VATT-KESKUSTELUALOITTEITA
VATT-DISCUSSION PAPERS

32

NONLINEARITY AND HETEROGENEITY IN WEIBULL DURATION ANALYSIS:

A test for misspecification with one degree of freedom

Ilpo Suoniemi
Juha Rantala

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 1992

ISBN 951-561-041-9

ISSN 0788-5016

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Hämeentie 3, 00530 Helsinki, Finland

Valtion painatuskeskus

Pasilan VALTIMO

Helsinki 1992

NONLINEARITY AND HETEROGENEITY IN WEIBULL DURATION ANALYSIS: A test for misspecification with one degree of freedom.

ILPO SUONIEMI & JUHA RANTALA
GOVERNMENT INSTITUTE FOR ECONOMIC RESEARCH

SUONIEMI, ILPO AND JUHA RANTALA: NONLINEARITY AND HETEROGENEITY IN WEIBULL DURATION ANALYSIS: A TEST FOR MISSPECIFICATION WITH ONE DEGREE OF FREEDOM. Helsinki: VATT, Valtion taloudellinen tutkimuskeskus, 1992. (C, ISSN 0788-5016, No 32). ISBN 951-561-041-9.

ABSTRACT: In contrast with regression models duration models are not robust against violation of the distributional assumptions. There is thus a need for specification tests which have power against a wide range of alternatives and are easy to implement. In this paper the transformation family introduced by MacKinnon and Magee is applied in the Weibull duration framework to derive a score test for misspecification with one degree of freedom. The test statistic is found to be sensitive for violation of three conditional moment restrictions. The test is compared with a form of the RESET-test as well as with score tests for sample heterogeneity in the location component and such type of duration dependence which is related to the location component. An empirical example is presented in which the derived test clearly indicated violation of the third moment restriction in the disturbances and was more powerful than the alternatives considered. A decomposition of the test statistic simultaneously accounts for the results given by the other two score tests. It is suggested that the test statistic should be routinely used in model diagnostics.

KEY WORDS: Weibull duration models, misspecification, functional form.

SUONIEMI, ILPO AND JUHA RANTALA: NONLINEARITY AND HETEROGENEITY IN WEIBULL DURATION ANALYSIS: A TEST FOR MISSPECIFICATION WITH ONE DEGREE OF FREEDOM. Helsinki: VATT, Valtion taloudellinen tutkimuskeskus, 1992. (C, ISSN 0788-5016, No 32). ISBN 951-561-041-9.

TIIVISTELMÄ: Tavallisista regressiomalleista poiketen duraatiomallit ovat herkkiä spesifikaatiovirheille jakaumaoletuksien ja otoksen heterogeenisuuden suhteen. Tästä syystä tarvitaan spesifikaatiotestejä, joilla on voimaa useiden vaihtoehtojen suhteen ja joita on helppo soveltaa. MacKinnonin ja Mageen esittämän muunnosperheen avulla johdetaan yhden vapausasteen pisteytystesti Weibull mallin spesifikaatiolle. Testisuure osoittautuu herkäksi virhetermin kolmen ehdollisten momenttien välisen rajoituksen pätemiselle. Testiä verrataan sekä eräänlaiseen RESET -testiin että yhden vapausasteen pisteystesteihin, jotka testaavat otoksen heterogeenisuutta sijaintiparametrin ja sellaisen duraatioriippuvuuden suhteen, joka riippuu sijaintikomponentin koosta. Esitettävässä empiirisessä esimerkissä testi paljastaa selvästi kolmannen momenttirajoituksen pätemättömyyden ja on samalla voimakkaampi kuin em. vaihtoehdot. Testitunnusluku voidaan jakaa tekijöihin, jotka paljastavat jälkimmäisten pisteystestien tulokset. Testitunnuslukua suositellaan rutiinikäyttöön osana mallin diagnostista tarkastelua.

AVAINSANAT: Weibull kestomallit, väärinspesifiointi, funktiomuoto.

CONTENTS

1. INTRODUCTION

Duration models based on parametric assumptions such as exponential, Weibull or generalized Gamma distribution are widely used to analyze count processes in econometrics. They have been particularly popular in applied labour economics. These models are usually estimated by solving numerically the nonlinear estimating equations given by the principle of maximum likelihood. Therefore, investigators often display a natural reluctance to test the specification of the model with the thoroughness and vigor comparable to the common practice in the normal regression framework. In contrast it is well known, however, that duration models are not robust against specification errors such as sample heterogeneity and violation of the distributional assumptions. There is thus a need for specification tests of duration models which have power against a wide range of alternatives and are easy to implement and inexpensive to compute to complement other methods of model evaluation.[1]

In this context it seems natural to consider the use of score, or LM, tests because they require estimates only under the null hypothesis and can often be computed by means of artificial linear regressions. In this paper a misspecification test is derived by considering a possible misspecification in the transformation applied to the observed values of the dependent variable in the Weibull duration model.

In the statistical analysis transformations of the dependent variable are used to obtain three objectives. The first is to normalize the random variable in question. The second objective is to stabilize its variance. A classic example of a simultaneously normalizing and variance stabilizing transformation concerns the sample correlation coefficient of a

---

[1] In evaluating duration models informal graphical methods which are based on plots of generalized residuals of the model have proven to be of great value, see Aitkin & Clayton (1981), Lancaster (1985) & (1990), Lancaster & Chesher (1985), and Kiefer (1985).

bivariate normal distribution, where the transformation $\tanh^{-1}$ is used. In regression (duration) models the third objective is to obtain linearity of the conditional mean (location or the proportional hazard component) for data which have been conveniently transformed.

In econometrics the family of monotonic power transformations is widely used to obtain the above objectives (Box and Cox, 1964). However, the Box-Cox transformation has a rather limited use in Weibull duration models, since it does not allow for a non-monotonous hazard function.

MacKinnon and Magee (1990) have proposed a family of transformations which can sensibly be applied to variables that can take negative values. It is interesting to apply their method of analyzing ordinary regression models to duration models with logarithmic dependent variables. MacKinnon and Magee derive score tests for the null hypothesis that the dependent variable has not been transformed against the alternative that a transformation of this family has been applied to it. These tests, which do not require that the exact form of the transformation is specified and are thus interpretable as implicit misspecification tests in the sense of Hausman (1978), are in this paper extended to Weibull duration models. To be more specific, this is accomplished by applying the transformation to logarithmic duration, i.e. to a variable which has the extreme value distribution. The misspecification test is based on the score under the null hypothesis and has one degree of freedom.

In case of misspecification in a duration model, the estimated disturbances of logarithmic durations would be expected to suffer from problems which would affect their conditional moment restrictions in a way not allowed by an ordinary extreme value model. The above implicit test for misspecification based on the MacKinnon-Magee family of transformations has some power against a wide range of alternative models. This is attributable to the way in which information on the non-linear model of the location is confounded with information concerning the distribution of disturbances. The lack of fit detected may be due to the location model, or the disturbance model, or both. This feature is common to all transformation families, eg. Box-Cox, that affect the values of the dependent variable. To cater for the above situation some additional score tests are discussed in the the paper. The tests specialize in testing non-linearity in the location component (RESET test) and sample heterogeneity in the duration shape parameter and the location component, respectively. The relative success of these three tests in rejecting the null would suggest whether modifications in modelling the location component, or sample heterogeneity, or the form of duration dependence in the data would be necessary.

In analogy with the results that MacKinnon and Magee derive in an ordinary regression model, the transformation test presented for Weibull duration models can be seen as testing simultaneously for three restrictions that affect conditional moments. The first restriction, which is closely related to the well-known RESET test of Ramsey (1969), is that there is no correlation between the squared conditional location component and the generalized residuals of the model. In this case one has to correct the residuals for their expected values under possible censoring.

The second restriction is that there is no correlation between the conditional location component and zero mean 'disturbance terms' of the model. The latter are implicitely defined by a second order conditional moment restriction, which is satisfied by the estimated generalized residuals and 'disturbances' of the model.

The third restriction directly tests for a third order conditional moment restriction under the null. The restriction is related to a well known test for sample heterogeneity, see eg. Burdett et. al (1985) and Lancaster (1985). The latter test is in this paper shown to test whether the moment generating function of the estimated disturbances of the model has its theoretical value at point two, $\varphi(2)$, while the disturbances follow the extreme value distribution under the null hypothesis. In contrast the third order moment restriction considered in this paper tests for the theoretical value of the second derivative of the moment generating function at the point one, $\varphi''(1)$, under the null. Keeping in mind that the estimated disturbances always satisfy the theoretical values of $\varphi(1)$ and $\varphi'(1)$, under the null, the two tests are not identical but closely related.

Sample heterogeneity causes the parameter estimates from a Weibull model to be inconsistent. It is thus a more severe problem than in the case of ordinary regression. In addition, because duration models are typically estimated using cross-sectional data it is a problem likely to be encountered quite often. Therefore a score test based on a specific form of sample heterogeneity which affects the duration shape parameter is discussed as an additional diagnostic tool. The test which turns out to be direcly related to the transformation test is inspired by an obvious analogy with heteroscedasticity tests in a standard regression model. The one degree of freedom form of the heteroscedasticity test is originally due to Anscombe (1961).

An additional interesting feature of the paper is that in deriving the test statistics and moment restrictions we explicitely allow for a censoring mechanism by an independent latent variable. The independence assumption allows us to consider the censoring mechanism as nuisance parameters which do not interfere neither with efficient estimation nor statistical inference on the parameters of interest.

In the final sections of the paper we present a straightforward extension of the test to a competing risks framework and illustrate the use of these tests by applying them to an empirical example of estimating a vacancy duration model using some Finnish data from the 1989. In this particular case the implicit misspecification test based on the idea introduced by MacKinnon and Magee (1990) is quite successful in narrowing the type of misspecification present due to violation of the third order conditional moment restriction and is more powerful than the alternatives considered.

2. THE MACKINNON-MAGEE TRANSFORMATION IN WEIBULL DURATION MODELS

The following family of transformations introduced by MacKinnon and Magee (1990) is considered in this paper

$$g(\gamma y)/\gamma, \tag{1}$$

where the function g is monotonic and satisfies the following properties:

$$g(0) = 0, \tag{2}$$

$$g'(0) = 1, \tag{3}$$

$$g''(0) \neq 0. \tag{4}$$

If one allows for γ to vary in a suitable way, the above transformation (1) is homogeneous of degree one in y. This property of scale invariance is not shared by the Box-Cox transformation. In addition transformation (1) is applicable to variables y that can have negative values.[2] These features make it interesting to examine the properties of (1) in duration models, where logarithmic duration is considered. Property (4) is needed because otherwise the partial derivative of the loglikelihood function w.r.t. γ would be zero at the point $\gamma = 0$. This, however, rules out skew-symmetric functions g (see the example at the end of this section; for detailed discussion, see MacKinnon and Magee, 1990).

Consider the following Weibull duration model, where the observed duration t, $t \geq 0$, has the hazard function,

$$h(t; X^T\beta) = \alpha t^{\alpha-1}\exp\{-X^T\beta\}, \tag{5}$$

---

[2] Note that applying the Box-Cox transformation $t \to t^{(\gamma)} = (t^\gamma - 1)/\gamma$ in a Weibull duration model (see eq. (5)) produces a random variable with a Weibull distribution if $\gamma = 0$ and with a Gompertz distribution if $\gamma = 1$. However, the transformation allows only for monotonic variation in the hazard function, since $dt^{(\gamma)}/dt = t^{\gamma-1} > 0$ (see Lancaster, 1979, Sec. 3.5).

and the survival function

$$S(t; X^T\beta) = \exp\{-t^{\alpha} e^{-X^T\beta}\}. \tag{6}$$

Here X is a column vector of k independent variables and β is a vector of parameters to be estimated together with a shape parameter α. For the purposes of this study it is convenient to work with the variable y = log(t). It is well known that y follows the type I extreme value distribution, $y \sim EV(X^T\beta, \alpha)$, with the survival function

$$S(y; X^T\beta) = \exp\{-e^{\alpha y - X^T\beta}\}. \tag{7}$$

The variable, $u_k = \alpha y_k - X_k^T\beta$ is a normalized disturbance term in the sense that $E(u_k) = \psi(0)$, and $Var(u_k) = \psi'(0)$ and has a moment generating function, $\varphi(s) = \Gamma(1+s)$.[3] The integrated hazard, or alternatively the generalized residual of the model, $\epsilon_k = \exp(u_k)$, has a standard exponential distribution with moments $E(\epsilon_k^m) = \Gamma(1+m) = m!$.

In the sequel we consider the following generalized Weibull model, where the observed log-duration y is generated by

$$r(\gamma, y) \equiv g(\gamma y)/\gamma \sim EV(X^T\beta, \alpha). \tag{8}$$

One important feature of (2) in logarithmic duration models is that the probability of an survival lasting for at least one time period, i.e. P(y > 0) is independent of the transformation (1). On the other hand, Weibull models are scale invariant in the sense that changing the unit of time measurement changes only the value of the constant term in expression $X^T\beta$. In addition it turns out that the relevant score used to test for $\gamma = 0$ is independent on the time scale. Therefore the investigator is in principle able to choose the point zero freely in order to maximize the local power of the test which is dependent on g"(0).

---

[3] Notation Γ refers to the gamma function and ψ to the psi function, $\psi(x) = d\log\Gamma(1+x)/dx$. The higher order cumulants of u_k are obtainable through the derivatives of the psi function at zero, $\psi(0) = -\gamma_0$ (Euler's constant) and $Var(u_k) = \psi'(0) = \pi^2/6$, etc., see Abramovitz & Stegun (1970).

In the following we consider the score test for the null hypothesis that $\gamma = 0$. Under this null $r(0, y)$ reduces to y and the model (8) reduces to the ordinary Weibull model. This is easily verified by taking the appropriate limit of $r(\gamma, y)$ and using $g'(0) = 1$.

The survival function of the log-duration y is given by

$$S(y) \equiv P(\underline{y} \geq y) = P\big(r(\gamma, \underline{y}) \geq r(\gamma, y)\big) = \exp\left(-e^{\frac{\alpha g(\gamma y)}{\gamma} - X^T\beta}\right). \quad (9)$$

The corresponding hazard function w.r.t $y = \log(t)$ is given by

$$h(y) \equiv \alpha g'(\gamma y) \exp\left(\frac{\alpha g(\gamma y)}{\gamma} - X^T\beta\right). \quad (10)$$

The above expression shows that the MacKinnon-Magee transformation maintains the proportional hazard property present in Weibull duration models.

The derivative of the hazard function w.r.t t is given by

$$h'(t) \equiv \alpha\left[\gamma g''(\gamma y) + \alpha\left(g'(\gamma y)\right)^2 - g'(\gamma y)\right] t^{-2} \exp\left(\frac{\alpha g(\gamma y)}{\gamma} - X^T\beta\right), \quad (11)$$

where $y = \log(t)$.

Let $\gamma > 0$ with no loss of generality. The above expression indicates that if g is a concave function of y, the hazard function can be an initially increasing and a finally decreasing function of t. Similarly, if g is a convex function, the hazard function can be an initially decreasing and a finally increasing function of t. In the ordinary Weibull model, when $\gamma = 0$, the hazard function is always a monotonic function of t, and is strictly increasing if and only if $\alpha > 1$.[4]

Example

Consider the transformation

$$g = \kappa\,(S^{-1} \circ \Lambda), \tag{12}$$

where κ is a scaling constant to be determined later, S is the survival function of the standardized extreme value distribution, with $S^{-1}(x) = \log(-\log(x))$, and Λ is a given standardized survival function, with $\Lambda(0) = S(0) = e^{-1}$. In this case condition (2) is satisfied and (3) holds if

$$g'(0) = \kappa \frac{-\lambda(0)}{\log(\Lambda(0))\,\Lambda(0)} = \kappa\lambda(0)\,e = 1, \tag{13}$$

where $\lambda = -d\Lambda$. In order to satisfy (13) one must set $\kappa = [e\,\lambda(0)]^{-1} = f(0)/\lambda(0) = h_S(0)/h_\Lambda(0)$, i.e. the ratio of the hazard rates at zero. Similarly

$$g''(0) = \frac{-\kappa}{[\log(\Lambda(0))]^2}\,[h_\Lambda^2(0) + h_\Lambda'(0)\log(\Lambda(0))]. \tag{14}$$

This expression reveals that in order to have $g''(0) \neq 0$ one must have $h_\Lambda'(0) \neq h_\Lambda^2(0)$, i.e. the distribution Λ must be such that the corresponding density function has no mode at zero. In this case $g''(0) = \lambda'(0)/\lambda(0)$.

Similarly the survival function of the latent variable y is given by

$$\begin{aligned} P(\underline{y}>y) &= P\big(\alpha r(\gamma,\underline{y}) - X^T\beta > \alpha r(\gamma,y) - X^T\beta\big) \\ &= \exp\Big(-e^{\frac{\alpha g(\gamma y)}{\gamma} - X^T\beta}\Big) = \exp\big(-(-\log(\Lambda(\gamma y)))^{\alpha\kappa/\gamma} e^{-X^T\beta}\big). \end{aligned} \tag{15}$$

---

4 It is often reasonable to consider alternatives with non-monotonic duration dependence, for instance the jobmatching model by Jovanovic (1979) predicts that hazard for leaving a job is an initially increasing and a finally decreasing function of job tenure.

In other words, one has taken a random variable distributed according to Λ up to a scale factor γ and transformed it to get a Weibull random variable, with a proportional hazard term $X_k^T \beta$, and shape parameter $\kappa\alpha/\gamma$. The corresponding hazard function is in this case given by

$$\kappa\alpha\,[-\log(\Lambda(\gamma y))]^{\kappa\alpha/\gamma\,-1}\,h_\Lambda(\gamma y)\,e^{-X^T\beta}. \tag{16}$$

The example shows that setting $\gamma = \kappa\alpha$ ensures that the score test derived below has some power if one considers the alternative hypotheses characterized by a base line hazard function h_Λ in the above example. The function h_Λ is essentially arbitrary up to the condition $h'(0) \neq h^2(0)$.

3. SCORE TEST FOR H_0: $\gamma = 0$, WHEN CENSORING IS PRESENT

Consider the following censoring mechanism. The logarithmic duration y^* is not directly observable. Instead we observe y, where

$$y = \min\{y^*, c\}, \tag{17}$$

The censoring variable c is independent of y^*, and has a density function λ and survival function Λ.

The loglikelihood of an individual observation, (I_k, y_k, X_k), $k = 1, \ldots, n$, where I_k is an indicator function for a completed, uncensored duration, is given by

$$\mathcal{L}_k = \mathcal{L}(\gamma, \alpha, \beta \mid I_k, y_k, X_k)$$

$$= I_k\left[\log(h(y_k; X_k^T\beta)) + \log(S(y_k; X_k^T\beta))\right] + (1 - I_k)\log(S(y_k; X_k^t\beta))$$
$$+ I_k \log(\Lambda_k(y_k)) + (1 - I_k)\log(\lambda_k(y_k)).$$

Substituting (9) and (10) gives

$$\mathcal{L}_k = I_k\left[\log(\alpha) + \log(g'(\gamma y_k)) + \frac{\alpha g(\gamma y_k)}{\gamma} - X_k^T\beta\right]$$
$$- \exp\left\{\frac{\alpha g(\gamma y_k)}{\gamma} - X_k^T\beta\right\} + I_k \log(\Lambda_k(y_k)) + (1 - I_k)\log(\lambda_k(y_k)). \tag{18}$$

In the above formulation the censoring mechanism may vary between the observations. The coefficients $\Lambda_k(y_k)$, and $\lambda_k(y_k)$ may be regarded as nuisance parameters which do not depend on the direct objects of interest, i.e. (α, β, γ). Therefore it suffices to maximize that part of the loglikelihood which involves the parameters of interest.

The derivative of $\mathcal{L}_k$ w.r.t γ is

$$\frac{\partial \mathcal{L}_k}{\partial \gamma} = I_k \left[\frac{g''(\gamma y_k)\, y_k}{g'(\gamma y_k)} + \alpha \left[\frac{g'(\gamma y_k)\, y_k}{\gamma} - \frac{g(\gamma y_k)}{\gamma^2} \right] \right]$$

$$- \alpha \left[\frac{g'(\gamma y_k)\, y_k}{\gamma} - \frac{g(\gamma y_k)}{\gamma^2} \right] \exp\left[\frac{\alpha g(\gamma y_k)}{\gamma} - X_k^T \beta \right]. \tag{19}$$

Applying l'Hopital's rule to the term in the last brackets implies

$$\lim_{\gamma \to 0} \frac{g'(\gamma y_k)\, y_k}{\gamma} - \frac{g(\gamma y_k)}{\gamma^2} = \frac{1}{2} g''(0)\, y_k^2, \tag{20}$$

giving

$$\left[\frac{\partial \mathcal{L}_k}{\partial \gamma} \right]_{\gamma = 0} = g''(0) \left[I_k \left(y_k + \frac{\alpha y_k^2}{2} \right) - \frac{\alpha y_k^2}{2} \exp(\alpha y_k - X_k^T \beta) \right]. \tag{21}$$

The other derivatives of the loglikelihood function under $\gamma = 0$ are the same as in the ordinary Weibull model:

$$\left[\frac{\partial \mathcal{L}_k}{\partial \alpha} \right]_{\gamma = 0} = I_k \left[\frac{1}{\alpha} + y_k \right] - y_k \exp(\alpha y_k - X_k^T \beta), \tag{22 i}$$

$$\left[\frac{\partial \mathcal{L}_k}{\partial \beta} \right]_{\gamma = 0} = \left[-I_k + \exp(\alpha y_k - X_k^T \beta) \right] X_k. \tag{22 ii}$$

Under H_0 the maximum likelihood estimates of β and α in the ordinary Weibull duration model satisfy

$$\left[\frac{\partial \mathcal{L}}{\partial \beta} \right]_{\gamma = 0} \equiv \sum_{k=1}^{n} \left[\frac{\partial \mathcal{L}_k}{\partial \beta} \right]_{\gamma = 0} = 0 = \sum_{k=1}^{n} \left[\frac{\partial \mathcal{L}_k}{\partial \alpha} \right]_{\gamma = 0} \equiv \left[\frac{\partial \mathcal{L}}{\partial \alpha} \right]_{\gamma = 0}. \tag{23}$$

A score test for $\gamma = 0$ can always be interpreted as a test for

$$plim_{n \uparrow \infty} \left[\frac{1}{n}\frac{\partial \mathcal{L}}{\partial \gamma}\right]_{\gamma = 0} = 0. \tag{24}$$

Denoting the residuals of completed uncensored durations by $I_k u_k = \alpha I_k y_k - I_k X_k^T \beta$, one can easily show by partial integration that under $\gamma = 0$

$$\mathscr{E}(I_k e^{u_k}) = \int (1 - S_{u_k}(z))\, \lambda_{c_k}(z)\, dz - \int e^z S_{u_k}(z)\, \lambda_{c_k}(z)\, dz, \tag{25 i}$$

$$\mathscr{E}((1 - I_k)\, e^{u_k}) = \int e^z S_{u_k}(z)\, \lambda_{c_k}(z)\, dz. \tag{25 ii}$$

Giving

$$\mathscr{E}(e^{u_k}) = \int (1 - S_{u_k}(z))\, \lambda_{c_k}(z)\, dz = \mathscr{E}(I_k). \tag{26}$$

In Weibull models some additional conditional moment identities are given by the recursion formula

$$\mathscr{E}(u_k^m e^{u_k}) = m\mathscr{E}(I_k u_k^{m-1}) + \mathscr{E}(I_k u_k^m), \quad m \geq 1. \tag{27}$$

If the constant term is included among the regressors X_k, the corresponding sample moments always satisfy (26) by the first order conditions given by (22 ii). Formula (22 i) implies that (27) holds for m = 1. In the case of no censoring i.e. $I_k \equiv 1$, for all k, the above formulas reduce to the moment recursions of the standard extreme value distribution.

One can write

$$\frac{1}{n}\left[\frac{\partial \mathcal{L}}{\partial \gamma}\right]_{\gamma = 0} = \left[\frac{g''(0)}{\alpha n}\right] \left[\sum_{k=1}^{n} I_k \left(X_k^T\beta + u_k + \frac{1}{2}(X_k^T\beta + u_k)^2\right) - \frac{1}{2}(X_k^T\beta + u_k)^2 e^{u_k}\right]$$

$$= \left[\frac{-g''(0)}{\alpha}\right] \left[\sum_{k=1}^{n} \frac{1}{2n} (X_k^T\beta)^2 \left(e^{u_k} - I_k\right)\right. \tag{28 i}$$

$$+ \sum_{k=1}^{n} \frac{X_k^T\beta}{n}\left(u_k e^{u_k} - I_k(1 + u_k)\right) \qquad \textbf{(28 ii)}$$

$$+ \sum_{k=1}^{n} \frac{1}{2n}\left(u_k^2 e^{u_k} - I_k(u_k^2 + 2u_k)\right)\Bigg]. \qquad \textbf{(28 iii)}$$

Using the properties (26) & (27), one easily shows that each of terms in (28 i-iii) have a probability limit of zero, as n ↑ ∞. In addition the above expression is independent of the time scale and therefore the investigator is in principle able to choose freely the point zero in logarithmic scale in order to maximize the local power of the test which is dependent on $g''(0)$.[5]

If one tests an ordinary Weibull model against an alternative defined by (8), one tests simultaneously for three restrictions that affect conditional moments. The first of these, (28 i), may in this case also be seen as a form of regression test, where one tests that there is no correlation between $(X_k^T\beta)^2$ and the generalized residual $\epsilon_k = \exp(u_k)$ when the latter is corrected for its expectation under the censoring mechanism (17). This is in analogy with what a popular RESET test by Ramsey (1969) tests for.

To see this consider instead of (8) the class of models

$$\alpha y_k = \frac{g(\delta X_k^T\beta)}{\delta} + u_k, \qquad \textbf{(29)}$$

where the MacKinnon-Magee transformation has been applied to the location term, and test for $\delta = 0$ in (29). As before u_k has the standardized extreme value distribution.

[5] To see this consider changing the time scale by dividing the time t by, say μ. Now all other parameters of the model stay unchanged expect for the constant term in $X_k^T\beta$, say β_0. The new estimated constant term is $\beta_0 + (1/\alpha)\log(\mu)$. Inserting this to (28 i)- (28 ii) gives the same value for the score as earlier because of (22 ii), (26), and (27) for m = 1.

Formulate the loglikelihood of an individual observation, (I_k, y_k, X_k), $k = 1, \ldots, n$, where I_k is an indicator function for a completed, uncensored duration. Then, in analogy with (21) the derivative of $\mathcal{L}_k$ w.r.t δ, under $\delta = 0$ is given by

$$\left[\frac{\partial \mathcal{L}_k}{\partial \delta}\right]_{\delta=0} = \frac{g''(0)\,(X_k^T\beta)^2}{2}\left(e^{u_k} - I_k\right). \tag{30}$$

The other derivatives of the loglikelihood function under $\delta = 0$ are the same as in the ordinary Weibull model (22 i & ii).

The score test for $\delta = 0$ is easily seen to be asymptotically equivalent to a form of RESET test in which one includes an additional explanatory variable $(X_k^T\beta)^2$ in the model. In duration models this form of the RESET test tests for $\delta = 0$ in a Weibull model where the proportional hazard component is written as

$$X_k^T\beta + \delta(X_k^T\tilde{\beta})^2, \tag{31}$$

where $\beta^{\sim}$ is the maximum likelihood estimate of β in a preliminary Weibull model, with $\delta = 0$.

The second restriction, (28 ii), is that there is no correlation between $(X_k^T\beta)$ and zero mean 'disturbances' which are implicitely defined by the conditional moment restriction (27) with $m = 1$.

The third restriction, (28 iii), is that the conditional moment restriction (27) holds for $m = 2$. The last two restictions will be considered in a more detailed way in the next section of the paper. In particular, we outline their relation to sample heterogeneity and duration dependence in Weibull models.

In summary, if the data were generated by (8) with $\gamma \neq 0$, and one estimated an ordinary Weibull model, the estimated conditional moments would be expected to suffer from problems which would affect their theoretical restictions (26)

& (27) in a way not allowed by the ordinary Weibull model. These observations, made originally by MacKinnon and Magee (1990) in ordinary regression models, carry over to the Weibull model, albeit in a modified form.

The RESET test will have some power against $\gamma \neq 0$ in (8) because of (28 i). If the fit of the Weibull model is "very good" by which it is meant that α is large relative to the variation of the location component $X_k^T b = (1/\alpha)(X_k^T\beta)$,[6] then any violation to the condition that the probability limits of (28 ii) and (28 iii) are zero, will contribute relatively little to the score test. If the fit of the Weibull model is poor, however, RESET test will have low power compared to the score test against the above specific alternative hypothesis (8). This follows from the observation that as $\alpha \uparrow \infty$ then typically $(X_k^T\beta)^2 \uparrow \infty$, and the first term (28 i) will dominate over the other two terms (28 ii) and (28 iii).

The easiest way to calculate the score test for $\gamma = 0$ is to replace the information matrix of the parameter estimates with its finite sample approximation by the "Outer Product of the Gradient," or OPG,

$$plim_{n\uparrow\infty} \left[\left[\frac{\partial \mathcal{L}}{\partial \theta}\right]\left[\frac{\partial \mathcal{L}}{\partial \theta}\right]^T\right]_{\gamma=0} = \mathscr{E}\left[-\frac{\partial^2 \mathcal{L}}{\partial \theta \partial \theta^T}\right]_{\gamma=0} \equiv \mathfrak{I}. \tag{32}$$

The OPG test statistic, popularized by Godfrey and Wickens (1981), can be computed as n minus the sum of squared residuals, or nR^2, from the artificial linear regression

$$\iota = \left[\frac{\partial \mathcal{L}}{\partial \theta}\right] c + remainder, \tag{33}$$

where ι is an n-vector of ones and the regressor matrix is an $n \times p$ matrix of the derivatives of the loglikelihood of each of the observations evaluated at the restricted ordinary Weibull model estimates. Here $n = k + 2$, and the cor-

[6] In Weibull model this implies a rapidly increasing hazard function with the shape parameter α large relative to $X_k^T b$.

responding components are given in the formulae (21) -(22). Note that the actual value of g"(0) has no effect on the test statistic and might as well be omitted.

Alternatively one can calculate the second derivatives of the loglikelihood

$$\left[\frac{\partial^2 \mathcal{L}_k}{\partial\gamma\partial\alpha}\right]_{\gamma=0} = \frac{g''(0)}{2}\left(I_k\, y_k^2 - (y_k^2 + \alpha y_k^3)\, e^{u_k}\right), \tag{34}$$

$$\left[\frac{\partial^2 \mathcal{L}_k}{\partial\gamma\partial\beta}\right]_{\gamma=0} = \frac{g''(0)\,\alpha y_k^2 e^{u_k}}{2} X_k. \tag{35}$$

Partitioning the information matrix as

$$\Im = \begin{bmatrix} \iota_{\gamma\gamma} & \Im_{\gamma 2} \\ \Im_{2\gamma} & \Im_{22} \end{bmatrix}, \tag{36}$$

one can write the score statistic LM

$$LM = \frac{\left[\frac{\partial \mathcal{L}}{\partial\gamma}\right]^2_{\gamma=0}}{\iota_{\gamma\gamma} - \Im_{\gamma 2}\Im_{22}^{-1}\Im_{2\gamma}}$$

$$= \frac{\left[\sum_{k=1}^{n} I_k\left(X_k^T\beta + u_k + \frac{1}{2}(X_k^T\beta + u_k)^2\right) - \frac{1}{2}(X_k^T\beta + u_k)^2 e^{u_k}\right]^2}{v_{\gamma\gamma} - 1_\gamma^T V 1_\gamma}, \tag{37}$$

where

$$v_{\gamma\gamma} = \frac{\alpha^2\ \iota_{\gamma\gamma}}{g''(0)^2} = \sum_{k=1}^{n} \mathcal{E}\left[I_k\left(X_k^T\beta + u_k + \frac{1}{2}(X_k^T\beta + u_k)^2\right) - \frac{1}{2}(X_k^T\beta + u_k)\, e^{u_k}\right]^2, \tag{38 i}$$

$$l_\gamma = \frac{\alpha}{g''(0)} \Im_{2\gamma} = \frac{1}{2} \begin{bmatrix} \alpha \sum_{k=1}^{n} \mathscr{E}\left((y_k^2 + \alpha y_k^3)\, e^{u_k} - I_k y_k^2 \right) \\ -\sum_{k=1}^{n} \mathscr{E}\left((\alpha y_k)^2 e^{u_k} x_k \right) \end{bmatrix}, \qquad \textbf{(38 ii)}$$

and V is the estimated covariance matrix of the maximum likelihood estimators of $(\alpha, \beta^T)^T$ in the ordinary Weibull duration model. The formulae for V, $v_{\gamma\gamma}$ and l_γ are given in the appendix. They involve conditional moments of uncensored observations up to the fourth degree and are straightforward, albeit relatively tedious to derive.

The above score test based on the MacKinnon-Magee family of transformations has some power against a wide range of alternative models. This is attributable to the way in which information on the non-linear model of the location is confouded with information concerning the distribution of disturbances, affected by e.g. sample heterogeneity. The lack of fit detected may be due to the location model, or sample heterogeneity, or both. This feature of model (1) is common to all transformation families, eg. Box-Cox, that affect only the values of the dependent variable.

To cater for the above situation two score tests are considered below. These two tests specialise in testing sample-heterogeneity in the constant term of the proportional hazard component, and such duration dependence in the shape parameter α of the model, which is related to the location component, respectively. If any of these tests rejects the null substantially more emphatically than the score test for $\gamma = 0$ in (8), that would suggest that (8) is not the appropriate model and modifications in modelling the proportional hazard component, or sample heterogeneity would be necessary.

4. SCORE TESTS FOR THE SAMPLE HETEROGENEITY

The standard procedure to introduce sample heterogeneity to Weibull duration model is through the constant term in $X_k^T\beta$ e.g. Burdett et. al (1985) and Lancaster (1985). One introduces a latent heterogeneity inducing variable η along with an extreme value random disturbance term u^* with the formula

$$u_k = u_k^* + \eta_k, \tag{39}$$

where η has the density function ψ, and mean zero and variance σ^2. The loglikelihood is obtained by replacing the density and survival functions by

$$\int f(u_k | \eta)\, \psi(\eta)\, d\eta \;, \; and \; \int S(u_k | \eta)\, \psi(\eta)\, d\eta. \tag{40}$$

For completeness one should also replace the nuisance parameters Λ, λ by

$$\int \lambda(u_k | \eta)\, \psi(\eta)\, d\eta \;, \; and \; \int \Lambda(u_k | \eta)\, \psi(\eta)\, d\eta. \tag{41}$$

Next the conditional density function and survival function are replaced by their second order Taylor expansions w.r.t. η, at point $\eta = 0$. We get

$$f(u | \eta) \approx f(u | 0) + \eta \left[\frac{\partial f(u | \eta)}{\partial \eta}\right]_{\eta=0} + \frac{\eta^2}{2}\left[\frac{\partial^2 f(u | \eta)}{\partial \eta^2}\right]_{\eta=0}, \tag{42 i}$$

$$S(u | \eta) \approx S(u | 0) + \eta \left[\frac{\partial S(u | \eta)}{\partial \eta}\right]_{\eta=0} + \frac{\eta^2}{2}\left[\frac{\partial^2 S(u | \eta)}{\partial \eta^2}\right]_{\eta=0}. \tag{42 ii}$$

Substituting the relevant derivatives and integrating w.r.t. η gives in Weibull model[7]

$$f(u^*) \approx f(u | 0)\left[1 + \frac{\sigma^2}{2}(1 - 3e^u + e^{2u})\right], \; and \tag{43 i}$$

[7] Note that the conditional density of u given η, $u|\eta$, at x is equal to the density of u^* at the point $x-\eta$.

$$S(u^*) \approx S(u\,|\,0)\left[1 + \frac{\sigma^2}{2}\,(e^{2u} - e^{u})\right]. \qquad \textbf{(43 ii)}$$

The derivative of $\mathcal{L}_k$ w.r.t σ^2, under $\sigma^2 = 0$ is given by

$$\left[\frac{\partial \mathcal{L}_k}{\partial \sigma^2}\right]_{\sigma^2=0} = \frac{1}{2}\left[I_k\,(1 - 3e^{u_k} + e^{2u_k}) + (1 - I_k)\,(e^{2u_k} - e^{u_k})\right]. \qquad \textbf{(44)}$$

We get mean score

$$\frac{1}{n}\left[\frac{\partial \mathcal{L}}{\partial \sigma^2}\right]_{\sigma^2=0} = \frac{1}{2n}\sum_{k=1}^{n}\left[e^{2u_k} - 2I_k e^{u_k}\right], \qquad \textbf{(45)}$$

because sample moments satisfy (26).

The other derivatives of the loglikelihood function under $\sigma^2 = 0$ are the same as in the ordinary Weibull model (22 i & ii). The score test for $\sigma^2 = 0$ is easily seen to be asymptotically equivalent to a conditional moment test in which one compares the second moment of the generalized residual $\epsilon = \exp(u)$ with its theoretical value under H_0. The expectation can be calculated by Taylor-series expansion and applying (27) to give

$$\mathscr{E}(e^{2u_k}) = \sum_{m=0}^{\infty}\frac{\mathscr{E}(u_k^m e^{u_k})}{m!} = 2\mathscr{E}(I_k e^{u_k}). \qquad \textbf{(46)}$$

More generally one can show that the moment generating function of u_k is

$$\varphi(r) = \mathscr{E}(e^{r u_k}) = \Gamma(r+1)\,\mathscr{E}(I_k e^{u_k}). \qquad \textbf{(47)}$$

Therefore score test can be seen as testing whether the second sample moment of generalized residuals is equal to its theoretical counterpart $\varphi(2)$.

On the other hand, part (28 iii) of the score test which is based on the MacKinnon-Magee family of transformations effectively tests for the sample moment restriction which corresponds to the value of $\varphi''(1)$. Note that by construction the Weibull model errors u_k always satisfy the theoretical values $\varphi(1)$ and $\varphi'(1)$ by (26)-(27). The above considerations show that these two tests are related though they refer to different type of moment restrictions.[8] A general sequence of conditional moment restrictions are obtainable through the recursion

$$\mathscr{E}(u_k^m e^{ru_k}) = r\mathscr{E}(I_k u_k^m e^{(r-1)u_k}) + m\mathscr{E}(I_k u_k^{m-1} e^{(r-1)u_k}), \qquad m \geq 1, \quad r \geq 1. \tag{48}$$

The above recursion effectively characterizes the theoretical counterparts of all the derivatives $\varphi^{(r)}(m)$, under the null, where $r,m = 1,2,\ldots,\infty$.

Another form of sample heterogeneity which is briefly discussed involves the Weibull shape parameter α. Consider a Weibull model, where the shape parameter is replaced by

$$\alpha = \alpha_0 g'(\delta X^T \beta), \tag{49}$$

where g is a monotonic function which satisfies properties (2)-(4). In the above model the shape parameter of the duration depends on the proportional hazard component through the derivative of function g.

Possible functional forms for g', which define locally equivalent alternative models, include the exponential function, e^x, giving

$$g'(\delta X^T \beta) = \exp(\delta X^T \beta), \tag{50 i}$$

considered in testing for heteroscedasticity in regression models, or the affine function, $1 + x$, with $x > -1$, giving

[8] An intuitive reason for not obtaining complete correspondence between the two restrictions is that the above sample heterogeneity test is derivable as an information matrix test, see Chesher (1984), which involves the second derivatives of the log likelihood.

$$g'(\delta X^T\beta) = 1 + \delta X^T\beta. \tag{50 ii}$$

In the sequel a score test for the hypothesis $\delta = 0$ is considered. The motivation for basing the test on model (49) is that duration dependence which is in Weibull models characterized by α may be related to values of some important explanatory variables in the model. These variables correlate with the location, $X_k^T\beta$, and the test for $\delta = 0$ is likely to have some power even in the case where duration dependence is related to the values of a single important explanatory variable. The above idea was proposed originally for testing heteroscedasticity in regression models by Anscombe (1961), see also Breusch & Pagan (1979).

Using (18) the derivative of $\mathcal{L}_k$ w.r.t δ, under $\delta = 0$ is given by

$$\left[\frac{\partial \mathcal{L}_k}{\partial \delta}\right]_{\delta=0} = \alpha_0 g''(0)\, X_k^T\beta \left[I_k \left(\frac{1}{\alpha_0} + y_k\right) - y_k \exp(\alpha y_k - X_k^T\beta) \right], \tag{51}$$

Giving mean score

$$\frac{1}{n}\left[\frac{\partial \mathcal{L}}{\partial \delta}\right]_{\delta=0} = -g''(0) \left[\sum_{k=1}^{n} \frac{(X_k^T\beta)^2}{n} \left(e^{u_k} - I_k\right) + \sum_{k=1}^{n} \frac{X_k^T\beta}{n} \left(u_k e^{u_k} - I_k(1 + u_k)\right) \right]. \tag{52}$$

The other derivatives of the loglikelihood function under $\delta = 0$ are the same as in the ordinary Weibull model (22 i & ii). The score test for $\delta = 0$ is equivalent with the first two terms in the expression characterizing the score test derived earlier (28 i-ii) with the exception that the second term should be split in half.

5. EXTENSION TO COMPETING RISKS FRAMEWORK

In the present case the duration may end in m alternative ways which are here called exit channels. Both the length of duration and the label of the corresponding exit channel are observed. The standard competing risks model for duration is obtained by defining m mutually independent random variates y_j^*, $j = 1, \ldots, m$, and setting the observed duration y,

$$y \equiv \min\{y_1^*, \ldots, y_m^*, c\}, \tag{53}$$

The censoring variable c is independent of all y^*'s, and has density function λ and survival function Λ.

In the specific case considered here the log-durations y_j^*, $j = 1, \ldots, m$, are generated by applying a set of MacKinnon-Magee transformations g_j, separately to each of them to get m random variables following the extreme value distribution. These variates are defined by (8) with shape parameters α_j and location terms where the latter are linear combinations of the explanatory variables X(j) with coefficients β_j.

Define hazard functions h_j and survival functions S_j for each individual exit channel. Now the loglikelihood of an individual observation, $(y_k, I_k(j) \mid X_k(j);\ j = 1, \ldots, m)$, $k = 1, \ldots, n$, where $I_k(j)$ is an indicator function for a completed, uncensored duration through exit channel j, is given by

$$\mathcal{L}_k \equiv \mathcal{L}(\gamma, \alpha_j, \beta_j \mid I_k(j), y_k, X_k(j))$$

$$= \sum_{j=1}^{m} I_k(j) \left[\log(h_j(y_k; X_k(j)^T\beta_j)) + \sum_{i=1}^{m} \log(S_i(y_k; X_k(i)^T\beta_i)) + \log(\Lambda_k(y_k)) \right]$$

$$+ \left(1 - \sum_{j=1}^{m} I_k(j)\right) \left[\sum_{i=1}^{m} \log(S_i(Y_k; X_k(i)^T\beta_i)) + \log(\lambda_k(y_k)) \right]. \tag{54}$$

Rearranging the terms, and substituting for the functional forms for hazard and survival functions shows that

$$\mathscr{L}_k = \sum_{j=1}^{m} \mathscr{L}_k(j) \text{ , where} \tag{55}$$

$$\mathscr{L}_k(j) = I_k(j) \left[\log(\alpha_j) + \log(g_j'(\gamma_j y_k)) + \frac{\alpha_j g_j(\gamma_j y_k)}{\gamma_j} - X_k(j)^T \beta_j \right]$$
$$- \exp\left[\frac{\alpha_j g_j(\gamma_j y_k)}{\gamma_j} - X_k(j)^T \beta_j \right] + I_k(j) \log(\Lambda_k(y_k)) + \left(\frac{1}{m} - I_k(j)\right) \log(\lambda_k(y_k)) .$$

The above expression shows that each set of the parameters of interest, i.e. $(\alpha_j, \beta_j, \gamma_j)$, can be estimated by maximizing that part of the loglikelihood which involves the parameters of interest, i.e. $\mathscr{L}(j) = \Sigma\, \mathscr{L}_k(j)$.[9] In particular the equations for first order conditions (22 i & ii) and the scores w.r.t. γ's (21) stay the same expect for the substitution for the indices j in the indicator variables, I_k, and parameters, etc. Considering each likelihood component separately we can derive m score tests for misspecification, one per each exit channel.[10]

[9] The coefficients $\Lambda_k(y_k)$, and $\lambda_k(y_k)$ may be regarded as nuisance parameters which do not depend on the parameters of interest, i.e. (α, β, γ)'s. Alternatively the censoring variable may have a fixed distribution with some estimable parameters.

[10] Note that equations (25 i)-(26) have to be modified to allow for an 'augmented censoring mechanism' which includes the other exit channels than the one we are explicitely concerned with. With these modifications all the moment restrictions given earlier in the paper remain true if one substitutes $I_k(j)$ for I_k.

6. AN EMPIRICAL EXAMPLE

The implicit misspecification test derived above is applied to an empirical example analyzed earlier by the second author (Rantala, 1991). The data concerns vacancies reported to employment offices in the province of Uusimaa, Finland in 1989.[11] Vacancies were for employees, with upper secondary education in science and technology programs.

In the data (2531 observations) two exit channels are considered, the first is recruitment through employment offices (30 per cent of the cases) and the second, recruitment through other channels (45 per cent of the cases). The remaining vacancies were withdrawn by the employers from the employment office registers. The latter are here considered as censored observations. In duration models the following explanatory variables were used, dummy variables accounting for occupation (11 categories), industry (6 categories), type of work (permanent/temporary), worktime (regular/shift work), and sub-region (3 categories). In addition U/V - ratios, i.e. the number of unemployed divided by the number of vacancies, were calculated separately for each sub-region and occupation at the end of the duration in question. The resulting continuous variable was used in the analysis to capture the effect of local labour market conditions.

In the analysis, we used twenty explanatory variables and 2531 observations, 25 per cent of which were considered as censored. In the bulk of the data recruitment was through the second channel and the mean duration in the data was 46 days. A competing risks Weibull duration model with (44 parameters) was subjected to the score tests for misspecification derived in the paper. Test results are given in Table 1, separately for the two exit channels considered.

[11] In the data complete durations are recorded for all vacancies. The province of Uusimaa is the most populated in Finland and the capital, Helsinki, is situated there.

TABLE 1

PERFORMANCE OF THE WEIBULL MODEL FOR VACANCIES[12]

Test	Mean score	Test statistic χ^2-form	t-form
Channel 1			
$LM(\gamma_1)$	-0.0998	119.20	-11.13
LM_1	0.0065	7.25	
LM_2	0.0067	0.69[13]	
LM_3	-0.1131	119.36[13]	
RESET		7.25	2.68
HETGEN		31.76	-5.65
HETDUR		10.23	3.19
Channel 2			
$LM(\gamma_2)$	-0.1526	208.18	-14.99
LM_1	-0.0021	4.73	
LM_2	0.0027	0.28[13]	
LM_3	-0.1532	228.12[13]	
RESET		4.73	-2.17
HETGEN		93.80	-9.82
HETDUR		0.83	-0.91

The score test statistics are asymptotically $\chi^2(1)$ when the underlying competing risks (Weibull) model is actually generating the data. The first row gives the score test for misspecification. The values LM_1, LM_2, and LM_3 correspond to the decomposition of the score, given in eqs. (28 i-iii), respectively, and indicate the contributions of the misspecification of the proportional hazard component, and two conditional moment restrictions in the implied disturbance distribution corrected for censoring. The RESET row gives the test statistic for misspecification in the functional form of the proportional hazard component, with score given in eq. (30). The last two rows give the test statistics for sample heterogeneity, with the scores given in eqs (45) and (52), respectively. In their calculation the outer product forms were used, and corresponding asymptotically equivalent statistics based on Student's t-distribution are given in the last column.

---

[12] A complete set of results obtainable by request.

[13] These are pseudo-statistics with the values calculated by a mechanical application of the OP-method, and have by themselves no interpretation as score statistics.

The highly significant test statistics LM(γ_1) and LM(γ_2) clearly indicate that the model is misspecified. Interestingly enough this rejection does not seem to have been markedly influenced by the misspecification of the location or alternatively the proportional hazard component, rows RESET, nor the second moment restriction (28 ii), row LM_2. In contrast the row LM_3 suggests that the violation of the third moment restriction (28 iii) is the main underlying problem. This restriction is closely related to sample heterogeneity as the test statistic reported in row HETGEN confirms. Here our test statistic and particularly the restiction (28 iii) seems to be more powerful in detecting misspecification than the heterogeneity test, which is derivable as an information matrix test (Chesher, 1984).

An additional interesting feature of the example is that the decomposition of the score (first column in Table 1) gives some useful information about the possible source of misspecification including the information in the other score-tests taylored to test separately for the misspecification of the proportional hazard and such sample heterogeneity which is related to the constant term and the scale (duration shape) parameter in the model (rows HETGEN and HETDUR).[14]

[14] In the original analysis of the data the poor fit of the Weibull model was detected by using the generalized gamma distribution with an additional parameter. This family nests both extreme value distributed and normally distributed disturbances in logarithmic duration.

7. CONCLUSION

In the paper the transformation family introduced by MacKinnon and Magee has been applied in Weibull duration framework to derive a score test for misspecification with one degree of freedom. The test statistic is found to be sensitive for violation of three conditional moment restrictions in the implied disturbance distribution of the model. The test is compared with a form of the RESET-test as well as with score tests which also have one degree of freedom and test for sample heterogeneity in the location component and such duration dependence which is related to the location component. Finally, an empirical example has been presented in which the derived test clearly indicates violation of the third conditional moments restriction and simultaneously accounts for the results given by the other two score tests mentioned above. Furthermore, our test statistic seems to be more powerful than the alternatives considered. Because the test statistic is easy to implement, it is recommended that it should be routinely calculated and used in model diagnostics in applications based on the Weibull model.[15]

Use of specification tests, such as the one developed in the paper, to examine model adequacy are best seen as a complementary rather than an alternative tool to other diagnostic methods such as informal residual plots, which are well established in the literature, eg. Aitkin & Clayton (1981), Lancaster (1985) & (1990), Lancaster & Chesher (1985), and Kiefer (1985). Similar considerations apply with respect to estimation methods with less reliance on distributional assumptions, such as semi- or nonparametric estimation of the base line hazard function, based on the Kaplan-Meier estimate of the survival function, eg. Lancaster (1990), Sec. 9.4.

Diagnostic methods may, however, result in the recognition

[15] It is relatively straightforward to derive the corresponding transformation test statistic for other parametric families such as lognormal, or generalized gamma distributions.

of important phenomena that might otherwise have gone undetected if one had relied exclusively on robust methods. In fact indentification of such phenomena that results in a marked lack of fit may have at least equal scientific importance than the analysis of the bulk of data.

REFERENCES

Abramowitz, M. and I.A. Stegun, "Handbook of Mathematical Functions", **Dover Publications**, New York (1970).

Aitkin, M. and D. Clayton, "The Fitting of Exponential, Weibull and Extreme Value Distributions to Complex Censored Survival Data using GLIM", **Applied Statistics**, 29 (1980), 156-163.

Anscombe, F.J., "Examination of Residuals", **Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability**, 4 (1961), 1-36.

Box, G.E.P. and D.R. Cox, "An Analysis of Transformations", **Journal of the Royal Statistical Society, Series B,** 26 (1964), 211-252.

Breusch, T.S. and A.R. Pagan, "A Simple Test for Heteroscedasticity and Random Coefficient Variation", **Econometrica**, 47 (1979), 1287-1294.

Burdett, K., N.M. Kiefer & S. Sharma, "Layoffs and Duration Dependence in a Model of Turnover", **Journal of Econometrics**, 28 (1985), 51-69.

Chesher, A., "Testing for Neglected Heterogeneity", **Econometrica**, 52 (1984), 865-872.

Hausman, J., "Specification Tests in Econometrics", **Econometrica**, 46 (1978), 1251-1271.

Kiefer, N.M., "Specification Diagnostics based on Laguerre Alternatives for Econometric Models of Duration", **Journal of Econometrics**, 28 (1985), 135-154.

MacKinnon, J.G. and L. Magee, "Transforming the Dependent Variable in Regression Models", **International Economic Review**, 31 (1990), 315-339.

Lancaster, T., "Generalised Residuals and Heterogenous Duration Models -with Applications to the Weibull Model", **Journal of Econometrics**, 28 (1985), 113-126.

Lancaster, T., "The Econometric Analysis of Transition Data", **Cambridge University Press**, Cambridge (1990).

Lancaster, T. & A.D. Chesher, "Residuals, Tests and Plots with a Job Matching Application", **Annales de L'INSEE** 59/60 (1985), 47-70.

Ramsey, J., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis", **Journal of the Royal Statistical Society, Series B**, 31 (1969), 350-371.

Rantala, J. "Vacancy Duration: A Competing Channels Model" (in Finnish), **Government Institute for Economic Research:** Discussion Paper 24 (1991).

APPENDIX

Below we give the necessary formulae for calculating the variance of the score, under H_0. These are derived by using the equations for the conditional moments restrictions in the Weibull model (26), (27), and (48). First, one can write

$$v_{\gamma\gamma} = \frac{\alpha^2 \iota_{\gamma\gamma}}{g''(0)^2} = \sum_{k=1}^{n} \mathscr{E}\left[I_k\left(X_k^T\beta + u_k + \frac{1}{2}(X_k^T\beta + u_k)^2\right) - \frac{1}{2}(X_k^T\beta + u_k)^2 e^{u_k}\right]^2,$$

$$= \sum_{k=1}^{n} \mathscr{E}\left[I_k\left((X_k^T\beta + u_k)^2 + (X_k^T\beta + u_k)^3 + \frac{1}{4}(X_k^T\beta + u_k)^4\right)\right]$$

$$- \sum_{k=1}^{n} \mathscr{E}\left[I_k\left((X_k^T\beta + u_k)^3 + \frac{1}{2}(X_k^T\beta + u_k)^4\right)e^{u_k}\right]$$

$$+ \sum_{k=1}^{n} \mathscr{E}\left[\frac{1}{4}(X_k^T\beta + u_k)^4 e^{2u_k}\right]$$

By (48) the last two terms cancel out, leaving

$$v_{\gamma\gamma} = \sum_{k=1}^{n} \mathscr{E}\left[I_k\left((X_k^T\beta + u_k) + \frac{1}{2}(X_k^T\beta + u_k)^2\right)^2\right]$$

$$= \sum_{k=1}^{n} \mathscr{E}\left[I_k(X_k^T\beta + u_k)^2 + \frac{1}{4}(X_k^T\beta + u_k)^4 e^{u_k}\right], \qquad \textbf{(A 1)}$$

by (27).

Similarly,

$$\mathit{1}_\gamma = \frac{\alpha}{g''(0)}\mathfrak{I}_{2\gamma} = \frac{1}{2}\begin{bmatrix} \alpha\sum_{k=1}^{n} \mathscr{E}\left((y_k^2 + \alpha y_k^3)\, e^{u_k} - I_k y_k^2\right) \\ -\sum_{k=1}^{n} \mathscr{E}\left((\alpha y_k)^2 e^{u_k} X_k\right) \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} \frac{1}{\alpha}\sum_{k=1}^{n} \mathscr{E}\left(((X_k^T\beta + u_k)^3 e^{u_k} + 2\,I_k(X_k^T\beta + u_k)\right) \\ -\sum_{k=1}^{n} \mathscr{E}\left((X_k^T\beta + u_k)^2 e^{u_k}\right)X_k \end{bmatrix}$$

$$= \frac{1}{2}\left[\begin{array}{c} \dfrac{1}{\alpha}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k\Big((X_k^T\beta + u_k)^3 + 3\,(X_k^T\beta + u_k)^2 + 2\,(X_k^T\beta + u_k)\Big)\Big) \\ -\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k\,\Big(X_k^T\beta + u_k)^2 + 2\,(X_k^T\beta + u_k)\Big)\Big)X_k \end{array}\right]. \qquad \textbf{(A 2)}$$

Further,

$$V^{-1} = -\begin{bmatrix} \mathcal{E}\dfrac{\partial^2\mathcal{L}}{\partial\alpha\,\partial\alpha} & \mathcal{E}\dfrac{\partial^2\mathcal{L}}{\partial\alpha\,\partial\beta^T} \\ \mathcal{E}\dfrac{\partial^2\mathcal{L}}{\partial\beta\,\partial\alpha} & \mathcal{E}\dfrac{\partial^2\mathcal{L}}{\partial\beta\,\partial\beta^T} \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{1}{\alpha^2}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k + (X_k^T\beta + u_k)^2 e^{u_k}\Big) & -\dfrac{1}{\alpha}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big((X_k^T\beta + u_k)\,e^{u_k}\Big)X_k^T \\ -\dfrac{1}{\alpha}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big((X_k^T\beta + u_k)\,e^{u_k}\Big)X_k & \displaystyle\sum_{k=1}^{n} \mathcal{E}(e^{u_k})\,X_k X_k^T \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{1}{\alpha^2}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k(1 + X_k^T\beta + u_k)^2\Big) & -\dfrac{1}{\alpha}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k(1 + X_k^T\beta + u_k)\Big)X_k^T \\ -\dfrac{1}{\alpha}\displaystyle\sum_{k=1}^{n} \mathcal{E}\Big(I_k(1 + X_k^T\beta + u_k)\Big)X_k & \displaystyle\sum_{k=1}^{n} \mathcal{E}(I_k)\,X_k X_k^T \end{bmatrix}. \qquad \textbf{(A 3)}$$

After collecting the terms in (A1)-(A3) we can write in terms of uncensored observations

$$\begin{aligned} v_{\gamma\gamma} - 1^T V 1 = \frac{1}{4}\Big[&A_4 + 4A_3 + 4A_2 \\ &- (B_2+2B_1)^T(M^{-1}+M^{-1}(B_1+B_0)D^{-1}(B_1+B_0)^T M^{-1})(B_2+2B_1) \\ &+ 2(B_2+2B_1)^T M^{-1}(B_1+B_0)D^{-1}(A_3+3A_2+2A_1) \\ &- (A_3+3A_2+2A_1)^T D^{-1}(A_3+3A_2+2A_1)\Big], \end{aligned} \qquad \textbf{(A 4)}$$

where

$$A_m = \sum_{k=1}^{n} \mathcal{E}\Big(I_k(X_k^T\beta + u_k)^m\Big),$$

$$B_m = \sum_{k=1}^{n} \mathcal{E}\Big(I_k(X_k^T\beta + u_k)^m\Big)X_k\,,$$

$$M = \sum_{k=1}^{n} \mathscr{E}(I_k)\, X_k X_k^T \,,$$

$$D = (A_2 + 2A_1 + A_0) - (B_1 + B_0)^T M^{-1} (B_1 + B_0) \,.$$

Finally, in the case of no censoring i.e. $I_k \equiv 1$, one may calculate

$$\mathscr{E}(X_k^T\beta + u_k)^m = \sum_{\nu=0}^{m} \binom{m}{\nu} (X_k^T\beta + \psi(0))^{m-\nu}\, \mathscr{E}(u_k - \psi(0))^\nu$$

$$= (X_k^T\beta + \psi(0))^m + \sum_{\nu=2}^{m} \binom{m}{\nu} (X_k^T\beta + \psi(0))^{m-\nu} \psi^{(\nu-1)}(0) \,, \qquad \textbf{(A 5)}$$

for $m \geq 1$. Note that

$$\psi^{(\nu)}(0) = (-1)^{\nu+1} \nu! \sum_{k=0}^{\infty} (1+k)^{-\nu-1}, \qquad \textit{for } \nu \geq 1.$$