hat{\varepsilon}_{c_i}^2 + 1 \right) \exp(-\hat{\varepsilon}_{c_i}) \quad (57)$$

$$\check{v}_{31} = N^{-1} \sum \left(-\frac{1}{6} \hat{\varepsilon}_{c_i}^4 + \hat{\varepsilon}_{c_i}^3 - \frac{3}{2} \hat{\varepsilon}_{c_i}^2 + \hat{\varepsilon}_{c_i} \right) \exp(-\hat{\varepsilon}_{c_i}) \quad (58)$$

$$\check{v}_{32} = N^{-1} \sum \left(\frac{1}{12} \hat{\varepsilon}_{c_i}^5 - \frac{2}{3} \hat{\varepsilon}_{c_i}^4 + 2\hat{\varepsilon}_{c_i}^3 - 2\hat{\varepsilon}_{c_i}^2 + \hat{\varepsilon}_{c_i} \right) \exp(-\hat{\varepsilon}_{c_i}) \quad (59)$$

$$\check{v}_{33} = 1 - N^{-1} \sum \left(\frac{1}{36} \hat{\varepsilon}_{c_i}^6 - \frac{1}{3} \hat{\varepsilon}_{c_i}^5 + \frac{19}{12} \hat{\varepsilon}_{c_i}^4 - 3\hat{\varepsilon}_{c_i}^3 + 3\hat{\varepsilon}_{c_i}^2 + 1 \right) \exp(-\hat{\varepsilon}_{c_i}) \quad (60)$$

where \check{v}_{ij} refers to the (i, j) element of the submatrix $\check{V}_{mm'}$, $\hat{\varepsilon}_{c_i} = \varepsilon(c_i, x_i, \hat{\theta})$, c_i is the right censoring point, and all summations run from 1 to N . The other elements of (14) for the censored case are

$$\check{V}_{mg'} = \begin{bmatrix} -N^{-1} \sum \hat{\varepsilon}_{c_i} \exp(-\hat{\varepsilon}_{c_i}) x'_i \\ N^{-1} \sum \frac{1}{2} (\hat{\varepsilon}_{c_i}^2 - 2\hat{\varepsilon}_{c_i}) \exp(-\hat{\varepsilon}_{c_i}) x'_i \\ -N^{-1} \sum \frac{1}{6} (\hat{\varepsilon}_{c_i}^3 - 6\hat{\varepsilon}_{c_i}^2 + 6\hat{\varepsilon}_{c_i}) \exp(-\hat{\varepsilon}_{c_i}) x'_i \end{bmatrix} \quad (61)$$

$$\check{V}_{gg'} = N^{-1} \sum [1 - \exp(-\hat{\varepsilon}_{c_i})] x_i x'_i \quad (62)$$

Note that when there is no censoring, c_i (and therefore ε_{c_i}) may be taken to be infinite, so that (55)–(60) simplifies to I and (61) is zero as claimed.

LM test for heterogeneity For the LM test from section 5, Σ_0 may be estimated by $\check{\Sigma}$ based on (14) and (19). The elements of $\check{\Sigma}$ for the test statistic based on (46)–(48) are the following:

$$\check{v}_{11} = N^{-1} \sum \left[2 - (\hat{\varepsilon}_{c_i}^2 + 2\hat{\varepsilon}_{c_i} + 2) e^{-\hat{\varepsilon}_{c_i}} \right] \quad (63)$$

$$\check{v}_{21} = -N^{-1} \sum \left(3 - \frac{1}{2} (\hat{\varepsilon}_{c_i}^3 + 3\hat{\varepsilon}_{c_i}^2 + 6\hat{\varepsilon}_{c_i} + 6) e^{-\hat{\varepsilon}_{c_i}} \right) \quad (64)$$

$$\check{v}_{31} = N^{-1} \sum \left(4 - \frac{1}{6} (\hat{\varepsilon}_{c_i}^4 + 4\hat{\varepsilon}_{c_i}^3 + 12\hat{\varepsilon}_{c_i}^2 + 24\hat{\varepsilon}_{c_i} + 24) e^{-\hat{\varepsilon}_{c_i}} \right) \quad (65)$$

$$\check{v}_{22} = N^{-1} \sum \left(6 - \frac{1}{4} (\hat{\varepsilon}_{c_i}^4 + 4\hat{\varepsilon}_{c_i}^3 + 12\hat{\varepsilon}_{c_i}^2 + 24\hat{\varepsilon}_{c_i} + 24) e^{-\hat{\varepsilon}_{c_i}} \right) \quad (66)$$

$$\check{v}_{32} = -N^{-1} \sum \left(10 - \frac{1}{12} (\hat{\varepsilon}_{c_i}^5 + 5\hat{\varepsilon}_{c_i}^4 + 20\hat{\varepsilon}_{c_i}^3 + 60\hat{\varepsilon}_{c_i}^2 + 120\hat{\varepsilon}_{c_i} + 120) e^{-\hat{\varepsilon}_{c_i}} \right) \quad (67)$$

$$\check{v}_{33} = N^{-1} \sum \left(20 - \frac{1}{36} (\hat{\varepsilon}_{c_i}^6 + 6\hat{\varepsilon}_{c_i}^5 + 30\hat{\varepsilon}_{c_i}^4 + 120\hat{\varepsilon}_{c_i}^3 + 360\hat{\varepsilon}_{c_i}^2 + 720\hat{\varepsilon}_{c_i} + 720) e^{-\hat{\varepsilon}_{c_i}} \right) \quad (68)$$

$$\check{V}_{mg'} = N^{-1} \sum \begin{bmatrix} 1 - (\hat{\varepsilon}_{c_i} + 1) e^{-\hat{\varepsilon}_{c_i}} \\ -1 + \left(\frac{1}{2} \hat{\varepsilon}_{c_i}^2 + \hat{\varepsilon}_{c_i} + 1 \right) e^{-\hat{\varepsilon}_{c_i}} \\ 1 - \left(\frac{1}{6} \hat{\varepsilon}_{c_i}^3 + \frac{1}{2} \hat{\varepsilon}_{c_i}^2 + \hat{\varepsilon}_{c_i} + 1 \right) e^{-\hat{\varepsilon}_{c_i}} \end{bmatrix} x'_i \quad (69)$$

$\check{V}_{gg'}$ is as in (62).

A.1.2 Weibull Model

For the Weibull duration model (Weibull I from table 1), we have $\theta_0 = (\beta'_0, \sigma_0)'$; σ_0 is a scalar shape parameter controlling duration dependence ($\dim(\theta_0) = \ell + 1$). The PDF and generalized residual are

$$f(y_i|x_i, \theta_0) = (\sigma_0 y_i)^{-1} \left(y_i e^{-x'_i \beta_0} \right)^{1/\sigma_0} \exp \left[- \left(y_i e^{-x'_i \beta_0} \right)^{1/\sigma_0} \right] \quad (70)$$

and

$$\hat{\varepsilon}_i = \left[y_i \exp(-x'_i \hat{\beta}) \right]^{1/\hat{\sigma}}. \quad (71)$$

Raw moments test The $(\ell + 1)$ -vector of scores for the auxiliary regression (section 2.2) are

$$\hat{g}_i = \frac{1}{\hat{\sigma}} \begin{bmatrix} [\hat{\varepsilon}_i - (1 - d_i)] x_i \\ \hat{\varepsilon}_i \log \hat{\varepsilon}_i - (1 - d_i) (1 + \log \hat{\varepsilon}_i) \end{bmatrix} \quad (72)$$

for the i th observation for the Weibull case.²² The moment conditions (25)–(27) are regressed via SUR on the scores and a constant and the constants are tested for significance.

Laguerre moments test Since the Weibull model with unknown shape parameter is not in the class of distributions for which $M_0 = 0$ for the Laguerre tests (section 4), the asymptotic variance of the Laguerre test does not simplify to I even when there is no censoring. For uncensored observations, $\check{\Sigma}$ may be calculated as for the exponential model above. $\check{V}_{mm'} = I$ when there is no censoring (this is true for any distribution). The other elements needed for $\check{\Sigma}$ for the Weibull model and moments (31)–(33) are:

$$\check{V}_{gg'} = \frac{1}{\hat{\sigma}^2 N} \sum \begin{bmatrix} x_i x'_i & (1 - \gamma) x_i \\ (1 - \gamma) x'_i & \frac{1}{6} \pi^2 + (1 - \gamma)^2 \end{bmatrix} = \frac{1}{\hat{\sigma}^2 N} \sum \begin{bmatrix} x_i x'_i & 0.42278 x_i \\ 0.42278 x'_i & 1.82368 \end{bmatrix} \quad (73)$$

$$\check{V}_{g\beta m'} = \mathbf{0} \quad (74)$$

$$\check{V}_{g\sigma m'} = \hat{\sigma}^{-1} \begin{bmatrix} 1 & 1/2 & 1/3 \end{bmatrix} \quad (75)$$

²²If a parameterization of σ such as $\varsigma = \log(\sigma)$ is chosen for the ML routine, the final row of (72) needs to be adjusted accordingly.

where $\gamma = 0.57721566$ is Euler's constant.

Estimates (73)–(75) do not apply when the data are censored. The analytical estimate $\check{\Sigma}$ in this case contains partial gamma, digamma, and trigamma functions, making it computationally unattractive. Therefore a simpler estimate of Σ_0 is obtained from $\check{\Sigma}$, the estimate of Σ_0 based on plugging $\hat{V}_{mm'}$, \hat{M} , and $\hat{\mathcal{J}}$ into (15) (see section 2.2). $\hat{V}_{mm'} \equiv N^{-1} \sum m_i m'_i$ is found by using the summands of (31)–(33) for m_i . The other pieces of $\check{\Sigma}$ are:

$$\hat{M}_{\beta'} = \frac{1}{\hat{\sigma}N} \sum \varepsilon_i \begin{bmatrix} -\varepsilon_i + 2 - d_i \\ \frac{1}{2}\varepsilon_i^2 - 3\varepsilon_i + 3 + d_i (\varepsilon_i - 2) \\ -\frac{1}{6}\varepsilon_i^3 + 2\varepsilon_i^2 - 6\varepsilon_i + 4 - d_i \left(\frac{1}{2}\varepsilon_i^2 - 3\varepsilon_i + 3\right) \end{bmatrix} x'_i \quad (76)$$

$$\hat{M}_{\sigma} = \frac{1}{\hat{\sigma}N} \sum \varepsilon_i \log(\varepsilon_i) \begin{bmatrix} -\varepsilon_i + 2 - d_i \\ \frac{1}{2}\varepsilon_i^2 - 3\varepsilon_i + 3 + d_i (\varepsilon_i - 2) \\ -\frac{1}{6}\varepsilon_i^3 + 2\varepsilon_i^2 - 6\varepsilon_i + 4 - d_i \left(\frac{1}{2}\varepsilon_i^2 - 3\varepsilon_i + 3\right) \end{bmatrix} \quad (77)$$

$$\hat{\mathcal{J}}_{\beta\beta'} = (\hat{\sigma}^2 N)^{-1} \sum \varepsilon_i x_i x'_i \quad (78)$$

$$\hat{\mathcal{J}}_{\sigma\beta'} = (\hat{\sigma}^2 N)^{-1} \sum [(\log \varepsilon_i + 1) \varepsilon_i - (1 - d_i)] x'_i \quad (79)$$

$$\hat{\mathcal{J}}_{\sigma\sigma} = (\hat{\sigma}^2 N)^{-1} \sum [(\log \varepsilon_i + 2) \varepsilon_i \log \varepsilon_i - (1 - d) (2 \log \varepsilon_i + 1)] \quad (80)$$

where $\hat{M}_{\beta'}$ contains the first ℓ columns (those pertaining to β_0) of \hat{M} , \hat{M}_{σ} is the final column of \hat{M} , and $\hat{\mathcal{J}}_{\beta\beta'}$, $\hat{\mathcal{J}}_{\sigma\beta'}$, and $\hat{\mathcal{J}}_{\sigma\sigma}$ are the obvious partitions of $\hat{\mathcal{J}}$.

LM test for heterogeneity For the LM test (section 5) with uncensored observations, $\check{\Sigma}$ may be calculated as for the exponential model above. $\check{V}_{mm'} = I$, $\check{V}_{gg'}$ is as in (73), and the other elements needed for $\check{\Sigma}$ for the Weibull model and moments (46)–(48) are:

$$\check{V}_{g\beta m'} = (\hat{\sigma}N)^{-1} \sum x_i \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \quad (81)$$

$$\check{V}_{g_{\sigma}m'} = \hat{\sigma}^{-1} \begin{bmatrix} 2 - \gamma \\ \gamma - \frac{5}{2} \\ \frac{17}{6} - \gamma \end{bmatrix}' = \hat{\sigma}^{-1} \begin{bmatrix} 1.4228 \\ -1.9228 \\ 2.2561 \end{bmatrix}' \quad (82)$$

When the data are censored, $\tilde{\Sigma}$ should be used instead. $\hat{V}_{mm'} \equiv N^{-1} \sum m_i m'_i$ is found by using the summands of (46)–(48) for m_i . $\hat{\mathcal{J}}$ is as in (78)–(80). The other pieces of $\tilde{\Sigma}$ are:

$$\hat{M}_{\beta} = \frac{1}{\hat{\sigma}N} \sum \varepsilon_i \begin{bmatrix} -\varepsilon_i + (1 - d_i) \\ \frac{1}{2}\varepsilon_i^2 - (1 - d_i)\varepsilon_i \\ -\frac{1}{6}\varepsilon_i^3 + \frac{1}{2}(1 - d_i)\varepsilon_i^2 \end{bmatrix} x'_i \quad (83)$$

$$\hat{M}_{\sigma} = \frac{1}{\hat{\sigma}N} \sum \varepsilon_i \log(\varepsilon_i) \begin{bmatrix} -\varepsilon_i + (1 - d_i) \\ \frac{1}{2}\varepsilon_i^2 - (1 - d_i)\varepsilon_i \\ -\frac{1}{6}\varepsilon_i^3 + \frac{1}{2}(1 - d_i)\varepsilon_i^2 \end{bmatrix} \quad (84)$$

where the definitions are as above.

A.2 Miscellaneous Results

A.2.1 Proof of Proposition 1

Consider $\nabla_{\beta}\lambda_p$, the p th element of m . We have $\nabla_{\beta}\lambda_p = \frac{1}{N} \sum L'_p(\varepsilon_i) \nabla_{\beta}\varepsilon_i = -\frac{1}{N}\alpha \sum \varepsilon_i L'_p(\varepsilon_i) x_i$, so it suffices to show that $E(\varepsilon_i L'_p(\varepsilon_i)) = 0$. For Laguerre polynomials, the recursion $\varepsilon_i L'_p(\varepsilon_i) = p[L_p(\varepsilon_i) - L_{p-1}(\varepsilon_i)]$ (Abramowitz and Stegun, 1964, p.783) and orthogonality property (28) imply that $E(\varepsilon_i L'_p(\varepsilon_i))$ is indeed zero for all p .

A.2.2 Expectations of the censored generalized residual

The expectation ε_i^* when the observed variable ε_i is censored is:

$$E(\varepsilon_i^* | \varepsilon_i, d = 1) = \int_{c_i}^{\infty} \varepsilon(t, x_i, \theta_0) \frac{f(t | x_i, \theta_0)}{1 - F(c_i | x_i, \theta_0)} dt$$

Using the identity $\varepsilon = -\log(1 - F)$ and change of variables $u = S(t)$ yields

$$E(\varepsilon_i^* | \varepsilon_i, d = 1) = [1 - F(c_i)]^{-1} \int_{1-F(c_i)}^0 (-\log u)(-du) = \varepsilon_i + 1.$$

Similar calculation for higher powers ε_i^* of leads to (23).

A.3 Monte Carlo Exercise Details

A particular simulation includes these steps:

1. Load the initially generated and fixed x matrix, and form $\lambda = e^{-x'\beta}$, an $N \times 1$ vector. This vector is held fixed through all iterations.
2. Monte Carlo loop, to be performed R times for each particular σ^2 :
 - (a) Generate heterogeneity term v , an $N \times 1$ vector, if $\sigma^2 > 0$.
 - (b) Generate N exponential random deviates of rate $v\lambda$.
 - (c) Censor the duration variable, if greater than the right-censoring point c .²³
 - (d) Compute the ML estimate $\hat{\beta}$ for β_0 .
 - (e) Form the generalized residuals and the moment conditions (and the score vector for the auxiliary regression tests).
 - (f) Form the test statistics using the desired moments:
 - i. Raw moments test statistic using (25) alone or both (25) and (26), generated via the auxiliary regression method described in sections 2.2 and A.1.1: regress the moment conditions on the scores and constants (a SUR if both (25) and (26) are used); form the joint test statistic for the significance of the constants. This statistic is referred to the asymptotic critical value for a $\chi^2(1)$ or $\chi^2(2)$ random variable as appropriate.

²³The censoring point is determined by a subroutine that performs steps (a) and (b) and picks the quantile of the resulting pseudo-data that leads to the desired level of censoring. These pseudo-data are then discarded. After getting two such quantiles, c is set to their average.

- ii. Same as previous, but the statistic is referred to a bootstrap critical value. The size of the bootstrap sample is 999 for the size calculations and 99 for the power curves.²⁴
 - iii. Laguerre polynomial test statistic using (31) alone or both (31) and (32) and the asymptotic variance as calculated in section A.1.1. This statistic is referred to the asymptotic critical value.
 - iv. LM test statistic using (46) alone or both (46) and (47) and the asymptotic variance as calculated in section A.1.1. This statistic is referred to the asymptotic critical value.
- (g) Test the statistics using the relevant critical value, and record acceptance or rejection.

3. Report the percentage of rejections as the power of the tests for the chosen σ^2 .

²⁴The method is the parametric bootstrap (Horowitz, 1997); the paired (natural) bootstrap yielded qualitatively similar results (including the persistence of the bootstrap bias).

Model	Integrated Hazard		Asymptotically Orthogonal
	$\varepsilon = \frac{[\exp(-\beta'x)]^\alpha \Lambda(y)}{\Lambda(y)}$	α	
exponential	y	1	Yes
Rayleigh	$y^{1/2}$	0.5	Yes
proportional hazards	$\Lambda_0(y_i)$	1	Yes*
Weibull 1 [†]	$y^{1/\sigma}$	$1/\sigma$	No [‡]
Weibull 2 [†]	$y^{1/\sigma}$	1	No [‡]

*The baseline hazard Λ_0 is taken to be known. If not, its estimation adds to the variance of the estimated generalized residual (Tsiatis, 1981) and the results of this paper do not apply.

[†]There are two forms of the Weibull model in the literature.

[‡]If the Weibull shape parameter σ is known, then Yes.

Table 1: Integrated Hazard and Asymptotic Orthogonality of the Laguerre Test for Various Distributions

Test	Raw Moment Test		Laguerre and LM Tests
	Asymptotic Critical Values	Bootstrap Critical Values	
<hr/>			
$N = 250$			
No censoring	0.116*	0.050	0.041*
25% censoring	0.107*	0.050	0.041*
50% censoring	0.110*	0.050	0.040*
$N = 1,000$			
No censoring	0.075*	0.050	0.048
25% censoring	0.073*	0.050	0.048*
50% censoring	0.074*	0.048	0.046*
$N = 10,000$			
No censoring	0.056*	0.048	0.051
25% censoring	0.055	0.050	0.050
50% censoring	0.052	0.047	0.051
<hr/>			

Table notes: Nominal size is 5%; * indicates significant (1% level) bias in the empirical size. N is sample size. Raw moment tests are performed via the auxiliary regression method (see section 2.2). Censoring is accomplished with a fixed right censoring point common to all observations. Bootstrap sample size = 999. For raw moment tests, number of Monte Carlo trials is 100,000 for $N=250$, 25,000 for $N=1,000$, and 10,000 for $N=10,000$. For Laguerre and LM tests, number of Monte Carlo trials is 100,000 for all sample sizes. The Laguerre and LM tests are numerically indistinguishable; the final column is the results from either test.

Table 2: Empirical Levels of the Tests with Second Moments

Test	Raw Moment Test		Laguerre and LM Tests
	Asymptotic Critical Values	Bootstrap Critical Values	
$N = 250$			
No censoring	0.313*	0.050	0.040*
25% censoring	0.296*	0.052*	0.038*
50% censoring	0.311*	0.050	0.033*
$N = 1,000$			
No censoring	0.192*	0.050	0.047*
25% censoring	0.190*	0.051	0.047*
50% censoring	0.212*	0.052	0.044*
$N = 10,000$			
No censoring	0.087*	0.049	0.050
25% censoring	0.092*	0.050	0.049
50% censoring	0.104*	0.049	0.051

Table notes: Nominal size is 5%; * indicates significant (1% level) bias in the empirical size. N is sample size. Raw moment tests are performed via the auxiliary regression method (see section 2.2). Censoring is accomplished with a fixed right censoring point common to all observations. Bootstrap sample size = 999. For raw moment tests, number of Monte Carlo trials is 100,000 for $N=250$, 25,000 for $N=1,000$, and 10,000 for $N=10,000$. For Laguerre and LM tests, number of Monte Carlo trials is 100,000 for all sample sizes. The Laguerre and LM tests are numerically indistinguishable; the final column is the results from either test.

Table 3: Empirical Levels of the Tests with Second and Third Moments

Moments	Laguerre Test	LM Test
Second	0.3951	0.3953
Second and Third	0.4901	0.4899
Second, Third, and Fourth	0.6357	0.6353

Table notes: $N = 250$. The true data generating process is lognormal with no heterogeneity; the (false) null hypothesis is an exponential model. Censoring is accomplished with a fixed right censoring point common to all observations so that 25% of the sample is censored. Number of Monte Carlo trials is 100,000.

Table 4: Power of the Laguerre and LM Tests vs. Incorrect Baseline Model as the Number of Moments Increases

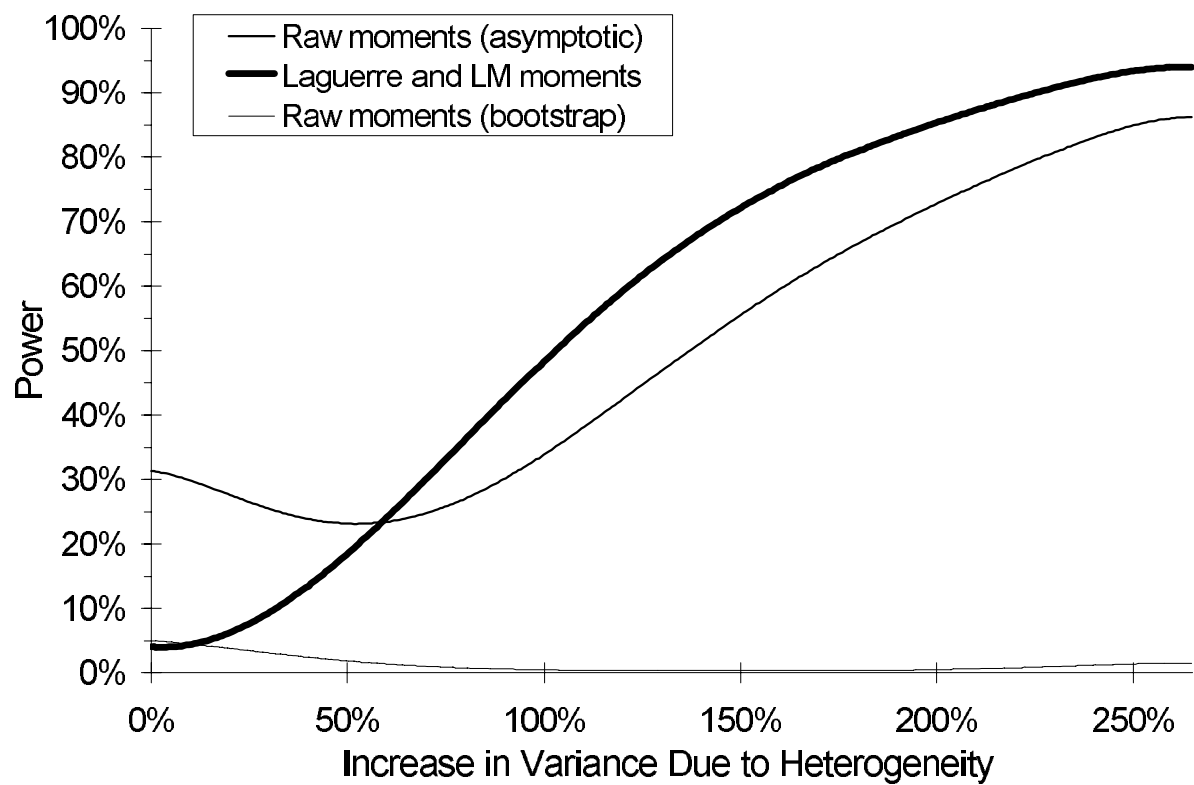


Figure 1: Power curves for the tests vs. lognormal multiplicative heterogeneity (no censoring, 2nd and 3rd moments, $N=250$)

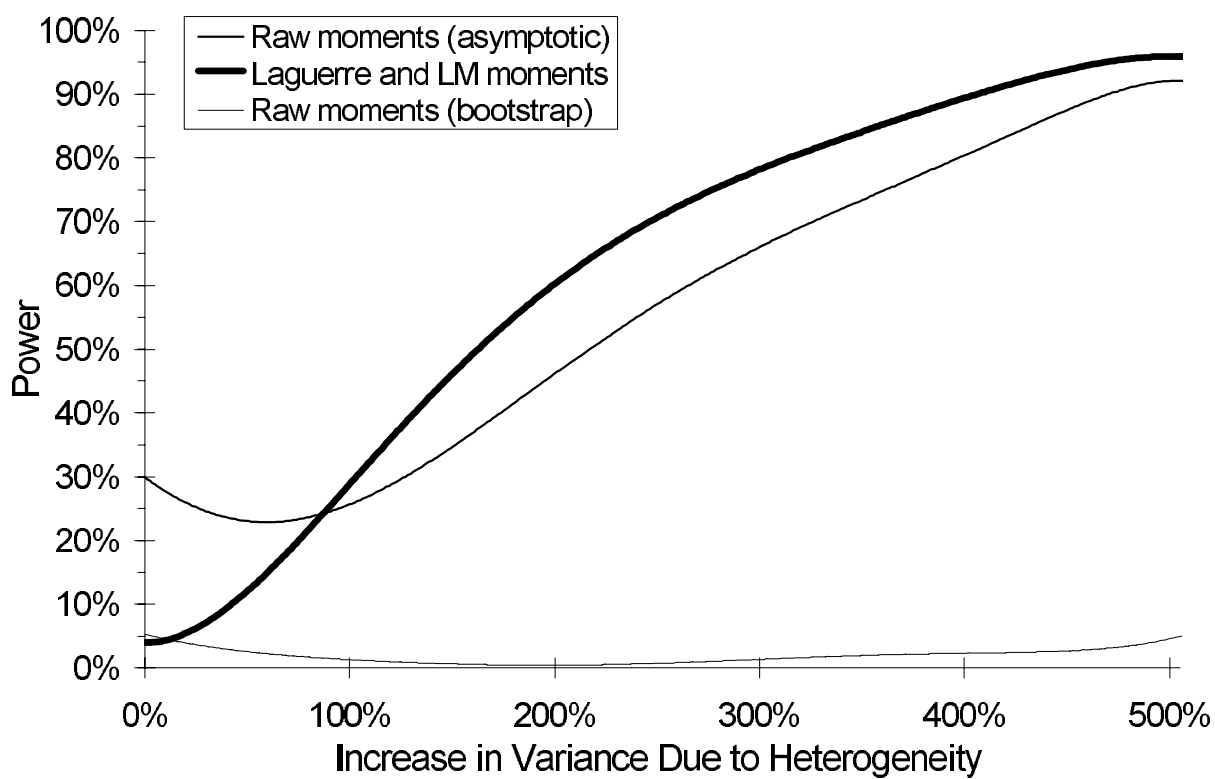


Figure 2: Power curves for the tests vs. lognormal multiplicative heterogeneity (25% censoring, 2nd and 3rd moments, $N=250$)

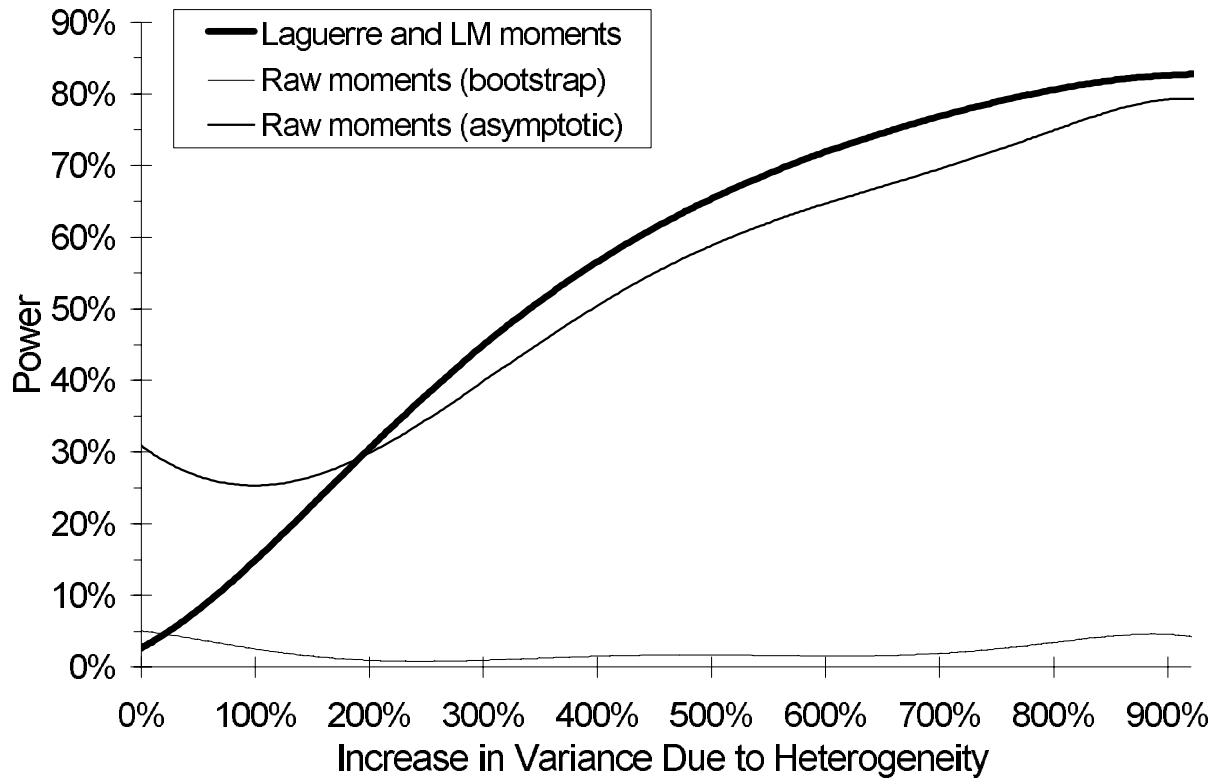


Figure 3: Power curves for the tests vs. lognormal multiplicative heterogeneity (50% censoring, 2nd and 3rd moments, $N=250$)

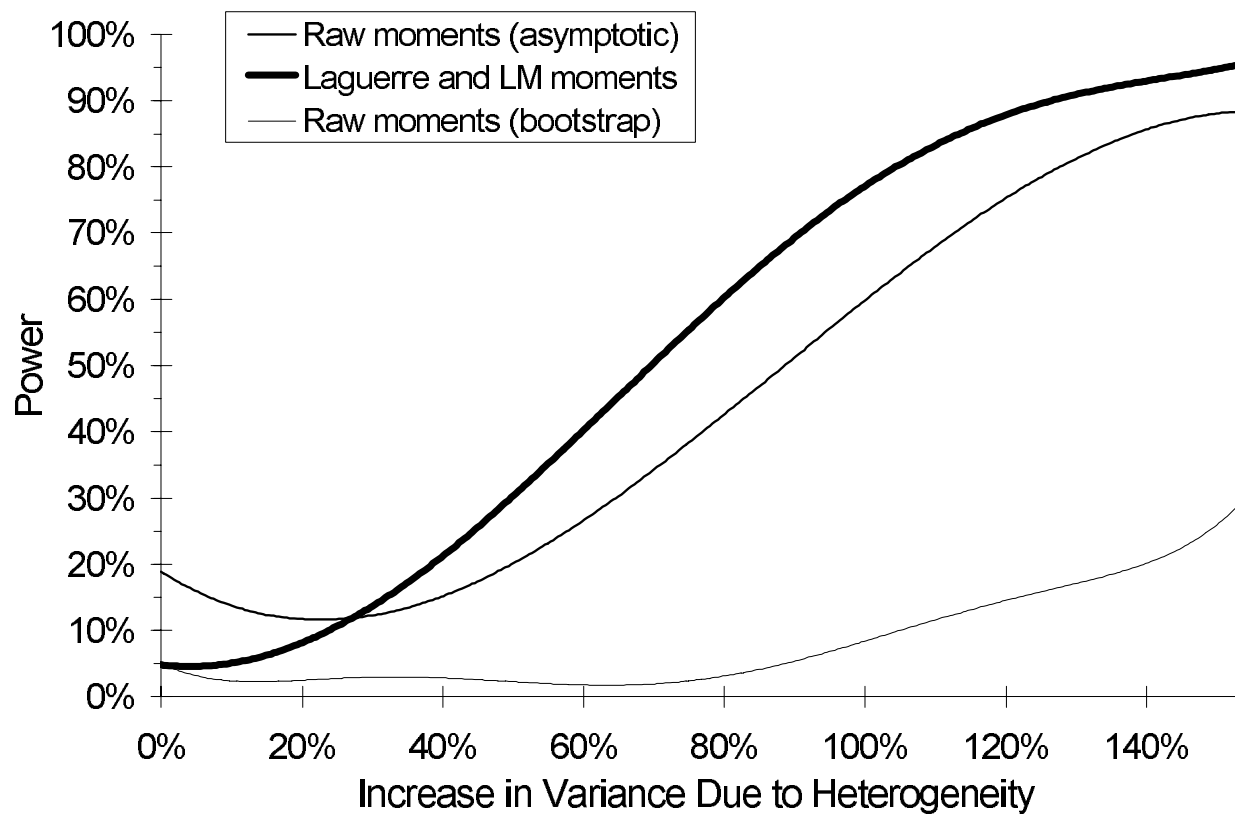


Figure 4: Power curves for the tests vs. lognormal multiplicative heterogeneity (25% censoring, 2nd and 3rd moments, $N=1,000$)

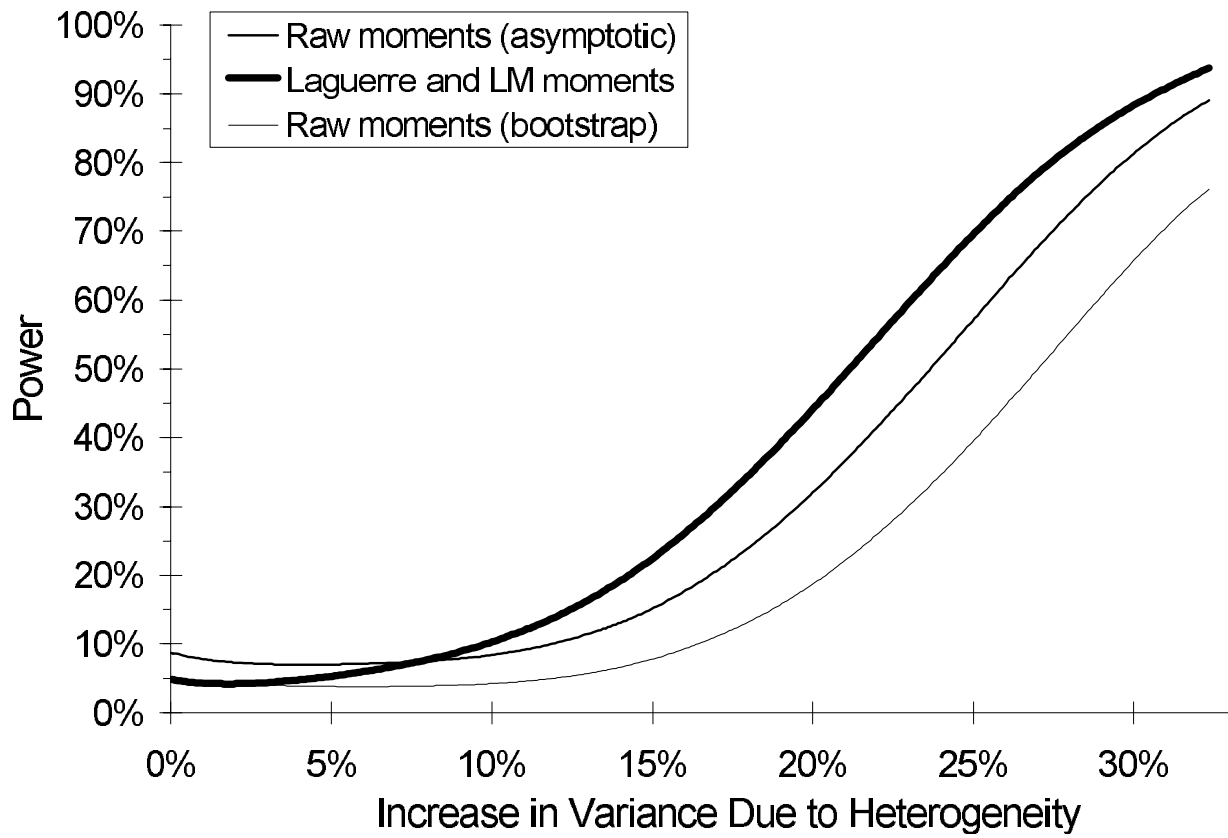


Figure 5: Power curves for the tests vs. lognormal multiplicative heterogeneity (no censoring, 2nd and 3rd moments, $N=10,000$)

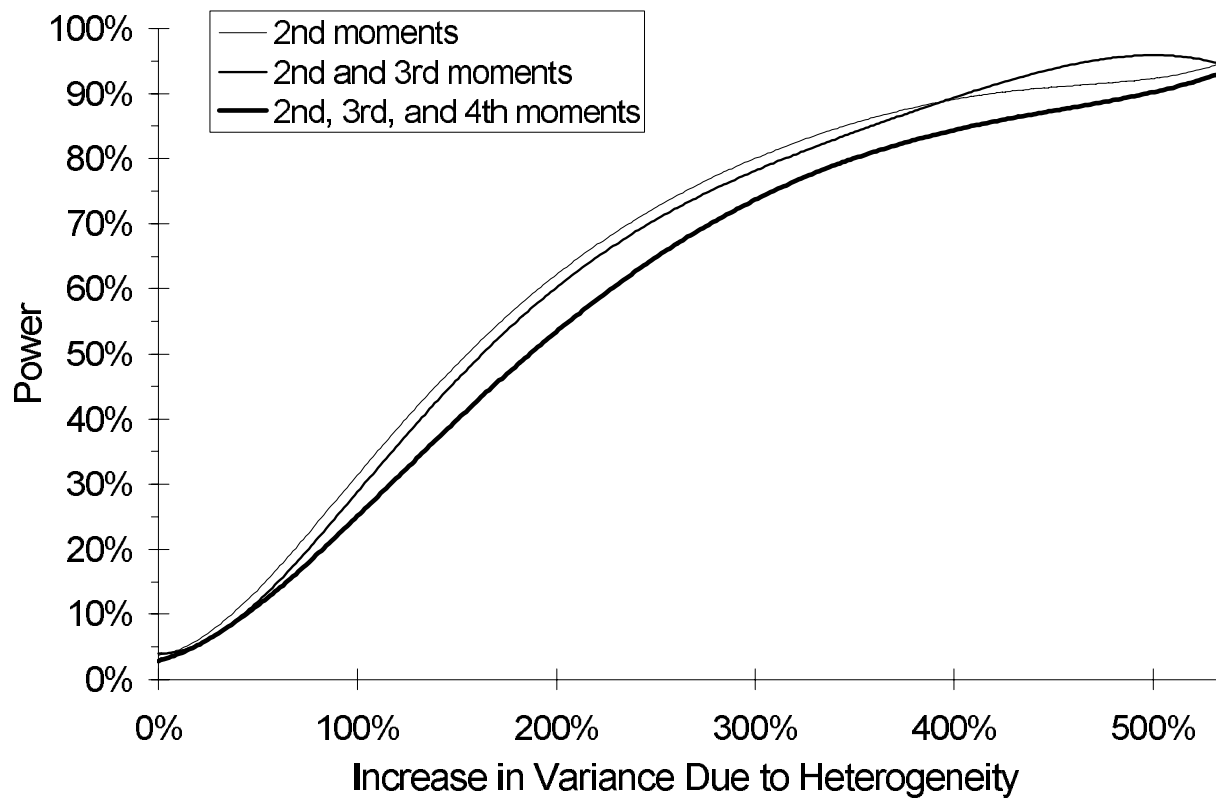


Figure 6: Power curves for the LM and Laguerre tests (25% censoring, N=250)

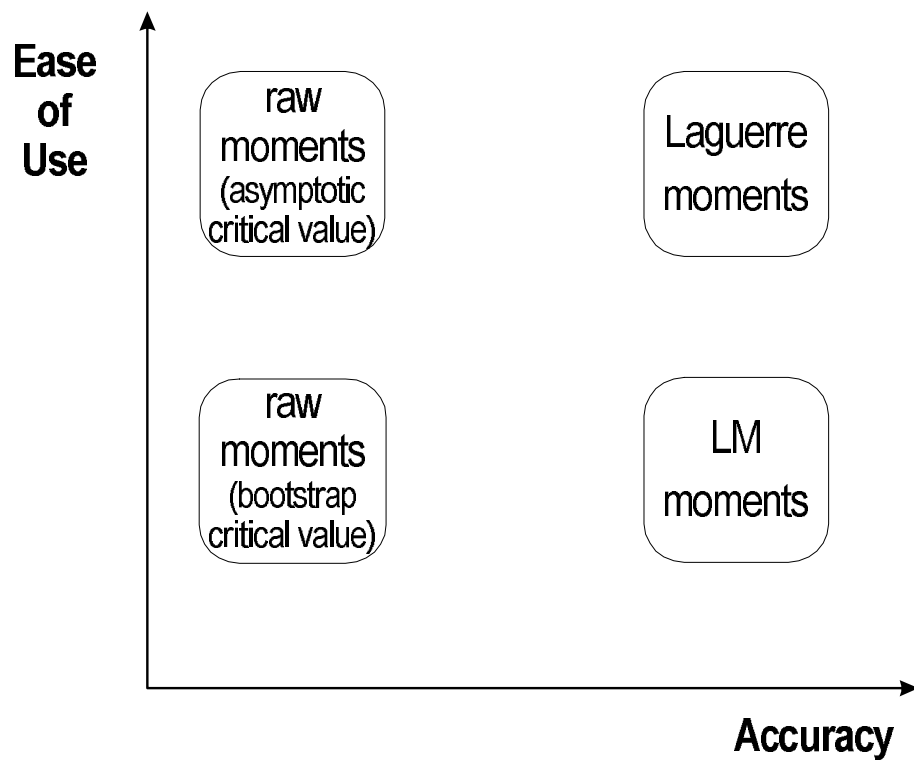


Figure 7: Comparison of the Tests when the Data are Uncensored

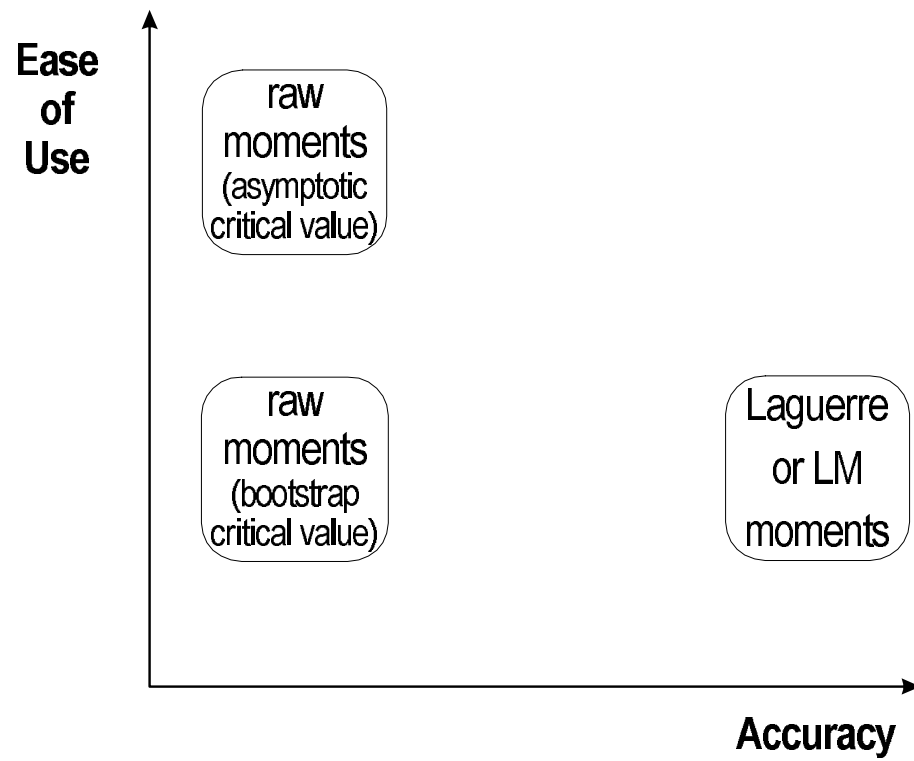


Figure 8: Comparison of the Tests when the Data are Censored