Augustin:

Correcting for measurement error in parametric duration models by quasi-likelihood

Sonderforschungsbereich 386, Paper 157 (1999)

Online unter: http://epub.ub.uni-muenchen.de/

Correcting for measurement error in parametric duration models by quasi-likelihood

Thomas Augustin

Department of Statistics
Ludwig-Maximilians University of Munich
Akademiestr. 1
D-80799 Munich
thomas@stat.uni-muenchen.de+

Abstract

In regression models for duration data it is usually implicitly assumed that all variables are measured and operationalized exactly. If measurement error is present, however, but not taken into account, parameter estimates may be severely biased. The present paper studies measurement error corrected estimation in the context of a huge class of parametric duration models. The proposed quasi-likelihood based method easily allows – as long as no censoring occurs – to deal simultaneously with covariate measurement error as well as with measurement error in the duration itself and yields estimates with sound asymptotic properties. A general formula for the measurement error corrected quasi-score function can be derived, which is valid for most of the commonly used parametric duration models.

Keywords. Measurement error in covariate and duration, error-in-variables, parametric survival analysis, parametric duration models, lifetime data, accelerated failure time models, Weibull model, log-logistic model, quasi-likelihood.

1 Introduction

Duration analysis studies for every $i = 1, \ldots, n$ the time T_i, in which the unit i changes from an original state into an absorbing state (death, first occurrence of a certain disease etc.). This paper concentrates on *parametric* regression models (cf. Section 2 for more details, references and examples) for the relationship between T_i and covariates. This means that T_i is taken to be distributed according to a certain type of distribution, which is assumed to be known up to few parameters, (some of) which may vary with the covariates.

A typical problem in applying regression analysis to real data is the presence of measurement error. Often there is (at least) one covariate X_i of theoretical interest, which cannot be directly observed or measured correctly.[1] However, if one ignores the measurement error by just plugging in substitutes or incorrect measurements W_i instead of X_i ('naive estimation'), then all the parameter estimates must be suspected to be severely biased. Error-in-variables modeling provides a methodology, which is serious about that fact and develops procedures for adjusting for the measurement error.

[1] According to the literature the term 'measurement error' is only applied to continuous variables. The corresponding problem for discrete variables ('misclassification') is not addressed here.

Already in the eighties the effects of covariate measurement errors had been extensively studied for the standard linear model and several powerful methods for adjustment in that case had been developed (e.g. the books of SCHNEEWEISS and MITTAG, 1986, and FULLER, 1987; for a recent monograph on this topic see: CHENG and VAN NESS, 1999). A summary of the state of the art up to 1995 in nonlinear models is provided by CAROLL, RUPPERT, and STEFANSKI (1995). Measurement error in case-control-studies is reviewed in a comprehensive way by THÜRIGEN et. al. (1999).

In contrast to its practical importance, measurement error has not yet attracted much attention in duration analysis. To the author's knowledge for quite a long time the papers by PRENTICE (1982), PEPE, SELF, and PRENTICE (1989), CLAYTON (1991), NAKAMURA (1992), and HUGHES (1993) had more or less formed a complete list of the work dedicated to covariate measurement error in duration analysis. Recently, however, the discussion is becoming a bit more vivid by the contributions of WANG, et. al. (1997), HU, TSIATIS, and DAVIDIAN (1998), and KONG, HUANG, and LI (1998) — some other work is said to appear soon.

All these papers concentrate on *covariate* measurement error in the *Cox*-model. A further issue, which is of particular importance for duration analysis, was raised by HOLT, MCDONALD, and SKINNER (1991) (see also SKINNER and HUMPHREYS (1999) and the references therein). They give empirical evidence that especially in retrospective studies the duration time itself may be subject to a measurement error, which is definitely not of a negligible size. Assuming correctly specified covariates SKINNER and HUMPHREYS (1999) discusses such measurement error in the dependent variable in a Weibull model.

One general methodological principle to deal with measurement error is quasi-likelihood based measurement error correction. In particular this idea underlies the work of ARMSTRONG (1985), LIANG & LU (1991), CAROLL et. al. (1995, Section 7.8 and Appendix A.4), and of THAMERUS (1997, 1998), which is closest to the development here.

An application or even a comprehensive, critical investigation of that methodology in the context of duration models is still lacking. The present study wants to contribute to this question by investigating measurement error corrected quasi-likelihood estimation in parametric duration models. More specifically, the paper is organized as follows: Firstly some aspects of parametric duration models are briefly reviewed. In section 3 the measurement model for the covariate and the assumption underlying it are described. Section 4 turns to covariate measurement error correction via quasi-likelihood. Using a super-model which comprises the most prominent duration models it is possible to arrive at general results, which can easily be specialized to cover the specific model considered. The general form of the measurement error corrected quasi-score function is derived. Solving it yields consistent and asymptotic normally distributed parameter estimates, which furthermore can be shown to be efficient in a generalized sense of the Gauss-Markov theorem.

Section 5 and section 6 are concerned with the problem of – separately or additionally – dealing with messy observations in the dependent variables T_i. Measurement error in the narrower sense as occurring in particular in retrospective studies (cf. the note above) can be easily dealt with in the framework developed, the general formula for the measurement error corrected quasi-score equation can be appropriately extended. In the case of (right-)censoring, however, the method comes up against limiting factors. While the usual likelihood approach is straightforwardly adopted to non-informative random censoring, this seems not to be possible for quasi-likelihood techniques. In the likelihood function the density of the censoring variable appears as a proportionality factor not influencing the place of the maximum, whereas the first and second moments needed for quasi-likelihood estimation depend on the censoring mechanism in quite a complex manner. Some attempts to extend the approach to deal with censored observations indicate that a general formula analogous to the one above can no longer be expected. But it may still be possible to derive solutions using specific properties of the parametric model assumed.

2 Common parametric duration models

2.1 Parametric duration analysis

In *parametric duration models* the individual duration time T_i, $i = 1, \dots, n$, is assumed to follow a distribution, which is taken to be completely known up to few parameters $\theta_1, \dots, \theta_q$. Typically one of them, θ_1 say, is allowed to vary with the individual characteristics, while $\theta_2, \dots, \theta_q$ are assumed to be constant among different cases.

As is also assumed throughout the paper, θ_1 is typically a parameter influencing location. In most cases it is linked to the covariates $X_i, Z_{1i}, \dots, Z_{qi}$ with $(Z_{1i}, \dots, Z_{qi})^T =: Z_i$ and the vector $\beta = (\beta_0, \beta_X, \beta_Z^T)^T$ of regression parameters by the relation

$$\theta_1 = \exp\left(\beta_0 + \beta_X \cdot X + \beta_Z^T \cdot Z\right) . \tag{1}$$

Easy accessible introductions to parametric duration models are provided among many others by the corresponding chapters in the textbooks of Kalbfleisch and Prentice (1980), Miller (1981), Lawless (1982), and Blossfeld, Hamerle, and Mayer (1989). A recent comparative study of fitting different parametric models to biometric data is Lindsey (1998).

By many scholars duration modeling had been equated for some time with an analysis along the lines of Cox's semiparametric model (Cox, 1972). In the last years, however, parametric duration analysis has experienced a slight renaissance and has again attracted much attention. As exemplified by the models below, in many situations parametric duration models provide an interesting alternative against an exclusive use of the Cox model. To mention only some reasons for this: In contrast to the Cox model many parametric duration models can deal with situations where the hazards of different units are not proportional among each other. Also the typical feature of the Cox model, namely to do without a specification of the baseline hazard, is not always advantageous. Often there is at least some material background knowledge about the basic form of the hazard rate, which should be incorporated in the model. In this sense parametric models allow a more detailed reflection of the knowledge, which will result in a more powerful estimation — all the more as parametric models can be constructed from qualitative assumptions for instance on the underlying diffusion process (see e.g. Diekmann, 1990, for the log-logistic model). Furthermore, often the time course of the hazard rate itself is of main interest. Parametric models allow to formulate and to test hypotheses on the overall time dependency like the question 'Does the average risk increase in time or is it decreasing?'.

2.2 Some prominent parametric duration models

In this section some of the most commonly used parametric duration models are briefly reviewed for further reference. In what follows, the distribution of the duration time T_i – conditional on the covariates – is always assumed to be absolutely continuous. For ease of illustration the models are mainly described in terms of the hazard rate

$$\lambda_{T_i}(t|X_i, Z_i; \theta) = \lim_{h \downarrow 0} \frac{P_{T_i}(t \le T_i < t + h|\{T_i \ge t\}, X_i, Z_i; \theta)}{h} .$$

The density of T_i will be denoted by $f_{T_i}(t|X_i, Z_i; \theta)$, the distribution function by $F_{T_i}(t|X_i, Z_i; \theta)$ and the survivor function by $S_{T_i}(t|X_i, Z_i; \theta)$. (Analogous notations will later be used for the other random variables.)

2.3 The log-logistic model

The first model to mention in this context is the log-logistic model with shape parameter $\nu > 2$.[2] Its hazard rate

$$\lambda_{T_i}(t_i|X_i, Z_i; \theta) = \frac{\nu}{t \cdot \left(1 + t^{-\nu} \exp\left(\nu\left(\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i\right)\right)\right)} ,$$

is non-proportional and of quite an appealing form. The hazard first increases and then decreases. The place of the maximum differs with the different covariate values, while for t going to infinity the hazard ratio for each pair of units tends to one. So this models maps situations where at the beginning of the time scale the risk strongly depends on the individual characteristics, while for all those, who have survived the risky period, the hazards are becoming more and more similar independent of the personal profile. Further illustrations of the model can be gained by studying its derivation from a simple diffusion model (e.g. Diekmann, 1990). An example of a biometric application of this model (in a slightly different parameterization) is Bennett (1983). Outside the biometric context it is quite prominent in particular in sociology (Cf. e.g. Blossfeld et. al., 1989), where it is used for instance for modeling divorce rates.

2.4 The Weibull model

A parametric model leading to proportional hazards is the Weibull model. Outside biometrics it is quite popular, e.g., in technometrics to describe the failure of machines. In econometrics it is perhaps *the* model for the distribution of unemployment spells. Biometric application is studied, for instance, in Rao, Talwalker, and Kundu (1991). The hazard rate

$$\lambda_{T_i}(t|X_i, Z_i; \theta) = \nu \left\{\exp(-\beta_0 - \beta_X \cdot X_i - \beta_Z^T \cdot Z_i)\right\}^{\nu} \cdot t^{\nu-1}, \qquad \nu > 0\,,$$

depends on time by a power of t. The monotonicity remains unchanged over time. The direction of time dependency is governed by the shape parameter ν, providing easy ways to test the hypothesis of increasing or decreasing risk. $\nu < 1$ leads to monotonely decreasing hazard, while $\nu > 1$ corresponds to monotonely increasing hazard containing the Rayleigh distribution with linear hazard ($\nu = 2$). The special case of constant hazard ($\nu = 1$) is the exponential model.

2.5 Accelerated failure time model

Both models, the log-logistic one as well as the Weibull model, can be derived from a linear relationship[3] for the logarithm of the lifetimes T_i,

$$\ln T_i = \beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i + \alpha \cdot \epsilon_i\,, \quad \alpha > 0, \tag{2}$$

by taking the error variables ϵ_i to be i.i.d standard logistic distributed or standard minimum extreme value distributed, respectively, and by substituting α by $\frac{1}{\nu}$.

Other choices of ϵ_i (i.i.d. and independent of X_i and Z_i, eventually depending on a parameter vector γ) lead to further prominent models, like the log-normal model, the (generalized) Gamma model, or, as the most comprehensive one, the model connected with the generalized F-distribution (cf. Kalbfleisch and Prentice, 1980, Chapter 2). More generally, all the models arising from relation (2) form a particular class of duration

[2] To exclude singularities of the first two moments $\nu > 2$ is assumed here.

[3] It might be noted that this is no standard linear model, because in general $\mathbb{E}(\epsilon_i)$ is not equal to 0.

models. With model-specific functions $h_0(\cdot)$, their hazard rates and survivor functions have the characteristic form (cf. especially KALBFLEISCH and PRENTICE, 1980, p. 34)

$$\begin{aligned}\lambda_{T_i}(t|X_i, Z_i;\theta) &= h_0\Big(t \cdot \exp\left(\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i\right)\Big) \\ &\quad \cdot \exp\left(\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i\right)\end{aligned}$$

and

$$S(t|X_i, Z_i;\theta) = \exp\left(-\int_0^{t\cdot\exp\left(\beta_0+\beta_X\cdot X_i+\beta_Z^T\cdot Z_i\right)} h_0(w)dw\right) .$$

This led to the name *accelerated failure time models* expressing the idea that the covariates are 'changing the time scale' in the hazard function and the survivor function. In what follows it is additionally assumed that the moment generating function of ϵ_i and its derivative exist in a sufficiently large neighborhood to the right of zero[4].

3 The measurement model for the covariate

The basic preparation for any measurement error correction is to lay carefully down (and check) the assumptions on the relationship between the unobserved variables (often called 'gold standard') and their observable counterparts. In detail, the assumptions listed below and commented later underlie the study of measurement error correction in section 4. They will be supplemented in section 5 to allow for measurement error in the duration time, too.

3.1 Assumptions

[A1] Only one covariate, X_i say, is measured with non-negligible error, whereas the other covariates $Z_{i1}\dots, Z_{iq}$ are taken to be (more or less) exactly known.

[A2] T_i is measured without error. (This will be relaxed in Section 5).

[A3] Additive measurement error: W_i, the incomplete measurement of X_i, is related to X_i via
$$W_i = X_i + U_i .$$

[A4] U_i is independent of T_i, X_i, Z_i and U_j $(i = 1,\dots,n;\ j \neq i)$.

[A5] (Possibly heteroscedastic) Normal error: $U_i \sim \mathcal{N}(0,\sigma_i^2)$ with σ_i^2 known, $i = 1,\dots,n$,

[A6] Structural measurement model: X_i is stochastic.

[A7] Normal distribution of the unobservable variable: $X_i \overset{i.i.d}{\sim} \mathcal{N}(\mu_X,\sigma_X^2)$.

[A8] The conditional distribution of X_i given W_i and Z_i does not depend on Z_i.

3.2 A short discussion of some of the assumptions

Ad [A1]: In practice, this assumption is not so restrictive as it may seem at a first glance. Often the degree of complexity is indeed rather different between the covariates. Typical examples for such situations are studies on the influence of a certain nutrition habit or the true dose-exposure on a certain disease. While these variables are quite difficult to obtain and may be subject to severe measurement error, many other covariates incorporated in

[4] This condition, which guarantees relation (12) to be well defined, sometimes causes restrictions on α.

the studies mainly serve as control factors (like socio-demographic characteristics) and may be determined comparatively exactly.

Ad [A5]: Taking the error to be normally distributed is quite usual in measurement error modeling and is – at the moment[5] – crucial for the method developed in the sections below. Nevertheless, of course, this can be criticized.

Often it makes sense to assume homoscedasticy. In principle the approach allows for homoscedastic measurement error, too. The requirement that the variance(s) should be known is needed to circumvent problems with the identifiability of the parameters. A real data example of a situation, where at least reliable estimates are available, is the radon study of THAMERUS (1997), where each W_i itself is the average of several measurements $W_{i1}, \ldots W_{iq_i}$.

Ad [A6]: Since the covariate measurement error correction studied below is based on the conditional distribution of X_i given W_i, this provides an indispensable requirement. The proposed procedure seems not to be transferable to so-called 'functional measurement error models', which assume a non-stochastic character of X_i.

Ad [A7]: The typically unknown parameters can be unbiasedly estimated from the sample mean $\bar{W}$ and the sample variance S_W^2 of the values $W_1, \ldots, W_n$ by $\hat{\mu}_X = \bar{W}$ and $\hat{\sigma}_X^2 = S_W^2 - \frac{n-1}{n^2}\sum_{i=1}^n \sigma_i^2$. In what follows these estimated values are treated as if they were the true parameter values. Uncertainty in their estimation is not taken into account.

It should be possible to relax the normality assumed along the lines of THAMERUS (1997) to allow for mixtures of normal distributions.

Ad [A8]: This requirement may be tricky. Its appropriateness may be difficult to justify empirically. Therefore, robustness of the results with respect to this assumption should be carefully investigated.

3.3 Some properties resulting from assumptions [A1] to [A8]

Mainly two properties deducible from [A1] to [A8] are needed repeatedly in what follows. Firstly it can be concluded that T_i and W_i are conditionally independent given X_i and Z_i. This means that W_i does not provide information on T_i which is not contained in X_i and Z_i and therefore would be superfluous if X_i were known. According e.g. to CAROLL et. al. (1995, p. 16f.) one speaks then of 'nondifferential measurement error' and calls W_i a 'surrogate variable' for X_i.

Secondly the conditional distribution of X_i given W_i (and Z_i) can be determined. One obtains

$$(X_i|W_i, Z_i) \sim (X_i|W_i) \sim N(\tilde{\mu}_i, \tilde{\tau}_i^2) \tag{3}$$

with

$$\tilde{\mu}_i = \mu_X + \frac{\sigma_X^2}{(\sigma_X^2 + \sigma_i^2)}(W_i - \mu_X)\,, \qquad \tilde{\tau}_i^2 = \sigma_X^2(1 - \frac{\sigma_X^2}{\sigma_X^2 + \sigma_i^2})\,. \tag{4}$$

Since, given W_i, the quantity $\exp(b \cdot X_i)$ is lognormally distributed for every real b, this leads to

$$\begin{aligned} I\!E(\exp(b \cdot X_i)|W_i, Z_i) &= \exp(b \cdot \tilde{\mu}_i) \cdot \exp(0.5b^2 \cdot \tilde{\tau}_i^2) && (5)\\ I\!V(\exp(b \cdot X_i)|W_i, Z_i) &= \exp(2b \cdot \tilde{\mu}_i) \cdot \exp(b^2 \cdot \tilde{\tau}_i^2) \cdot \big(\exp(b^2 \cdot \tilde{\tau}_i^2) - 1\big) && (6) \end{aligned}$$

[5]Some further research to extend the method developed below to other error distribution like the skew normal distribution in the sense of AZZALINI (1985) is still in its infancy.

4 Covariate measurement error corrected quasi-likelihood estimation

While the 'ideal model' without measurement error is specified in terms of Y_i, Z_i and the unknown variable X_i, the estimation of the parameter vector $\theta = (\alpha, \gamma^T, \beta_0, \beta_X, \beta_Z^T)^T$ can of course only be based on the observable quantities Y_i, W_i and Z_i. Since, as already mentioned in the introduction, naive estimates gained by just plugging in W_i instead of X_i may be severely biased, more elaborate estimation procedures are needed, which are serious about the measurement error. They should be able to take the measurement error into account in the estimation procedure by using the relation between the observable and the unobservable quantities formulated in the measurement model.

The likelihood function, which is based on the observable quantities only, is not available in the models considered here. One has to search for a method, which does not use the full likelihood but is nevertheless efficient.

4.1 Some principles of quasi-likelihood based measurement error correction

A promising device seems to adopt the quasi-likelihood approach. Quasi-likelihood – as introduced by WEDDERBURN (1974) – is a general method for parameter estimation based on the first two moments only rather than on the full likelihood. Given mean functions $m_i(\theta)$ and variance functions $v_i(\theta)$ of independent random variables Y_i, $i = 1, \ldots, n$, the *quasi-likelihood estimate* $\hat{\theta}$ is defined to be the root of the so-called *quasi-score equation*

$$\sum_{i=1}^{n} \frac{\partial m_i(\theta)}{\partial \theta} \cdot \frac{(Y_i - m_i)}{v_i(\theta)} = 0\,.$$

Under some regularity conditions quasi-likelihood estimates have a lot of desirable properties (cf. below).

In several different settings quasi-likelihood has proven to be a powerful, easy to handle, tool to adjust for measurement error in a covariate. To prepare the discussion of its effectiveness and its limitations in parametric duration models, the general principle of the way how it will be used here has to be briefly sketched (cf. especially ARMSTRONG, 1985 and THAMERUS, 1997, 1998).

The leitmotif of quasi-likelihood based measurement error correction is the insight that the mean functions and the variance functions of the data with measurement error can be directly expressed in terms of the mean and variance function of the 'ideal model' without measurement error and its parameters. One applies the theorem of iterated expectation and the decomposition formula for variances and uses the non-differentiability of the measurement error (cf. section 3.3). Then the conditional moments $\mathbb{E}(T_i|W_i, Z_i; \theta)$ and $\mathbb{V}(T_i|W_i, Z_i; \theta)$ with respect to the observable quantities can be related to their counterparts $\mathbb{E}(T_i|X_i, Z_i; \theta)$ and $\mathbb{V}(T_i|X_i, Z_i; \theta)$ based on the unobservable quantities:

$$\begin{aligned} \mathbb{E}(T_i|W_i, Z_i; \theta) &= \mathbb{E}\Big(\mathbb{E}(T_i|X_i, W_i, Z_i; \theta)\Big|W_i, Z_i; \theta\Big) = \\ &= \mathbb{E}\Big(\underbrace{\mathbb{E}(T_i|X_i, Z_i; \theta)}_{\text{ideal model}}\Big|\ \underbrace{W_i, Z_i}_{\text{observable}}; \theta\Big) \end{aligned} \tag{7}$$

and

$$\begin{aligned} &\mathbb{V}(T_i|W_i, Z_i; \theta) = \\ &= \mathbb{V}\Big(\mathbb{E}(T_i|X_i, W_i, Z_i; \theta)\Big|W_i, Z_i; \theta\Big) + \mathbb{E}\Big(\mathbb{V}(T_i|X_i, W_i, Z_i; \theta)\Big|W_i, Z_i; \theta\Big) \end{aligned}$$

$$= \mathbb{V}\Big(\underbrace{\mathbb{E}(T_i|X_i,Z_i;\theta)}_{\text{ideal model}} \Big| \underbrace{W_i, Z_i}_{\text{observable}};\theta\Big) + \mathbb{E}\Big(\underbrace{\mathbb{V}(T_i|X_i,Z_i;\theta)}_{\text{ideal model}} \Big| \underbrace{W_i, Z_i}_{\text{observable}};\theta\Big) \tag{8}$$

or

$$\begin{aligned} \mathbb{V}(T_i|W_i,Z_i;\theta) &= \mathbb{E}(T_i^2|W_i,Z_i;\theta) - (\mathbb{E}(T_i|W_i,Z_i;\theta))^2 = \qquad (9) \\ &= \mathbb{E}\Big(\underbrace{\mathbb{E}(T_i^2|X_i,Z_i;\theta)}_{\text{ideal model}} \Big| \underbrace{W_i, Z_i}_{\text{observable}};\theta\Big) - \Big(\mathbb{E}\Big(\underbrace{\mathbb{E}(T_i|X_i,Z_i;\theta)}_{\text{ideal model}} \Big| \underbrace{W_i, Z_i}_{\text{observable}};\theta\Big)\Big)^2 . \end{aligned}$$

Solving the corresponding quasi-score equation, which will be called *measurement error corrected*[6] *quasi-score equation*,

$$\sum_{i=1}^{n} \frac{\partial\ \mathbb{E}(T_i|W_i,Z_i;\theta)}{\partial\ \theta} \cdot \frac{\{T_i - \mathbb{E}(T_i|W_i,Z_i;\theta)\}}{\mathbb{V}(T_i|W_i,Z_i;\theta)} = 0 \tag{10}$$

yields $\hat{\theta} = (\hat{\alpha}, \hat{\gamma}^T, \hat{\beta}_0, \hat{\beta}_X, \hat{\beta}_Z^T)^T$, the *measurement error corrected quasi-likelihood estimate.*

Therefore, for each of the models considered, measurement error corrected quasi-likelihood estimation consists of the following steps:

- Calculate the 'ideal' mean and variance function!
- Correct them along the lines of (7), (8) and (9)!
- Solve the quasi-score equation (10) – for instance by the IRLS algorithm (cf. e.g. McCullagh, 1991)!

4.2 Covariate measurement error corrected quasi-like- lihood estimates

To apply the agenda just described to arrive at covariate measurement error corrected quasi-likelihood estimation in duration models one can take profit from the fact that many of the models possess a common structure — namely that they are of accelerated failure time type. In the context considered here this is of great importance. It can easily be used to calculate the general form of the 'ideal' mean and variance functions as well as – based on this – to derive common expressions for the measurement error corrected mean and variance functions and therefore for the measurement error corrected score function. So a framework is provided, which allows to deal simultaneously with many duration models.

From the linear relationship (2) between log-duration time and the covariates characterizing accelerated failure time models one immediately obtains the mean and the variance functions of the 'ideal model'. Presupposing its existence, the moment of any order r about the origin, conditional on Z_i and the unobservable X_i, is

$$\begin{aligned} \mathbb{E}(T_i^r|X_i,Z_i;\theta) &= \mathbb{E}\Big(\exp\big(r\cdot(\beta_0+\beta_X\cdot X_i+\beta_Z^T\cdot Z_i)\big)\cdot\exp\left(r\alpha\epsilon_i\right)\Big) \\ &= c_r(\alpha)\cdot\exp\Big(r\cdot(\beta_0+\beta_X\cdot X_i+\beta_Z^T\cdot Z_i)\Big), \qquad (11) \end{aligned}$$

where

$$c_r(\alpha) = M_\epsilon(r\cdot\alpha) \tag{12}$$

and $M_\epsilon(\cdot)$ is the moment generating function of one of ϵ_1. It is worth noticing that this factor does not depend on the regression parameters and the covariates. Furthermore, it is the only term which varies with the underlying parametric distribution.

For the two models explicitly exemplified in section 2.2 one obtains

[6]The justification for the term 'measurement error corrected' is handed in later in Proposition 1, where the properties of the estimates are described.

- $c_r(\alpha) = r\pi(\alpha \cdot \sin(r\pi/\alpha))^{-1}$ for the log-logistic model and
- $c_r = \Gamma\left(\frac{\alpha+r}{\alpha}\right)$ for the Weibull model[7].

For instance also the following prominent duration models (cf. the references in section 2) are included. One takes

- $c_r(\alpha) = \gamma^{-\alpha} \cdot \Gamma(\gamma)^{-1} \cdot \Gamma(\gamma + \alpha \cdot r)$ in the generalized gamma model with shape parameter γ, containing the usual gamma distribution by specializing α to 1.
- $c_r(\alpha) = (\Gamma(\gamma_1) \cdot \Gamma(\gamma_2))^{-1} \cdot \Gamma(\gamma_1 + \alpha \cdot r) \cdot \Gamma(\gamma_1 - \alpha \cdot r) \cdot (\gamma_1/\gamma_2)^{r \cdot \alpha}$ in the model connected with the generalized F-distribution with parameter γ_1 and γ_2.
- $c_r(\alpha) = (1 - 0.5 \cdot r^2 \cdot \alpha^2)^{-1}$, $\alpha < (\sqrt{0.5} \cdot r)^{-1}$, in the (standard) log-Laplace model.

It should be noted that the arguments and the calculation following are also valid if $c_r(\alpha)$ is not exactly of the form (12). It is sufficient that $c_r(\alpha)$ does not depend on the covariates X_i subject to measurement error and is differentiable with respect to α. Important examples for this are accelerated failure time models with multiplicative measurement error in the duration time (cf. section 5.2).

To derive the measurement error corrected mean and variance functions for all these models one brings together the general relation between mean and variance functions of the 'ideal model' and their counterparts from the model with measurement error formulated in (7), (8) and (9) with the results on the conditional distribution of the unobservable X_i given the observable W_i concluded from the assumptions [A1] to [A8] on the measurement model (cf. sections 3.1 and 3.3). Accordingly, using (5) and (6) as well as (7) and (8) or (9) together with (4) and (3) one obtains – by additionally taking into account that $c_r(\alpha)$ does not depend on the covariates – the measurement error corrected mean and variance functions

$$\mathbb{E}(T_i|W_i, Z_i; \theta) = c_1(\alpha) \cdot \exp\left(\beta_0 + \beta_X \cdot \tilde{\mu}_i + \frac{\beta_X^2 \cdot \tilde{\tau}_i^2}{2} + \beta_Z^T \cdot Z_i\right) \tag{13}$$

$$\begin{aligned} \mathbb{V}(T_i|W_i, Z_i; \theta) = {} & \Big(c_2(\alpha) \cdot \exp\left(\beta_X^2 \cdot \tau_i^2\right) - c_1^2(\alpha)\Big) \cdot \\ & \cdot \Big(\exp\left(2\beta_0 + 2\beta_X \cdot \tilde{\mu}_i + \beta_X^2 \cdot \tau_i^2 + 2\beta_Z^T \cdot Z_i\right)\Big) \end{aligned} \tag{14}$$

They have to be inserted into the measurement error corrected quasi-score equation (10).

4.3 Properties of the measurement error corrected quasi-likelihood estimates

One important advantage of using quasi-likelihood is that one can simply utilize known general results on the properties of quasi-likelihood estimates in order to establish the properties of the estimates produced by the method for measurement error correction proposed in this paper. Using the results of McCullagh (1983, p. 62) and ensuring that the regularity assumptions mentioned there are satisfied[8], one obtains especially the confirmation that the measurement error bias is indeed asymptotically removed.

Proposition 1 *Let $\hat{\theta}$ be the root (assumed to be unique) of the measurement error corrected quasi-score equation (10). Then*

[7]Note that in this case the distribution of ϵ_i is an extrem-value distribution for the minimum – and not for the maximum, which is usually tabulated. Therefore, in most formula for the moment generating function from the literature ϵ_i has to be substituted by $-\epsilon_i$.

[8]In some situations this may impose some additional restrictions on α to guarantee the existence of the third moment.

- $\hat{\theta}$ *is consistent.*
- $\hat{\theta}$ *is asymptotically normal,*

$$\sqrt{n}(\hat{\theta} - \theta) \overset{n \to \infty}{\sim} \mathcal{N}(0, F^{-1}) \tag{15}$$

with

$$F := I\!\!E\left(-\frac{\partial s(\theta)}{\partial \theta^T}\right) = \sum_{i=1}^{n} \frac{d_i(\theta) \cdot (d_i(\theta))^T}{I\!\!V(T_i|W_i, Z_i; \theta)}$$

denoting the quasi-likelihood analogue to the information matrix and with $d_i(\theta)$ as the vector of the partial derivatives $\frac{\partial\ I\!\!E(T_i|W_i,Z_i;\theta)}{\partial\ \theta}$.

(15) immediately suggests an estimate for the covariance matrix of the asymptotic distribution. It is, however, not clear, whether a sandwich formula may be preferable. For instance Caroll et al. (1995, Appendix A4) and Thamerus (1998, p. 20) come to different conclusions with respect to this question.

The asymptotic distribution from (15) further immediately yields a way to test the significance of the parameters. The application of e.g. McCullagh (1991, p. 271f.) provides another possibility: It guarantees that in the case considered here also a quasi-likelihood function (i.e. a function, whose derivatives lead to the measurement error corrected quasi-score function) must exist. Therefore, alternatively, a test analogous to the likelihood ratio test can be used. This leads indeed to an asymptotically χ^2-distributed statistic (cf. McCullagh, 1983, p. 62).

General quasi-likelihood theory yields furthermore a type of Gauss-Markov theorem for the estimates. They are optimal under all estimates arising from linear estimation functions (McCullagh, 1983, p. 62, cf. also especially Heyde, 1997, p. 21f.). Nevertheless the efficiency might be improved by incorporating measurement error corrected higher order moments.[9] Slightly generalizing (13) one obtains via the same arguments used there

$$\begin{aligned} I\!\!E(T_i^r|W_i, Z_i; \theta) \;=\; c_r(\alpha) \cdot \exp\big(\, & r \cdot \beta_0 + r \cdot \beta_X \cdot \tilde{\mu}_i \;+\; 0.5 r^2 \cdot \beta_X^2 \cdot \tilde{\tau}_i^2 \\ & + r \cdot \beta_Z^T \cdot Z_i \,\big)\,. \end{aligned}$$

5 Measurement error in duration time

In many applications it seems to be quite realistic to suspect the duration times T_i themselves to be incorrectly measured or reported. Indeed, examining as an example data on the age at menarche (first menstruation), Holt et. al. (1991) and Skinner and Humphreys (1999) come to the conclusion that measurement error in the dependent variable may be an important issue, especially in retrospective studies.

Assuming in a Weibull model all covariates to be measured correctly Skinner and Humphreys (1999) developed a method to correct maximum likelihood estimates for measurement error in the duration time. It is a nice feature of the method under study here that additive or multiplicative measurement error in the dependent variable can be handled simultaneously with additive measurement error in one covariate.

[9] To circumvent singularities in the case of the log-logistic distribution some caution is needed. A sufficient condition for the r-th moment to exist is to take $\alpha > r$.

5.1 Additive measurement error in duration time

At first an additive measurement error is considered: Instead of the true duration times T_i, $i = 1, \ldots, n$, the quantities T_i^* are observed according to the relation

$$T_i^* = T_i + V_i \,. \tag{16}$$

Additionally the following assumptions are made:

[A9] V_i, $i = 1, \ldots, n$ are identically and independently distributed with mean $m = \mathbb{E}(V_i)$ and variance $v = \mathbb{V}(V_i)$.

[A10] V_i is 'independent of the rest': V_i, $i = 1, \ldots, n$ is independent of T_j, X_j, W_j, Z_j, ϵ_j and U_j, $j = 1, \ldots, n$, $j \neq i$ and also independent of T_i given X_i and Z_i or given W_i and Z_i.

Then one immediately gets

$$\begin{aligned}
\mathbb{E}(T_i^*|W_i, Z_i; \theta) &= \mathbb{E}(T_i|W_i, Z_i; \theta) + \mathbb{E}(V_i|W_i, Z_i; \theta) = \\
&= \mathbb{E}(T_i|W_i, Z_i; \theta) + m \,, \qquad (17)\\
\mathbb{V}(T_i^*|W_i, Z_i; \theta) &= \mathbb{V}(T_i|W_i, Z_i; \theta) + \mathbb{V}(V_i|W_i, Z_i; \theta) = \\
&= \mathbb{V}(T_i|W_i, Z_i; \theta) + v \,. \qquad (18)
\end{aligned}$$

These corrected mean and variance functions can directly be used to get the quasi-likelihood estimate for θ in the case of no covariate measurement error, i.e. where $X_i = W_i$.

If covariate measurement error is present fulfilling the assumption of section 3, then (17) and (18) can be combined with the relations (7) and (8) or (9) to calculate quasi-likelihood estimates from (10). For any accelerated failure time model the general formulas (13) and (14) for the measurement error corrected mean and variance functions straightforwardly extend to

$$\begin{aligned}
\mathbb{E}(T_i^*|W_i, Z_i; \theta) &= c_1(\alpha) \cdot \exp\left(\beta_0 + \beta_X \cdot \tilde{\mu}_i + \frac{\beta_X^2 \cdot \tilde{\tau}_i^2}{2} + \beta_Z^T \cdot Z_i\right) + m \\
\mathbb{V}(T_i^*|W_i, Z_i; \theta) &= \left(c_2(\alpha) \cdot \exp\left(\beta_X^2 \cdot \tau_i^2\right) - c_1^2(\alpha)\right) \cdot \\
&\quad \cdot \left(\exp\left(2\beta_0 + 2\beta_X \cdot \tilde{\mu}_i + \beta_X^2 \cdot \tau_i^2 + 2\beta_Z^T \cdot Z_i\right)\right) + v \,.
\end{aligned}$$

Solving the corresponding quasi-score equation based on the observation T_i^*,

$$\sum_{i=1}^{n} \frac{\partial\, \mathbb{E}(T_i^*|W_i, Z_i; \theta)}{\partial\, \theta} \cdot \frac{\{T_i^* - \mathbb{E}(T_i^*|W_i, Z_i; \theta)\}}{\mathbb{V}(T_i^*|W_i, Z_i; \theta)} = 0 \,,$$

produces estimates, which are corrected for additive measurement error in duration times T_i and in the covariates X_i yielding mutatis mutandis the asymptotic properties stated in proposition 1.

5.2 Multiplicative measurement error in duration time

In many application another measurement model might be more plausible. The difference $(T_i^* - T_i)$ between the observed and the true duration time may be (taken as) proportional to the true duration time T_i itself. (The longer the spell lasted the higher the measurement error tends to be.) Then T_i is subject to a *multiplicative measurement error*, i.e. T_i^* with

$$T_i^* = V_i \cdot T_i \tag{19}$$

instead of T_i is observed. Again $[A9]$ and $[A10]$ are assumed, but they have to be supplemented by the requirement $V_i > 0$ a.s., $i = 1, \ldots, n$, to exclude senseless results.

If the 'ideal model' in terms of T_i, X_i and Z_i is of accelerated failure time type then the model with incorrectly measured duration time T_i^* is of quite a similar form: Substituting (19) in (2) leads to

$$\ln T_i^* = \ln T_i + \ln V_i = \beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i + \alpha \cdot \epsilon_i + \ln V_i$$

resulting in

$$\begin{aligned} &\mathbb{E}((T_i^*)^r | X_i, Z_i; \theta) \ = \\ &= \exp\Big(r \cdot (\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i)\Big) \ \cdot \ \mathbb{E}\Big(\exp(r \cdot \alpha \cdot \epsilon_i) \cdot V_i^r \Big| X_i, Z_i; \theta\Big). \end{aligned}$$

From [A10] one knows that V_i and ϵ_i are independent. Therefore (e.g. KARR, 1993, Corollary 3.11; p. 77) the variables $\exp(r \cdot \alpha \cdot \epsilon_i)$ and V_i^r are independent, too. Since they further are independent of X_i and Z_i, their conditional distribution given X_i and Z_i and their unconditional distribution coincide, and $\exp(r \cdot \alpha \cdot \epsilon_i)$ and V_i^r are also conditionally independent given X_i and Z_i. Because of this, one obtains

$$\begin{aligned} & \mathbb{E}\Big(\exp(r \cdot \alpha \cdot \epsilon_i) \cdot V_i^r \Big| X_i, Z_i; \theta\Big) \\ = \ & \mathbb{E}\Big(\exp(r \cdot \alpha \cdot \epsilon_i) \Big| X_i, Z_i; \theta\Big) \cdot \mathbb{E}\Big(V_i^r \Big| X_i, Z_i; \theta\Big) \\ = \ & \mathbb{E}\Big(\exp(r \cdot \alpha \cdot \epsilon_i)\Big) \cdot \mathbb{E}(V_i^r). \end{aligned}$$

This finally leads to

$$\mathbb{E}_\theta((T_i^*)^r | X_i, Z_i; \theta) = \bar{c}_r(\alpha) \cdot \exp\Big(r \cdot (\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i)\Big) \tag{20}$$

with

$$\bar{c}_r(\alpha) = M_\epsilon(r \cdot \alpha) \cdot \mathbb{E}(V_i^r), \tag{21}$$

which can be used for quasi-likelihood estimation if no covariate measurement error is present.

To incorporate additional covariate measurement error it is of importance that one is, cum grano salis, led back to the general formula (11). The moment-functions of the mismeasured, observable duration times with respect to the unobservable covariates are just of a form very similar to the 'ideal' moment functions in the model without measurement error in the duration time. Only the quantities $c_r(\alpha)$ appearing in (12) have to be substituted by the expression (21). Since $\bar{c}_r(\alpha)$ still does not depend on the covariates, the results derived in section 4 on incorporating additive measurement error in a covariate fulfilling assumption [A1] to [A8] (cf. section 3.1) remain valid without any restriction. Therefore, after having appropriately substituted $c_r(\alpha)$ by $\bar{c}_r(\alpha)$, $r = 1, 2$, inserting the modified versions of (13) and (14) into the corresponding quasi-score equation yields quasi-likelihood estimates for θ. They take into account the multiplicative measurement error in duration time T_i as well as the additive measurement error in the covariates X_i. Since the general framework is not left, especially the useful asymptotic properties formulated in proposition 1 are still guaranteed.

6 Difficulties arising under censorship

While, as the preceding section showed, measurement error in the dependent variable is easily incorporated, another type of 'messy observations', namely right-censoring, will cause serious problems, which immediately then questions the elegance and efficiency of the methodology. The main aim of this section is to point out, why under censorship the approach comes up against limiting factors. Also an idea how to eventually overcome the difficulties is briefly discussed.

6.1 On the lacking of an 'ideal' mean and variance function under random censorship

Now observations might be right-censored, i.e. for $i = 1, \ldots, n$,

$$Y_i := \min(T_i, C_i) \qquad \text{and} \qquad \Delta_i := I\Big(\{T_i < C_i\}\Big)$$

is observed instead of T_i, where C_i is a random variable describing the censoring process.

If covariate measurement error is present, it seems to be straightforward to extend the principal considerations of section 4.1 to bivariate observations (Y_i, Δ_i) (which is not the main difficulty, however some care is needed with respect to the regularity conditions for the asymptotics) and then to proceed along the lines of the agenda listed at the end of section 4.1. But even the first of the steps listed there becomes problematic. As will be clear soon, under censorship the 'ideal' mean and variance functions prove to be difficult to obtain.

The problem one is confronted with here is that, quite surprisingly, one of the most convenient properties of the likelihood does not carry over to quasi-likelihood. While – under standard assumption on the censoring pattern – the likelihood does not depend on the typically unknown censoring distribution, the mean and variance functions do so in quite an unpleasant way.

To see this, it is useful to briefly recapitulate the derivation of the joint density $f_{Y_i,\Delta_i}(t_i, \delta_i | X_i, Z_i; \theta)$ of an observation (Y_i, Δ_i), which is also equal to the likelihood-contribution of the unit i. If $\delta_i = 1$, then the observation t_i of Y_i corresponds to the event $\{T_i = t_i\} \cap \{C_i > t_i\}$, while otherwise ($\delta_i = 0$) the value t_i is equivalent to $\{C_i = t_i\} \cap \{T_i \geq t_i\}$. Assuming random censorship (i.e. stochastic independence of T_i and C_i given the covariates) and stochastic independence between C_i and the covariates (a 'non-informative censoring-pattern'), yields the well known relation

$$\begin{aligned} & f_{Y_i,\Delta_i}(t_i, \delta_i | X_i, Z_i; \theta) = \\ & = \left(f_{T_i}(t_i | X_i, Z_i; \theta) \cdot S_{C_i}(t_i)\right)^{\delta_i} \cdot \left(f_{C_i}(t_i) \cdot S_{T_i}(t_i | X_i, Z_i; \theta)\right)^{1-\delta_i}, \end{aligned}$$

where, in this subsection, C_i is taken as absolutely continuous with density $f_{C_i}(\cdot)$, $i = 1, \ldots, n$.

Assuming non-informative censorship, the characteristics $S_{C_i}(\cdot)$ and $f_{C_i}(\cdot)$ of the censoring distribution enter the likelihood as proportionality factors. So the distributions of the censoring variables do not influence the place of the maximum and therefore the inference on the parameters.

For the moments of Y_i and Δ_i, however, the marginal densities given X_i and Z_i are needed. One obtains for instance for Y_i

$$\begin{aligned} f_{Y_i}(t_i | X_i, Z_i; \theta) &= f_{Y_i,\Delta_i}(t_i, 1 | X_i, Z_i; \theta) + f_{Y_i,\Delta_i}(t_i, 0 | X_i, Z_i; \theta) = \\ &= f_{T_i}(t_i | X_i, Z_i; \theta) \cdot S_{C_i}(t_i) + f_{C_i}(t_i) \cdot S_{T_i}(t_i | X_i, Z_i; \theta) \end{aligned}$$

leading to

$$\mathbb{E}(Y_i^r | X_i, Z_i; \theta) = \int_0^\infty t^r \cdot f_{Y_i}(t | X_i, Z_i; \theta)\, dt = \tag{22}$$

$$\int_0^\infty t^r \cdot f_{T_i}(t | X_i, Z_i; \theta) \cdot S_{C_i}(t)\, dt \; + \; \int_0^\infty t^r \cdot f_{C_i}(t) \cdot S_{T_i}(t | X_i, Z_i; \theta)\, dt\,,$$

which depends on the unknown distribution of C_i in quite a complex way. The 'ideal' mean and variance functions cannot be determined, a straightforward extension of the procedure of section 4.2 and of the results gained there to random censorship is not possible.

6.2 A look at single censoring

The previous subsection has shown without any ifs and buts that one cannot follow the approach of section 4.2 without further assumptions or knowledge on the distribution of the censoring variables. So one will attempt to alleviate the problem by introducing some additional assumptions on the censoring process, which are as non-restrictive as possible.

One idea, which seems to suggest itself, is to try whether assuming type-I-censorship does help to circumvent the problem. To prepare this and to demonstrate a particular limitation only a simplified version is discussed here, namely the case of *single censoring* (truncation). There the maximal time η under study is taken as fixed (because of ethical or financial restrictions or because for all units the study started and ended at the same time). Then the distributions of the censoring variables C_i reduce to the one-point mass in η, i.e.

$$P(C_i = t) = \begin{cases} 1 & \text{if} \quad t = \eta \\ 0 & \phantom{\text{if}} \quad \text{else} \end{cases} \qquad S_{C_i}(t) = \begin{cases} 1 & \text{if} \quad t \leq \eta \\ 0 & \phantom{\text{if}} \quad \text{else.} \end{cases}$$

The analogue to (23) for the 'ideal' moments then has the form

$$\mathbb{E}(Y_i^r | X_i, Z_i; \theta) = \int_0^\eta t^r f_{T_i}(t|X_i, Z_i; \theta) dt + \eta^r \cdot S_{T_i}(\eta | X_i, Z_i; \theta) \,. \tag{23}$$

The first term is the partial moment $\mathbb{E}^\eta(T_i^r | X_i, Z_i; \theta)$ of T_i given X_i and Z_i with respect to the upper limit η. Trying to bring it into a form analogous to (11) yields

$$\mathbb{E}^\eta(T_i^r | X_i, Z_i; \theta) = \tilde{c}_r \cdot \exp\Big(r \cdot (\beta_0 + \beta_X \cdot X_i + \beta_Z^T \cdot Z_i)\Big)$$

with

$$\tilde{c}_r = \mathbb{E}^{\tilde{\eta}}\Big(\exp(r\alpha\epsilon_i)\Big)$$

and $\tilde{\eta}$ as the appropriate transformation of η according to the rules for partial moments.

At a first glance this formula looks quite similar to (11), and the general form underlying the results of the preceding sections seems to be reproduced again. Nevertheless, the arguments used there are not applicable here. The important difference to above is that now – in contrast to $c_r(\alpha)$ in (11) – the factor $\tilde{c}_r$ regularly depends (via $\tilde{\eta}$!) on the unobservable covariate X_i (and the unknown parameters). Because, however, the derivation of the formulas (13) and (14) for measurement error corrected mean and variance functions was essentially based on treating $c_r(\alpha)$ as fixed with respect to X_i, the general framework developed in section 4 is not extensible to the situation considered now. The dependence of $\tilde{\eta}$ on the covariates can be quite complex so that a general solution comprising all the models looks to be unreachable. For some particular distributions it seems nevertheless possible to derive explicit solutions. Generally, situation-specific results gained by approximations or numerical integration seem to be the best one can hope for, the more as also the second summand in (23) may cause similar difficulties.[10]

7 Conclusion

The quasi-likelihood based method discussed here showed to be a promising tool for measurement error correction in accelerated failure time models without censoring. So the

[10] To avoid misunderstandings it may be noted that the mean and variance functions are aimed at providing a manageable basis for measurement error correction. In the case of correctly measured covariates the effort to determine the mean and variance functions seems to be superfluous, since then a full likelihood analysis is feasible promising more efficient parameter estimation.

measurement error problem in parametric duration models proves to be an instance supporting McCallagh's thesis that via quasi-likelihood "[...] useful inferences are possible even in problems for which a full likelihood-based analysis is either intractable or impossible with the given assumptions." (McCallagh, 1991, p. 265.) A general framework was derived, which even allows to deal simultaneously with covariate measurement error and measurement error in the duration time itself. The method is computationally tractable. Sound asymptotic properties of the resulting estimates were established, proving that the inconsistency in naive estimation is overcome and the method properly adjusts for measurement error. Additionally, also measurement error corrected moments of higher order were derived, which might be used to improve the efficiency of the estimation further. The principles of the method are not restricted to accelerated failure time models, the procedure described here can be adopted to other duration models, for instance to the inverse Gaussian model.

Here the whole argumentation was given in terms of the duration time T_i. This is so-to-say the 'natural scale', which is more vivid for the generalization to other duration models and for understanding the censoring process. For accelerated failure time models it is an open question, in which situations it might be advantageous to base the consideration on the related distribution ϵ_i and to use the linear relationship between $\ln T_i$ and ϵ_i directly for estimation.

Some extensions of the measurement model, which may be worth a more detailed study, have already been mentioned in section 3. Further work is needed to elaborate the handling of censored observations. Section 6 has made it clear that a framework of a generality comparable to the results of section 4.2 must not be expected. In some situations (e.g. in the Weibull model) the approach briefly discussed in section 6.2 seems to provide feasible explicit solutions, which have to be further investigated (Augustin, 1999). Two other attempts to handle censored observations should be finally mentioned. It may be successful to process semiparametric estimates of the censoring distribution (based on a piecewise exponential assumption) or to work with Buckley-James-like pseudo-observations. Both ideas showed some preliminary encouraging features but that research is too much in its infancy to be reported here in more detail.

Acknowledgments. I thank Hans Schneeweiß for stimulating discussions on the topic and helpful comments on an earlier draft of this paper.

References

Armstrong, B., 1985: Measurement error in the generalized linear model. *Communications in Statistics, Part B – Simulation and Computation.* **14**, 529–544.

Augustin, T., 1999: Conditional likelihood and quasi-likelihood estimation in the Weibull model under censoring and covariate measurement error. In preparation.

Azzalini, A., 1985: A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics.* **12**, 171–178.

Bennett, S. 1983: Log-logistic Regression Models for Survival Data. *Applied Statistics* **32**, 165–171.

Blossfeld, H. P., Hamerle, A., and Mayer, K. U., 1989: *Event History Analysis: Statistical Theory and Application in the Social Sciences.* Lawrence Erlbaum, Hillsdale.

Caroll, R.J., Ruppert, D., and Stefanski, L.A., 1995: *Measurement Error in Nonlinear Models.* Chapman and Hall, London.

Cheng, C.-L., and van Ness, J.W., 1999: *Statistical Regression with Measurement Error.* Arnold, London.

Clayton, D. G., 1991: Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In: J. H. Dwyer, M. Feinleib, and P. Lipsert et al. (eds.): *Statistical Models for Longitudinal Studies of Health.* Oxford University Press, New York, 301–331.

DIEKMANN, A. 1990: Diffusion and survival models for the process of entry into marriage. In: K. U. Mayer, N. B. Tuma (eds.): *Event History Analysis in Life Course Research.* University of Wisconsin Press, Wisconsin, 170–184.

FULLER, W. A., 1987: *Measurement Error Models.* Wiley, New York.

HEYDE, C. F., 1997: *Quasi-Likelihood and its Application. A General Approach to Parameter Estimation.* Springer, New York.

HOLT, D., MCDONALD, J.W., and SKINNER, C.J., 1991: The effect of measurement error on event history analysis. In: P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, S. Sudman (eds.): *Measurement Errors in Surveys.* Wiley. New York, 665–685.

HU, P., Tsiatis, A., and DAVIDIAN, M., 1998: Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics.* **54**, 1407–1419.

HUGHES, M. D., 1993: Regression dilution in the proportional hazards model. *Biometrics.* **49**, 1056–1066.

KARR, A. F., 1993: *Probability.* Springer, New York.

KALBFLEISCH, J. D., and PRENTICE, R. L., 1980: *The Statistical Analysis of Failure Time Data.* Wiley, New York.

KONG, F. H., HUANG, W., and LI, X., 1998: Estimating survival curves under proportional hazards model with covariate measurement errors. *Scandinavian Journal of Statistics.* **25**, 573–587.

LAWLESS, J. F., 1982: *Statistical Models and Methods for Lifetime Data.* Wiley, New York.

LIANG, K.-Y. and LIU, X.-H., 1991: Estimating equations in generalized linear models with measurement error. In: V. P. Godambe (ed.): Estimating Functions Clarendon Press, Oxford, 47–63.

LINDSEY, J. K., 1998: A study of interval censoring in parametric regression models. *Lifetime Data Analysis* **4**, 329–354.

MCCULLAGH, P., 1983: Quasi-likelihood functions. *The Annals of Statistics.* **11**, 59–67.

MCCULLAGH, P., 1991: Quasi-likelihood and estimating functions. In: D. V. Hinkley, N. Reid and E. J. Snell (eds.): *Statistical Theory and Modelling.* Chapman and Hall, London, 265–286.

MILLER, R. G. Jr., 1991: *Survival Analysis.* Wiley. New York.

NAKAMURA, T., 1992: Proportional hazards model with covariates subject to measurement error. *Biometrics.* **48**, 829–838.

PEPE, M. S., SELF, S. G., and PRENTICE, R. L., 1989: Further results in covariate measurement errors in cohort studies with time to response data. *Statistics in Medicine.* **8**, 1167–1178.

PRENTICE, R. L., 1982: Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika.* **69**, 331–342.

RAO, B.R., TALWALKER, S., and KUNDU, D., 1991: Confidence intervals for the relative relapse rate of placebo VS 6-Mercaptopurine group of acute Leukemia patients under a random censorship model. *Biometrical Journal.* **33**, 579–598.

SKINNER, C. J., and HUMPHREYS, K., 1999: Weibull regression for lifetimes measured with error. *Lifetime Data Analysis.* **5**, 23–37.

SCHNEEWEISS, H. and Mittag, H. J., 1986: *Lineare Modelle mit fehlerbehafteten Daten.* Physica, Heidelberg.

THAMERUS, M., 1997: Modelling count data with heteroscedastic measurement error in the covariates. *SFB Discussion Paper.* **58**, University of Munich; `www.stat.uni-muenchen.de/sfb386/publikation.html` To appear in: *Computational Statistics.*

THAMERUS, M., 1998: Nichtlineare Regressionsmodelle mit heteroskedastischen Meßfehlern. (Non-linear Regressions Models with Heteroscedastic Measurement Error, in German). Logos, Berlin.

THÜRIGEN, D., SPIEGELMAN, D., BLETTNER, M., HEUER, C., and BRENNER, H., 1999: Measurement error correction using validation data – a review of methods and their applicability in case-control. *Submitted.*

WEDDERBURN, R. W. M., 1974: Quasi-likelihood functions, generalised linear models and the Gauss-Newton method. *Biometrika*. **61**, 439–447.

WANG, C. Y., HSU, L., FENG, Z. D. and PRENTICE, R L., 1997: Regression calibration in failure time regression. *Biometrics*. **53**, 131–145.