Universidade de Vigo

Censored regression when censoring indicators are missing at random.

Jorge Mendonça and Jacobo de Uña Álvarez

Report 11/01

Discussion Papers in Statistics and Operation Research

Departamento de Estatística e Investigación Operativa

Facultade de Ciencias Económicas e Empresariales
Lagoas-Marcosende, s/n · 36310 Vigo
Tfno.: +34 986 812440 - Fax: +34 986 812401
http://webs.uvigo.es/depc05/
E-mail: depc05@uvigo.es

Universidade de Vigo

Censored regression when censoring indicators are missing at random.

Jorge Mendonça and Jacobo de Uña Álvarez

Report 11/01

Discussion Papers in Statistics and Operation Research

Imprime: GAMESAL

Edita:

Universidade de Vigo
Facultade de CC. Económicas e Empresariales
Departamento de Estatística e Investigación Operativa
As Lagoas Marcosende, s/n 36310 Vigo
Tfno.: +34 986 812440

I.S.S.N: 1888-5756

Depósito Legal: VG 1402-2007

Censored regression when censoring indicators are missing at random

Jorge Mendonça[1] and Jacobo de Uña-Álvarez[2,*]

[1]Polytechnic Institute of Oporto, School of Engineering, Portugal. E-mail: jpm@isep.ipp.pt

[2]Department of Statistics and OR, University of Vigo. E-mail: jacobo@uvigo.es

[*] Corresponding author. *Full postal address:* Jacobo de Uña-Álvarez, Facultad de Ciencias Económicas y Empresariales, Universidad de Vigo, Campus Universitario Lagoas-Marcosende, 36310 Vigo, Spain

Abstract

In this paper we propose a new regression method with censored responses when some censoring indicators are missing at random. We assume a parametric (linear or nonlinear) model for the regression function, and the vector parameter is estimated by randomly weighted least squares. The weights in the least squares criterion are presmoothed Kaplan-Meier weights which take the missing censoring indicator problem into account. We give the conditions under which the proposed method is consistent. Extensive simulations to investigate the finite sample performance of the regression parameter estimator and a real data illustration are included. The proposed estimator is compared to a multiple imputation type estimator. One of the conclusions is that no one of the methods dominates the other.

Key words: Kaplan-Meier, multiple imputation, presmoothing, semiparametric censoring

1 Introduction

In Survival Analysis and in other fields, censored data appear. Let Y be a response variable (e.g. a lifetime) which is observed under right-censoring, so one is only able to observe the pair (Z, δ) where $Z = \min(Y, C)$ is the observed response, $\delta = I(Y \leq C)$ is the censoring indicator, and C stands for the potential censoring time. Besides, let X be a p-dimensional vector of covariates. In many applications, one will be interested in the regression model

$$Y = f(X; \theta_0) + \varepsilon$$

where ε is a zero-mean error term and $f(x;\theta)$ is a parametric (linear, nonlinear) specification for the regression function $f(x) = E[Y|X = x]$, with θ varying in a subset Θ of the q-dimensional Euclidean space; θ_0 is reserved for the true (unknown) regression parameter. In Survival Analysis this model is usually identified as an accelerated failure time model; in such a case, the Y represents a logarithmic transformation of the lifetime. Given an iid sample (X_i, Z_i, δ_i), $i = 1, ..., n$, of (X, Z, δ), consistent estimation of the regression parameter θ is given by the minimizer of

$$\theta \mapsto \sum_{i=1}^{n} W_i \left[Z_i - f(X_i; \theta)\right]^2,$$

where W_i is the Kaplan-Meier weight attached to the lifetime Z_i, $i = 1, ..., n$ (see Stute, 1999, and references therein).

In some instances, however, the censoring indicators δ_i may not be observed completely. For example, in clinical trials, individuals may fail from multiple causes, one of which is of interest. The time to death from the cause of interest may be censored by a death from a different cause. However, sometimes the cause of death may be unavailable; for example documenting whether or not death is attributable to the cause of interest may require information that is lost or not collected to save expenses, or it may be difficult to determine the cause of death for some patients. In such situations, some censoring indicators are missing, and the sampling information is restricted to $(X_i, Z_i, \xi_i\delta_i, \xi_i)$, $i = 1, ..., n$, iid copies of $(X, Z, \xi\delta, \xi)$, where $\xi = 1$ when δ is available (and $\xi = 0$ otherwise). Under the assumption of missingness at random, namely

$$P(\xi = 1|Z, \delta) = P(\xi = 1|Z),$$

efficient estimation of the marginal survival function of Y was proposed by van der Laan and McKeague (1998). These authors introduced a nonparametric maximum likelihood estimator based on reduced data produced by a discretization of Y. Later, Wang and Ng (2008) (see also Subramanian, 2004) provided alternative estimators based on some preliminary presmoothing of the data. Basically, the alternative method consists in replacing each (missing) censoring indicator δ_i by some fitted value $m_n(Z_i)$ to the probability of uncensoring for Z_i. Comparisons between this approach and that in van der Laan and McKeague (1998) suggest that the presmoothing approach may be preferable in practice (same references). Recently, Subramanian (2009) compared parametric presmoothing methods to multiple imputation methods in the same setup. See our Section 2 for further details. However, none of these papers consider the regression framework.

Our "missing at random" (MAR) condition is given by

$$P(\xi = 1|X, Z, \delta) = P(\xi = 1|X, Z).$$

This condition states that the variables ξ and δ are conditionally independent given X and Z. This assumption is weaker than "missing completely at random" (MCAR), which states the independence between ξ and (X, Z, δ). Wang and Shen (2008) introduced a substitute for Beran's estimator of the conditional survival function of Y given X by using MAR. Gijbels, Lin and Ying (1993) investigated the Cox regression model under MCAR, while McKeague and Subramanian (1998) provided an alternative approach to estimation. Subramanian (2000) considered estimation under the Cox model and proportionality of the conditional hazards of Y and C given X. Goetghebeur and Ryan (1995) analyzed competing risks survival data with proportional hazards regression models under MAR, and Chen et al. (2009) extended their approach to deal with possibly non-proportional baseline hazards. Finally, Zhou and Sun (2003) adapted the additive hazards regression model (Lin and Ying, 1994) to the missing censoring indicators problem. However, for the best of our knowledge, the accelerated failure time model has not been addressed in this context.

The rest of the paper is organized as follows. In Section 2 we introduce the proposed estimator of the vector parameter θ, and we give the conditions under which it is consistent. In Section 3 we provide an extensive simulation study which allows to investigate the finite sample behaviour of the proposed estimator. A real data illustration is given in Section 4. Main conclusions and a final discussion is reported in Section 5.

2 The estimators. Consistency

Starting with the complete case, let $Z_{(1)} \leq ... \leq Z_{(n)}$ be the ordered Z-sample, where (by convention) uncensored cases preceed the censored ones in the case of ties. We denote by $\left(X_{[i]}, \delta_{[i]}\right)$ the (X, δ) value attached to $Z_{(i)}$ (i.e. the i-th concomitant), so the Kaplan-Meier weight attached to $Z_{(i)}$ is given by

$$W_{(i)} = \frac{\delta_{[i]}}{n-i+1} \prod_{j=1}^{i-1} \left[1 - \frac{\delta_{[j]}}{n-j+1}\right].$$

Stute (1999) proposed the minimizer of

$$\theta \mapsto \sum_{i=1}^{n} W_{(i)} \left[Z_{(i)} - f(X_{[i]};\theta)\right]^2$$

as a consistent estimator of θ. Of course, this estimator can not be computed when some censoring indicators are missing. In the following, we propose a possible modification of the weights $W_{(i)}$ to overcome this issue.

Introduce $m(x,z) = P(\delta = 1|X = x, Z = z)$, and let $m(.;\beta)$ denote a parametric family containing the true $m(.)$. Under MAR we have $m(x,z) = P(\delta = 1|X = x, Z = z, \xi = 1)$ and hence the model $m(.;\beta)$ may be consistently fitted via maximization of the conditional likelihood of the observed censoring indicators, this is

$$\beta \mapsto \prod_{\xi_i=1} m(X_i, Z_i; \beta)^{\delta_i} \left[1 - m(X_i, Z_i; \beta)\right]^{1-\delta_i}. \tag{1}$$

(Lu and Tsiatis, 2001). Let β_n be the corresponding maximizer, and introduce $m_n(x, z) = m(x, z; \beta_n)$. Put

$$W_{(i)}(m_n) = \frac{m_n(X_{[i]}, Z_{(i)})}{n-i+1} \prod_{j=1}^{i-1} \left[1 - \frac{m_n(X_{[j]}, Z_{(j)})}{n-j+1}\right].$$

When $P(\xi = 1) = 1$, these are the presmoothed Kaplan-Meier weights introduced in de Uña-Álvarez and Rodríguez-Campos (2004). These authors established the strong consistency of general empirical integrals based on these weights. Here we give the same result in the context of missing censoring indicators. As in Stute (1999), two identifiability conditions are needed to cope with the censoring; these are:

(i) Y and C are independent,
(ii) $P(\delta = 1|X, Y) = P(\delta = 1|Y)$

On the other hand, for the estimator m_n, a uniform convergence condition is needed, namely

(U) $\sup_{x,z} |m_n(x, z) - m(x, z)| \to 0$ with probability 1

See Dikta (1998, 2000) for conditions under which (U) holds for $m_n(x, z) = m(x, z; \beta_n)$. For completeness, we report here the consistency result. We denote by H the distribution function of Z. Introduce

$$F^0_{XY}(x, y) = P(X \le x, Y \le y \wedge \tau_H)$$

where $\tau_H = \inf\{z : H(z) = 1\}$.

Theorem 1. Assume that H is continuous, (i), (ii) and (U) hold, and

$$\int \frac{|\varphi(u, v)| \, F^0_{X,Y}(du, dv)}{m(u, v)(1 - H(v))^\rho} < \infty$$

is satisfied for some $\rho > 0$. Then, with probability 1,

$$\sum_{i=1}^{n} W_{(i)}(m_n)\varphi(X_{[i]}, Z_{(i)}) \to \int \varphi dF^0_{X,Y}.$$

Proof. See the proof to Theorem 2.1 in de Uña-Álvarez and Rodríguez-Campos (2004).■

As in Stute (1999), Theorem 1 is enough to get consistency of the minimizer of

$$\theta \mapsto \sum_{i=1}^{n} W_{(i)}(m_n) \left[Z_{(i)} - f(X_{[i]};\theta) \right]^2 , \tag{2}$$

say $\theta_n(m_n)$, provided that the function $f(x;\theta)$ fulfills a number of regularity conditions. Introduce the following assumptions:

(iii) $EY^2 < \infty$,
(iv) Θ is compact,
(v) $f(x;\theta)$ is continuous in θ for each x,
(vi) $f(x;\theta)^2 \leq M(x)$ for some integrable function M,
(vii) $L(\theta,\theta_0) = E\left\{ (f(X;\theta) - f(X;\theta_0))^2 \right\} > 0$ for each $\theta \neq \theta_0$.

The second moment assumption (iii) is always needed in least-squares estimation. Assumptions (iv) and (v) guarantee that $\theta_n(m_n)$ exists. (iv) and (vi) are not needed for linear regression, while (v) is evidently true in this case. Condition (vi) together with (v) and dominated convergence guarantee that all relevant integrals (such as L) are continuous in θ. Finally, (vii) guarantees that θ_0 may be identified from a sample of (X,Y)'s. It holds true for linear f if X has a finite second moment and is not concentrated on a hyperplane. For notational convenience, we assume without further mention that τ_H equals the upper limit of the lifetime distribution; in such a case, F^0_{XY} is just the joint distribution function of (X,Y). When the support of C is contained in that of Y, a proper modification of L is needed; see equation (1.2) in Stute (1999) for further details.

Theorem 2. Under (i)-(vii) and (U), $\theta_n(m_n) \to \theta_0$ with probability 1.

Proof. See the proof to Theorem 1.1. in Stute (1999).■

An alternative weighted least-squares estimator for θ under MAR may be introduced through a multiple imputation strategy. Let K be a fixed, positive integer (the number of imputations) and, for each i, let $\widehat{\delta}_{[i],k}$ be a $Bernoulli(m_n(X_{[i]}, Z_{(i)}))$ random variable, $k = 1, ..., K$. Put

$$\widehat{\delta}_{[i]} = \xi_{[i]}\delta_{[i]} + (1 - \xi_{[i]})\frac{1}{K}\sum_{k=1}^{K}\widehat{\delta}_{[i],k}$$

where $\xi_{[i]}$ is the value of ξ attached to $Z_{(i)}$, and let $\widehat{W}_{(i)}$ be the imputed Kaplan-Meier weight which is obtained by replacing $\delta_{[i]}$ by $\widehat{\delta}_{[i]}$ in the definition of $W_{(i)}$. Let $\widehat{\theta}_n = \arg\min_\theta \sum_{i=1}^{n} \widehat{W}_{(i)} \left[Z_{(i)} - f(X_{[i]};\theta) \right]^2$. For the problem of estimating the marginal distribution of Y, Subramanian (2009) suggested that the multiple imputation strategy may be preferable when the parametric model $m(.;\beta)$ is miss-specified. However, when the parametric model is correct, it was shown that multiple imputation suffers from a larger variance. In the following Section, we compare by simulations $\theta_n(m_n)$ to $\widehat{\theta}_n$. It will be seen that, in our regression context, no one of these two methods dominates the other.

3 Simulation study

In this Section we provide a simulation study in order to illustrate the relative performance of the presmoothed and the multiple imputation estimators, $\theta_n(m_n)$ and $\widehat{\theta}_n$ respectively. We consider the scenario simulated in de Uña-Álvarez and Rodríguez-Campos (2004). Namely, we proceed as follows: fix real numbers θ_{10}, θ_{20}, and $b_1 > 0$;

(I) Draw $X \sim U(0,1)$
(II) Given X, draw $Y \sim Weibull\,(a_1(X), b_1)$ where $a_1(x) = \exp(\theta_{10} + \theta_{20}x)$
(III) Draw independently $C \sim Weibull(a_2, b_2)$ where $a_2 = \exp(\theta_{10})$ and $b_2 = b_1$
(IV) Compute $Z = Y \wedge C$ and $\delta = 1_{\{Y \leq C\}}$

Note that assumptions (i) and (ii) in the previous Section hold due to the independence between C and (X, Y). Our simulation plan results in a logistic model for m, under which δ and Z are conditionally independent given X; namely

$$m(x,z) = \frac{\exp(b_1\theta_{20}x)}{1+\exp(b_1\theta_{20}x)}. \tag{3}$$

It is easily seen that the simulated (X, Y) follow the loglinear regression model

$$\ln Y = -\theta_{10} - \theta_{20}X + \frac{1}{b_1}\varepsilon$$

where ε follows an extreme value distribution, independent of the covariate. Hence, in our simulated example we have $f(x;\theta_0) = -\theta_{10} - \theta_{20}x$ after performing a log-transformation of the response, i.e. the log-linear case. Since $E(\varepsilon) = -0.57722$, a systematic bias in the estimation of the intercept will appear; this bias is removed in Tables referred below. Parametric presmoothing was considered, under the logistic assumption

$$m(x,z;\beta_0) = \frac{\exp(\beta_{10}+\beta_{20}x+\beta_{30}z)}{1+\exp(\beta_{10}+\beta_{20}x+\beta_{30}z)}. \tag{4}$$

Note that the true m given in (3) belongs to this parametric family. Under the simulated example, the proportion of uncensored responses equals

$$E(\delta) = \frac{1}{b_1\theta_{20}} \ln\left[\frac{1+\exp(b_1\theta_{20})}{2}\right].$$

Table 1: True parameter values in the logistic model for the probability of non-missingness and cases (b), (c) and (d). CP and MP stand for censoring percentage and missing rate respectively.

		MP=33%			MP=67%		
	CP(%)	γ_{10}	γ_{20}	γ_{30}	γ_{10}	γ_{20}	γ_{30}
$\pi(x,z)=\pi(x)$ (Case (b))	—	0	-1.4436	0	0	1.4436	0
$\pi(x,z)=\pi(z)$	10	-1	0	2.8	1	0	2.8
(Case (c))	28	-1	0	1	1	0	-1
	45	-0.5	0	-0.5	0.5	0	0.5
	62	0	0	-1.35	0	0	1.35
$\pi(x,z)=\pi(x,z)$	10	0	-2	2	0	2	-2
(Case (d))	28	0	-0.92	-0.92	0	0.92	0.92
	45	0	-4	2.3	0	4	-2.3
	62	0	1	-2.55	0	-1	2.55

We considered the case $b_1 = 1$. Given b_1, the value of the slope θ_{20} was fixed in order to obtain five censoring percentages, about 10, 28, 45 and 62% of censoring. The intercept θ_{10} was chosen to be zero in all the cases

In order to introduce some missingness in the censoring indicators, we considered four different situations for the function $\pi(x,z) = P(\xi = 1|X = x, Z = z)$: (a) $\pi(x,z) = \pi$, which means that the value of (X,Z) does not influence the probability of missingness (MCAR); (b) $\pi(x,z)$ free of z but dependent on x; (c) $\pi(x,z)$ free of x but dependent on z; and (d) $\pi(x,z)$ depending on both x and z. A logistic model of the form

$$\pi(x,z;\gamma_0) = \frac{\exp(\gamma_{10} + \gamma_{20}x + \gamma_{30}z)}{1 + \exp(\gamma_{10} + \gamma_{20}x + \gamma_{30}z)}$$

was used in all the situations, for imposing a percentage of missingness of 33% or 67%. In Table 1 we report the γ's for each situation. Note that the obvious situation (a) is not reported in the Table 1, while for situation (b) (unlike for (c) and (d)) the same triplet $(\gamma_{10}, \gamma_{20}, \gamma_{30})$ leads to a given missingness rate independently of the censoring degree.

Given the scenario with missing censoring indicators, we fitted the presmoothing function $m(x,z;\beta_0)$ via maximization of (1) and then we computed the minimizer $\theta_n(m_n) = (\theta_{1n}, \theta_{2n})$ of the weighted least squares criterion (2). We also computed the multiple imputation estimator $\widehat{\theta}_n = \left(\widehat{\theta}_{1n}, \widehat{\theta}_{2n}\right)$ described in the previous Section. Sample sizes of n =100 and n =500 were considered. For each case, M =1,000 trials were performed. In Tables 2-5 we report the bias, standard deviation (SD), and mean square error (MSE) of the estimator of (minus) the intercept θ_{10} along the simulations, these are

$$Bias(\theta_{1n}) = \theta_{10} - \frac{1}{M}\sum_{m=1}^{M}\theta_{1n}^{(m)}, \qquad SD(\theta_{1n}) = \left[\frac{1}{M-1}\sum_{m=1}^{M}\left(\theta_{1n}^{(m)} - \theta_{1n}^{(\bullet)}\right)^2\right]^{1/2},$$

$$MSE(\theta_{1n}^{(m)}) = Bias(\theta_{1n})^2 + SD(\theta_{1n})^2,$$

where $\theta_{1n}^{(m)}$ stands for the estimator θ_{1n} when based on the m-th trial and $\theta_{1n}^{(\bullet)} = M^{-1}\sum_{m=1}^{M}\theta_{1n}^{(m)}$. For comparison purposes, we included the figures pertaining to $\widehat{\theta}_{1n}$; more specifically, we report the results corresponding to the best (in the sense of achieving the minimum MSE) and the worst multiple imputation estimators among those using $K = 25$ imputations at maximum. In all the considered cases, the worst multiple imputation method was obtained for $K = 1$, that is, the single imputation estimator. However, the number of imputations needed to achieve the smallest MSE varied. Tables 6-9 report the same kind of figures but for the estimators of the slope (θ_{2n} and $\widehat{\theta}_{2n}$). We also included in Tables 2-9 the results corresponding to the case with non missingness. For this particular case, the Kaplan-Meier based least-squares estimator in Stute (1999) (see also Stute, 1993) can be computed too; the results of this estimator can be regarded as a 'gold standard' to which the remaining estimators can be compared to. The figures of the Kaplan-Meier estimator are reported separately in Table 10.

From Tables 2-10 we see that (for all the estimation methods) the bias, the standard deviation, and the MSE for $n = 500$ are smaller than for $n = 100$, and that they tend to increase when the censoring becomes heavier. These features were expected. Also, the MSEs for the estimators of the slope are greater than those of the intercept. We also see that, generally speaking, the bias increases with the proportion of missingness. This is not always the case when the probability of missingness is influenced by the observed lifetime Z (Tables 4, 5, 8, and 9); in such cases, less bias may be obtained with higher rates of missingness. In most of the situations, the standard deviation increases with the missingness proportion; but some exceptions to this are found when $\pi(x, z)$ depends on both arguments, for a 10% of censoring. Up to moment we have no explanation to these somehow counter-intuitive results.

The contribution of the standard deviation to the final MSEs is more important than that associated to the bias terms, which are in general of a smaller order. When comparing the several estimation methods through the attained MSEs, we see that the multiple imputation method $\widehat{\theta}_n$ performs better than the presmoothed estimator $\theta_n(m_n)$ only in 20.8% of the cases; besides, all theses cases correspond to light censoring. This suggests that, in general, $\theta_n(m_n)$ could be preferable to $\widehat{\theta}_n$, particularly when working under moderate to large censoring levels. Moreover, when $\widehat{\theta}_n$ performs better, the optimal number of imputations varies between 10 and 25; this variation makes the practical usage of $\widehat{\theta}_n$ complicated.

Table 2: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (a) $\pi(x,z) = \pi$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	0	-0.0903	0.2879	0.0910	-0.0911	0.2882	0.0913	21	-0.0939	0.3356	0.1213
		33	-0.0984	0.3207	0.1124	-0.0970	0.3210	0.1123	22	-0.0954	0.3578	0.1370
		67	-0.1021	0.4083	0.1770	-0.1018	0.4081	0.1767	11	-0.0969	0.4310	0.1950
100	28	0	-0.1538	0.2915	0.1085	-0.1520	0.2935	0.1092	21	-0.1663	0.3753	0.1683
		33	-0.1569	0.3159	0.1243	-0.1540	0.3159	0.1234	21	-0.1593	0.4002	0.1854
		67	-0.2040	0.4062	0.2064	-0.2034	0.4076	0.2073	18	-0.2190	0.4709	0.2695
100	45	0	-0.1617	0.3101	0.1222	-0.1596	0.3148	0.1245	21	-0.1780	0.4347	0.2205
		33	-0.1684	0.3233	0.1328	-0.1669	0.3284	0.1356	25	-0.2063	0.4496	0.2445
		67	-0.1951	0.3839	0.1853	-0.1955	0.3844	0.1859	24	-0.2372	0.5095	0.3156
100	62	0	-0.0907	0.3540	0.1334	-0.0921	0.3585	0.1369	24	-0.1126	0.5151	0.2778
		33	-0.0965	0.3665	0.1435	-0.0980	0.3747	0.1498	24	-0.1212	0.5392	0.3052
		67	-0.1468	0.4187	0.1967	-0.1503	0.4234	0.2017	25	-0.1673	0.5646	0.3464
500	10	0	-0.0380	0.1257	0.0172	-0.0377	0.1253	0.0171	24	-0.0340	0.1458	0.0224
		33	-0.0330	0.1289	0.0177	-0.0352	0.1348	0.0194	14	0.0249	-0.0366	0.0236
		67	-0.0432	0.1552	0.0259	-0.0426	0.1552	0.0259	15	-0.0451	0.1754	0.0328
500	28	0	-0.0626	0.1189	0.0180	-0.0613	0.1200	0.0181	25	-0.0616	0.1569	0.0284
		33	-0.0680	0.1282	0.0210	-0.0675	0.1308	0.0217	17	-0.0705	0.1601	0.0306
		67	-0.0760	0.1527	0.0291	-0.0750	0.1539	0.0293	25	-0.0784	0.1885	0.0417
500	45	0	-0.0672	0.1488	0.0266	-0.0668	0.1503	0.0270	22	-0.0732	0.2167	0.0522
		33	-0.0682	0.1526	0.0279	-0.0677	0.1545	0.0284	23	-0.0763	0.2124	0.0509
		67	-0.0756	0.1665	0.0334	-0.0747	0.1685	0.0339	24	-0.0844	0.2465	0.0678
500	62	0	-0.0053	0.1916	0.0367	-0.0062	0.1962	0.0385	25	0.2962	0.2963	-0.0130
		33	-0.0094	0.1938	0.0376	-0.0079	0.1979	0.0392	24	0.2984	0.2983	-0.0064
		67	-0.0154	0.2082	0.0435	-0.0153	0.2107	0.0446	24	0.3035	0.3038	-0.0164

Table 3: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (b) $\pi(x,z) = \pi(x)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	-0.1040	0.3293	0.1192	-0.1034	0.3290	0.1188	20	-0.1070	0.3658	0.1451
		67	-0.1061	0.3513	0.1346	-0.1031	0.3516	0.1342	18	-0.1093	0.3895	0.1635
100	28	33	-0.1672	0.3286	0.1359	-0.1650	0.3310	0.1367	18	-0.1699	0.4167	0.2024
		67	-0.2067	0.3784	0.1858	-0.2053	0.3791	0.1857	22	-0.2198	0.4418	0.2433
100	45	33	-0.1744	0.3366	0.1436	-0.1737	0.3406	0.1461	22	-0.1968	0.4524	0.2432
		67	-0.1951	0.3850	0.1862	-0.1938	0.3870	0.1872	24	-0.2182	0.4966	0.2939
100	62	33	-0.0995	0.3647	0.1427	-0.0987	0.3711	0.1474	25	-0.1191	0.5377	0.3031
		67	-0.1356	0.4097	0.1861	-0.1332	0.4158	0.1904	17	-0.1510	0.5583	0.3342
500	10	33	-0.0404	0.1459	0.0229	-0.0400	0.1463	0.0230	23	-0.0421	0.1664	0.0295
		67	-0.0439	0.1450	0.0229	-0.0434	0.1452	0.0230	16	-0.0458	0.1616	0.0282
500	28	33	-0.0654	0.1293	0.0210	-0.0655	0.1297	0.0211	23	-0.0681	0.1673	0.0326
		67	-0.0719	0.1475	0.0269	-0.0707	0.1479	0.0268	24	-0.0702	0.1855	0.0393
500	45	33	-0.0698	0.1529	0.0282	-0.0665	0.1575	0.0292	19	-0.0747	0.2189	0.0535
		67	-0.0768	0.1712	0.0352	-0.0758	0.1736	0.0358	24	-0.0798	0.2365	0.0622
500	62	33	-0.0119	0.1930	0.0373	-0.0117	0.1975	0.0391	20	-0.0271	0.3019	0.0918
		67	-0.0222	0.2096	0.0444	-0.0202	0.2140	0.0462	25	-0.0294	0.3194	0.1028

Table 4: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (c) $\pi(x,z) = \pi(z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	-0.1422	0.3309	0.1296	-0.1407	0.3321	0.1300	25	-0.1445	0.3699	0.1576
		67	-0.0658	0.3533	0.1291	-0.0655	0.3534	0.1291	19	-0.0704	0.3874	0.1549
100	28	33	-0.1831	0.3194	0.1354	-0.1825	0.3214	0.1365	22	-0.1814	0.4068	0.1982
		67	-0.1555	0.4027	0.1862	-0.1549	0.4039	0.1869	12	-0.1551	0.4575	0.2332
100	45	33	-0.1587	0.3198	0.1274	-0.1572	0.3224	0.1286	19	-0.1779	0.4439	0.2285
		67	-0.1650	0.3295	0.1357	-0.1656	0.3338	0.1388	21	-0.1921	0.4453	0.2349
100	62	33	-0.1001	0.3619	0.1409	-0.0977	0.3667	0.1439	16	-0.1329	0.5226	0.2904
		67	-0.1560	0.4734	0.2482	-0.1507	0.4782	0.2512	25	-0.1856	0.5995	0.3935
500	10	33	-0.0692	0.1500	0.0273	-0.0686	0.1504	0.0273	15	-0.0725	0.1726	0.0350
		67	-0.0363	0.1528	0.0246	-0.0371	0.1524	0.0246	10	-0.0378	0.1681	0.0296
500	28	33	-0.0738	0.1302	0.0224	-0.0743	0.1310	0.0227	19	-0.0815	0.1737	0.0368
		67	-0.0629	0.1526	0.0272	-0.0623	0.1537	0.0275	24	-0.0651	0.1831	0.0377
500	45	33	-0.0710	0.1526	0.0283	-0.0709	0.1550	0.0290	23	-0.0760	0.2134	0.0512
		67	-0.0849	0.1692	0.0358	-0.0859	0.1715	0.0368	24	-0.0904	0.2321	0.0620
500	62	33	-0.0068	0.1940	0.0376	-0.0137	0.1980	0.0394	16	-0.0066	0.3136	0.0983
		67	-0.0524	0.2188	0.0506	-0.0508	0.2213	0.0515	17	-0.0549	0.3221	0.1067

Table 5: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (d) $\pi(x,z)$ depends on both x and z

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	-0.1385	0.3614	0.1496	-0.1385	0.3606	0.1491	16	-0.1280	0.4013	0.1773
		67	-0.0839	0.3353	0.1194	-0.0831	0.3354	0.1193	23	-0.0890	0.3752	0.1486
100	28	33	-0.1586	0.3141	0.1237	-0.1577	0.3156	0.1244	24	-0.1668	0.4026	0.1897
		67	-0.2331	0.4082	0.2208	-0.2312	0.4104	0.2217	23	-0.2586	0.4764	0.2935
100	45	33	-0.2155	0.3417	0.1631	-0.2171	0.3480	0.1681	19	-0.2334	0.4608	0.2666
		67	-0.1832	0.3960	0.1903	-0.1799	0.3998	0.1920	25	-0.2112	0.5036	0.2980
100	62	33	-0.1057	0.3623	0.1423	-0.1073	0.3668	0.1459	25	-0.1330	0.5427	0.3119
		67	-0.1310	0.4253	0.1979	-0.1299	0.4317	0.2031	19	-0.1425	0.5822	0.3589
500	10	33	-0.0678	0.1582	0.0296	-0.0676	0.1577	0.0294	20	-0.0686	0.1761	0.0357
		67	-0.0377	0.1374	0.0203	-0.0374	0.1378	0.0204	22	-0.0416	0.1506	0.0244
500	28	33	-0.0634	0.1324	0.0215	-0.0631	0.1334	0.0218	18	-0.0676	0.1689	0.0331
		67	-0.0820	0.1574	0.0315	-0.0808	0.1569	0.0311	20	-0.0880	0.2018	0.0484
500	45	33	-0.0812	0.1553	0.0307	-0.0802	0.1547	0.0303	20	-0.0800	0.2248	0.0569
		67	-0.0675	0.1748	0.0351	-0.0675	0.1768	0.0358	21	-0.0709	0.2337	0.0596
500	62	33	-0.0088	0.1953	0.0382	-0.0121	0.1990	0.0397	24	-0.0095	0.2952	0.0871
		67	-0.0746	0.2262	0.0567	-0.0728	0.2294	0.0579	22	-0.0729	0.3163	0.1053

Table 6: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (a) $\pi(x,z) = \pi$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	0	0.1367	0.4799	0.2488	0.1378	0.4793	0.2484	21	0.1413	0.5370	0.3080
		33	0.1483	0.5192	0.2912	0.1469	0.5186	0.2902	20	0.1446	0.5654	0.3402
		67	0.1552	0.6057	0.3906	0.1548	0.6053	0.3899	11	0.1483	0.6345	0.4242
100	28	0	0.2073	0.4771	0.2704	0.2055	0.4803	0.2727	21	0.2267	0.5762	0.3831
		33	0.2128	0.4976	0.2926	0.2081	0.4981	0.2912	21	0.2137	0.6139	0.4221
		67	0.2655	0.5629	0.3870	0.2627	0.5642	0.3870	18	0.2822	0.6611	0.5162
100	45	0	0.1139	0.5266	0.2900	0.1101	0.5367	0.2999	16	0.1164	0.7353	0.5537
		33	0.1173	0.5265	0.2907	0.1169	0.5353	0.2999	21	0.1493	0.7385	0.5671
		67	0.1207	0.5396	0.3055	0.1211	0.5438	0.3101	24	0.1634	0.7661	0.6130
100	62	0	-0.3877	0.6031	0.5137	-0.3888	0.6189	0.5338	24	-0.4481	0.9796	1.1596
		33	-0.3959	0.6080	0.5261	-0.3947	0.6285	0.5504	17	-0.4781	0.9856	1.1990
		67	-0.4163	0.6364	0.5779	-0.4160	0.6493	0.5942	25	-0.4800	0.9721	1.1745
500	10	0	0.0584	0.2105	0.0477	0.0580	0.2095	0.0472	24	0.0528	0.2320	0.0566
		33	0.0552	0.2203	0.0515	0.0513	0.0545	0.0484	14	0.0617	0.0562	0.0586
		67	0.0662	0.2429	0.0633	0.0652	0.2423	0.0629	15	0.0687	0.2687	0.0769
500	28	0	0.0894	0.2003	0.0481	0.0873	0.2012	0.0481	25	0.0877	0.2504	0.0703
		33	0.0953	0.2060	0.0515	0.0943	0.2061	0.0513	17	0.1023	0.2493	0.0725
		67	0.1063	0.2242	0.0615	0.1050	0.2262	0.0622	25	0.1095	0.2723	0.0861
500	45	0	0.0505	0.2668	0.0737	0.0479	0.2706	0.0754	20	0.0551	0.3795	0.1469
		33	0.0512	0.2668	0.0737	0.0498	0.2705	0.0756	22	0.0592	0.3680	0.1388
		67	0.0565	0.2693	0.0756	0.0569	0.2731	0.0778	23	0.0632	0.4099	0.1719
500	62	0	-0.3061	0.3378	0.2077	-0.3024	0.3506	0.2142	18	-0.3215	0.5758	0.4346
		33	-0.3086	0.3380	0.2093	-0.3088	0.3479	0.2163	25	-0.3517	0.5651	0.4428
		67	-0.3199	0.3392	0.2173	-0.3190	0.3518	0.2254	21	-0.3612	0.5690	0.4539

Table 7: Bias,SD and MSE for the estimator of the slope θ_{20}. Case (b) $\pi(x,z) = \pi(x)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	0.1557	0.5273	0.3020	0.1540	0.5266	0.3007	22	0.1595	0.5756	0.3564
		67	0.1601	0.5532	0.3314	0.1562	0.5531	0.3300	18	0.1641	0.5975	0.3835
100	28	33	0.2230	0.5042	0.3037	0.2260	0.5046	0.3054	14	0.2220	0.6172	0.4299
		67	0.2658	0.5319	0.3533	0.2620	0.5339	0.3535	13	0.2778	0.6255	0.4680
100	45	33	0.1224	0.5331	0.2989	0.1275	0.5365	0.3037	25	0.1238	0.7455	0.5706
		67	0.1297	0.5509	0.3200	0.1247	0.5545	0.3227	24	0.1283	0.7570	0.5889
100	62	33	-0.3924	0.6119	0.5281	-0.3939	0.6260	0.5467	25	-0.4424	0.9826	1.1602
		67	-0.4120	0.6648	0.6112	-0.4112	0.6840	0.6365	24	-0.4851	1.0277	1.2905
500	10	33	0.0620	0.2313	0.0573	0.0614	0.2317	0.0574	23	0.0641	0.2561	0.0696
		67	0.0668	0.2341	0.0592	0.0660	0.2340	0.0590	13	0.0689	0.2528	0.0686
500	28	33	0.0933	0.2067	0.0514	0.0936	0.2070	0.0516	23	0.0944	0.2556	0.0742
		67	0.1004	0.2188	0.0579	0.0985	0.2197	0.0579	24	0.0975	0.2737	0.0843
500	45	33	0.0509	0.2659	0.0732	0.0499	0.2704	0.0755	22	0.0480	0.3848	0.1503
		67	0.0532	0.2667	0.0739	0.0501	0.2718	0.0763	19	0.0482	0.3793	0.1460
500	62	33	-0.3107	0.3346	0.2084	-0.3071	0.3469	0.2145	21	-0.3300	0.5588	0.4209
		67	-0.3242	0.3357	0.2177	-0.3200	0.3528	0.2267	15	-0.3512	0.5981	0.4807

Table 8: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (c) $\pi(x,z) = \pi(z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	0.2099	0.5306	0.3253	0.2080	0.5308	0.3247	25	0.2128	0.5770	0.3779
		67	0.1043	0.5533	0.3167	0.1052	0.5530	0.3166	22	0.1109	0.5946	0.3655
100	28	33	0.2421	0.4893	0.2978	0.2418	0.4905	0.2989	22	0.2468	0.6180	0.4424
		67	0.2084	0.5557	0.3519	0.2086	0.5577	0.3543	23	0.2074	0.6500	0.4651
100	45	33	0.1110	0.5220	0.2845	0.1083	0.5258	0.2880	19	0.1087	0.7355	0.5523
		67	0.1144	0.5324	0.2962	0.1157	0.5387	0.3033	21	0.1296	0.7309	0.5506
100	62	33	-0.3916	0.6036	0.5174	-0.3915	0.6156	0.5319	25	-0.4551	0.9414	1.0925
		67	-0.4323	0.6580	0.6194	-0.4307	0.6803	0.6479	21	-0.4931	1.0222	1.2870
500	10	33	0.1024	0.2404	0.0682	0.1017	0.2404	0.0681	15	0.1078	0.2690	0.0839
		67	0.0561	0.2409	0.0611	0.0572	0.2400	0.0608	10	0.0582	0.2585	0.0701
500	28	33	0.1024	0.2072	0.0534	0.1032	0.2087	0.0542	19	0.1143	0.2639	0.0826
		67	0.0906	0.2205	0.0568	0.0903	0.2221	0.0574	24	0.0951	0.2667	0.0801
500	45	33	0.0524	0.2670	0.0740	0.0520	0.2701	0.0756	23	0.0586	0.3755	0.1443
		67	0.0579	0.2657	0.0739	0.0589	0.2699	0.0762	24	0.0654	0.3866	0.1535
500	62	33	-0.3071	0.3400	0.2098	-0.2980	0.3499	0.2111	16	-0.3450	0.5922	0.4694
		67	-0.3555	0.3290	0.2345	-0.3529	0.3433	0.2423	25	-0.3982	0.5874	0.5033

Table 9: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (d) $\pi(x,z)$ depending on both x and z

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	Bias	SD	MSE	Bias	SD	MSE		Bias	SD	MSE
100	10	33	0.2052	0.5708	0.3675	0.2053	0.5688	0.3653	16	0.1889	0.6234	0.4240
		67	0.1292	0.5412	0.3093	0.1283	0.5412	0.3091	23	0.1363	0.5916	0.3682
100	28	33	0.2142	0.4853	0.2812	0.2126	0.4889	0.2840	24	0.2187	0.6111	0.4209
		67	0.2927	0.5607	0.3998	0.2911	0.5636	0.4020	23	0.3272	0.6621	0.5450
100	45	33	0.1430	0.5224	0.2931	0.1375	0.5355	0.3054	23	0.1307	0.7117	0.5230
		67	0.1497	0.5694	0.3463	0.1437	0.5758	0.3519	25	0.1545	0.7917	0.6500
100	62	33	-0.3953	0.6099	0.5279	-0.3936	0.6196	0.5385	24	-0.4583	0.9790	1.1676
		67	-0.4169	0.6362	0.5781	-0.4174	0.6575	0.6061	24	-0.4979	0.9909	1.2287
500	10	33	0.1005	0.2469	0.0710	0.1004	0.2460	0.0705	20	0.1017	0.2692	0.0827
		67	0.0584	0.2229	0.0530	0.0571	0.2235	0.0532	11	0.0636	0.2368	0.0601
500	28	33	0.0908	0.2088	0.0518	0.0899	0.2096	0.0520	18	0.0963	0.2608	0.0772
		67	0.1136	0.2235	0.0628	0.1123	0.2235	0.0625	20	0.1230	0.2840	0.0957
500	45	33	0.0573	0.2624	0.0721	0.0562	0.2617	0.0716	20	0.0467	0.3889	0.1533
		67	0.0566	0.2700	0.0760	0.0570	0.2742	0.0784	21	0.0602	0.3735	0.1430
500	62	33	-0.3080	0.3383	0.2092	-0.3044	0.3488	0.2142	24	-0.3496	0.5572	0.4324
		67	-0.3600	0.3180	0.2306	-0.3534	0.3336	0.2360	24	-0.4243	0.5458	0.4776

Table 10: Bias,SD and MSE for Kaplan-Meier based estimator of the intercept θ_{10} and the slope θ_{20}

		θ_{10}			θ_{20}		
n	CP(%)	Bias	SD	MSE	Bias	SD	MSE
100	10	-0.0865	0.2936	0.0936	0.1303	0.4879	0.2548
	28	-0.1406	0.3319	0.1298	0.1884	0.5538	0.3419
	45	-0.1707	0.4381	0.2209	0.1306	0.7704	0.6100
	62	-0.0993	0.5189	0.2789	-0.4193	0.9795	1.1342
500	10	-0.0330	0.1289	0.0177	0.0508	0.2160	0.0492
	28	-0.0564	0.1461	0.0245	0.0773	0.2455	0.0662
	45	-0.0723	0.1995	0.0450	0.0632	0.3673	0.1388
	62	-0.0087	0.2880	0.0830	-0.3258	0.5596	0.4190

The comparison between the Kaplan-Meier estimator and the various estimators adapted to missigness is of interest. Note that, with light censoring, the MSE of the Kaplan-Meier (0% of missingness) is in general smaller than that corresponding to the estimators working under a 33% or a 67% of missingness. This is of course intuitive. But the opposite occurs when we move to the heavily censored case; then, the presmoothed and the multiple imputation methods outperform the Kaplan-Meier estimator, even when they work with fewer known values of δ. This is because the informative censorship mechanism which is introduced through the parametric specification of the function $m(x, z)$ (see e.g. Dikta, 1998, 2000) compensates for the loosing of information on the censoring indicators.

In order to investigate the robustness of the proposed method against miss-specifications in the parametrization of the function m, we considered a second simulated scenario in which C follows a uniform distribution on the $(0, 3)$ interval. As discussed in de Uña-Álvarez and Rodríguez-Campos (2004), this automatically leads to a function $m(x, z)$ which does not belong to the logistic family. The results obtained (shown in Appendix A) were very similar to those corresponding to the Weibull censoring. The MSE of the slope estimators greatly increased in the heavily censored case, a fact that could be related to the relative difficulties of the Kaplan-Meier estimator in the new scenario, see Table 20 (note that, due to the changing in the censoring distribution, the value of θ_{20} needs to be adjusted in order to maintain a given censoring level). In these cases with high censoring degree, the presmoothed and the multiple imputation estimators again outperformed the Kaplan-Meier based estimator. The proportion of cases in which multiple imputation reduced the MSE of the presmoothed estimator was 25%, all of them corresponding to light censoring (10%). Again, the optimal number of imputations varied largely. The conclusion is that the proposed estimator $\theta_n(m_n)$ for the regression parameter θ may be quite robust against miss-specifications of the function m, and under moderate to heavy censoring it performs at least as well as the multiple imputation method.

4 Real data illustration

For illustration purposes, we have considered the PBC data set reported and widely explained in Fleming and Harrington (1991), with $n = 312$ individuals. In this example, the Y variable denotes survival time (in years) for primary biliary cirrhosis (PBC) patients. Censoring from the right is provoked by the end of following-up or by liver transplantation (187 censored times, or about 60% of censoring). We assumed a loglinear regression model on the bilirubin level, taken in log scale (the X variable), that is

$$E[\ln Y \mid X] = \theta_{10} + \theta_{20} X.$$

Table 11 reports the (randomly weighted) least squares estimators for $(\theta_{10}, \theta_{20})$ *via* ordinary Kaplan-Meier, parametric presmoothed given by (2), and multiple imputation. Note that for this data set we have no missing censoring indicator; hence, we need to introduce some missingness in an artificial way. This is interesting, since in this manner we can see if the estimators adapted to the missing censoring indicator model provide values similar to those corresponding to the case with completely observed censoring indicators. Given (X_i, Z_i), the missingness indicator ξ_i was drawn from a Bernoulli model with parameter $\pi(X_i, Z_i)$, where the function $\pi(x, z)$ was chosen in four different ways: (a), (b), (c), (d) cases explained in Section 3, according to the type of dependence on the (x, z) variables. In Table 11, labels 1-4 for presmoothed (P) and multiple imputation (MI) methods refer to these four scenarios. A logistic specification for $\pi(x, z)$ was used in all the situations, with the parameters fixed to give approximately 33% and also 67% of missing proportion.

For the estimation of $m(x, z)$, the logistic specification (4) was assumed; some preliminary goodness-of-fit testing for the logistic model was performed from the complete data, and the Hosmer-Lemeshow statistic (see Hosmer and Lemeshow, 1989) reported a p-value of 0.902. For multiple imputation, we considered a maximum number of imputations of $K = 25$. We only report the minimum and the maximum estimates provided by the 25 multiple imputation values. Both parametric presmoothing and multiple imputation methods were applied also to the complete data (0% of missingness); this makes sense, since it is known that presmoothing provides variance reduction with respect to Kaplan-Meier weights (e.g. de Uña-Álvarez and Rodríguez-Campos, 2004).

From Table 11 we see that the estimates of the intercept θ_{10} from the complete data range from 1.7261 to 2.1876, and that presmoothing and multiple imputation leads to a slight modification of the value reported by the Kaplan-Meier weighted LSE (1.9308). The same happens for the slope θ_{20}, for which the Kaplan-Meier based estimator gives -0.4909 and the modified Kaplan-Meier weights give values ranging from -0.6782 to -0.4620. When considering the cases with missing censoring indicators, we see that neither the missingness proportion nor the possible dependencies on X and/or Z in $\pi(X, Z)$ seem to provoke significant departures of the estimates with respect to the complete data case.

Table 11: Estimators of intercept and slope for the PBC data,based on ordinary Kaplan-Meier(KM), presmoothed(P1-P4), and multiple imputation (MI1-MI4) weights

	θ_{10} (intercept)			θ_{20} (logbilirubin)		
MP(%)	0%	33%	67%	0%	33%	67%
KM	1.9308	_	_	-0.4909	_	_
P1	2.0482	1.9952	2.0379	-0.5691	-0.5719	-0.5686
MI1	1.7261	1.7216	1.9301	-0.6782	-0.6121	-0.6212
	2.1876	2.1520	2.1861	-0.4620	-0.4690	-0.5255
P2	_	2.0508	2.0204	_	-0.5627	-0.5658
MI2	_	1.9377	1.9113	_	-0.5873	-0.6078
	_	2.1070	2.1526	_	-0.4935	-0.5126
P3	_	2.0506	2.0540	_	-0.5718	-0.5685
MI3	_	1.8749	1.9796	_	-0.6167	-0.6329
	_	2.1982	2.1738	_	-0.4963	-0.5314
P4	_	2.1276	2.0176	_	-0.5786	-0.5687
MI4	_	2.0588	1.9294	_	-0.6376	-0.6130
	_	2.1920	2.0979	_	-0.5422	-0.5114

Finally, we mention that all the estimates for the slope support previous statistical analysis for the PBC data set; that is, high bilirubin levels are associated to poor survival prognosis.

5 Main conclusions

In this work a new estimator of the regression parameters in the censored accelerated failure time model has been introduced. The estimator is constructed by using some preliminary estimation (or presmoothing) of the conditional probability of censoring given the observed time and the covariate vector. The method proposed is adapted to the missing censoring indicators situation. The strong consistency of the introduced estimator has been formally established. Besides, a simulation study to investigate the finite sample performance of the estimator has been conducted. An alternative estimator based on a multiple imputation strategy has been considered in the simulations too. One of the main conclusions of our study is that the presmoothed estimator may perform well even when some miss-specification occurs. Besides, for moderate to large censoring proportions it performs better than the multiple imputation method.

It should be noted that preliminary presmoothing may be performed through nonparametric (e.g. kernel) methods too. To this regard, we mention that our consistency result applies provided that the uniform consistency assumption (U) is fulfilled. Finally, it would be interesting to derive the asymptotic distribution of the presmoothed estimator. This topic is currently under investigation.

Acknowledgements. Work supported by the Grants MTM2008-03129 of the Spanish Ministerio de Ciencia e Innovacin and 10PXIB300068PR of the

Xunta de Galicia. Financial support from the INBIOMED project (DXPCT-SUG, Ref. 2009/063) is also acknowledged.

6 References

Chen P, He R, Shen J-S, Sun J-G (2009) Regression analysis of right-censored failure time data with missing censoring indicators. Acta Mathematicae Applicatae Sinica, English Series 25, 415-426.

de Uña-Álvarez J, Rodríguez-Campos C (2004) Strong consistency of presmoothed Kaplan-Meier integrals when covariables are present. Statistics 38, 483-496.

Dikta G (1998) On semiparametric random censorship models. Journal of Statistical Planning and Inference 66, 253-279.

Dikta G (2000) The strong law under semiparametric random censorship models. Journal of Statistical Planning and Inference 83, 110.

Fleming TR, Harrington DP (1991) *Counting Processes and Survival Analysis.* Wiley.

Gijbels I, Lin D, Ying Z (1993) Non- and semi-parametric analysis of failure time data with missing failure indicators. Tech. Report 039-93, Mathematical Sciences Research Institute, Berkeley

Goetghebeur EJ, Ryan L (1995) Analysis of competing risks survival data when some failure types are missing. Biometrika 82, 821-833.

Hosmer, D. W. and Lemeshow, S. (1989). *Applied Logistic Regression.* Wiley, New York.

Lin YD, Ying Z (1994) Semiparametric analysis of the additive risk model. Biometrika 81, 61-71.

Lu K, Tsiatis AA (2001). Multiple imputation methods for estimating regression coefficients in proportional hazards models with missing cause of failure. Biometrics 57, 1191-1197.

McKeague IW, Subramanian, S (1998) Product-limit estimators and Cox regression with missing censoring information. Scandinavian Journal of Statistics 25, 589-601.

Stute, W. (1993). Consistent estimation under random censorship when covariables are present. Journal of Multivariate Analysis 45, 89-103.

Stute W (1999) Nonlinear censored regression. Statistica Sinica 9, 1089-1102.

Subramanian, S (2000) Efficient estimation of regression coefficients and baseline hazard under proportionality of conditional hazards. Journal of Statistical Planning and Inference 84, 81-94.

Subramanian S (2004) Asymptotically efficient estimation of a survival function in the missing censoring indicator model. Journal of Nonparametric Statistics 16, 797-817.

Subramanian S (2009) The multiple imputations based Kaplan-Meier estimator. Statistics & Probability Letters 79, 1906-1914.

van der Laan MJ, McKeague IW (1998) Efficient estimation from right-censored data when failure indicators are missing at random. Annals of Statistics 26, 164-182.

Wang Q, Ng KW (2008) Asymptotically efficient product-limit estimators with censoring indicators missing at random. Statistica Sinica 18, 749-768.

Wang Q, Shen J (2008) Estimation and confidence bands of a conditional survival function with censoring indicators missing at random. Journal of Multivariate Analysis 99, 928-948.

Zhou X, Sun L (2003) Additive hazards regression with missing censoring information. Statistica Sinica 13, 1237-1257.

A Appendix

Simulation results when the censoring variable follows a $U(0,3)$ distribution. In this case, the presmoothing function is miss-specified. Values of the slope θ_{20} to get censoring levels 10%, 28%, 45% and 62% change from 3.2, 0.3, -0.91, -2.25 (Weibull censoring) to 6.9, 2, 0.4, -1, respectively.

Table 12: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (a) $\pi(x,z)=\pi$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	0	-0.0870	0.2683	0.0795	-0.0863	0.2683	0.0793	21	-0.0817	0.2997	0.0964
		33	-0.0954	0.2810	0.0880	-0.0966	0.2806	0.0880	12	-0.0998	0.3151	0.1092
		67	-0.0926	0.3497	0.1307	-0.0930	0.3495	0.1307	15	-0.0966	0.3805	0.1540
100	28	0	-0.1254	0.2726	0.0900	-0.1282	0.2740	0.0914	24	-0.1503	0.3459	0.1421
		33	-0.1349	0.2816	0.0974	-0.1342	0.2830	0.0980	21	-0.1440	0.3557	0.1472
		67	-0.1417	0.3266	0.1266	-0.1384	0.3292	0.1274	22	-0.1476	0.4125	0.1918
100	45	0	-0.0822	0.2910	0.0914	-0.0820	0.2951	0.0938	24	-0.1014	0.4261	0.1917
		33	-0.0920	0.2988	0.0977	-0.0909	0.3016	0.0992	24	-0.1111	0.4156	0.1849
		67	-0.1038	0.3428	0.1282	-0.1041	0.3466	0.1309	22	-0.1250	0.4476	0.2158
100	62	0	0.0178	0.3102	0.0964	0.0182	0.3177	0.1012	21	0.0017	0.4661	0.2171
		33	0.0134	0.3203	0.1027	0.0142	0.3297	0.1088	20	-0.0080	0.4731	0.2236
		67	-0.0074	0.3685	0.1357	-0.0093	0.3741	0.1399	17	-0.0162	0.5000	0.2500
500	10	0	-0.0408	0.1152	0.0149	-0.0406	0.1151	0.0149	15	-0.0453	0.1277	0.0183
		33	-0.0408	0.1204	0.0162	-0.0402	0.1210	0.0162	15	-0.0401	0.1342	0.0196
		67	-0.0459	0.1342	0.0201	-0.0454	0.1334	0.0198	13	-0.0469	0.1508	0.0249
500	28	0	-0.0719	0.1270	0.0213	-0.0710	0.1282	0.0215	25	-0.0742	0.1688	0.0340
		33	-0.0732	0.1299	0.0222	-0.0718	0.1314	0.0224	10	-0.0771	0.1730	0.0358
		67	-0.0765	0.1389	0.0251	-0.0762	0.1392	0.0252	18	-0.0806	0.1826	0.0398
500	45	0	-0.0346	0.1426	0.0215	-0.0344	0.1466	0.0227	17	-0.0359	0.2296	0.0540
		33	-0.0343	0.1458	0.0224	-0.0364	0.1499	0.0238	24	-0.0397	0.2310	0.0549
		67	-0.0385	0.1583	0.0265	-0.0388	0.1602	0.0271	21	-0.0471	0.2533	0.0663
500	62	0	0.1016	0.1761	0.0413	0.0980	0.1846	0.0436	18	0.0965	0.3006	0.0996
		33	0.0985	0.1796	0.0419	0.0973	0.1841	0.0433	24	0.0828	0.2926	0.0924
		67	0.0939	0.1916	0.0455	0.0936	0.1931	0.0460	24	0.0688	0.2895	0.0885

Table 13: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (b) $\pi(x,z) = \pi(x)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	-0.0972	0.2963	0.0972	-0.0971	0.2956	0.0967	24	-0.0988	0.3297	0.1184
		67	-0.0950	0.3119	0.1062	-0.0941	0.3124	0.1064	19	-0.0945	0.3355	0.1214
100	28	33	-0.1328	0.2878	0.1004	-0.1349	0.2883	0.1012	21	-0.1483	0.3716	0.1599
		67	-0.1490	0.3160	0.1220	-0.1459	0.3192	0.1231	24	-0.1588	0.3908	0.1778
100	45	33	-0.0858	0.3029	0.0990	-0.0806	0.3068	0.1005	15	-0.1010	0.4275	0.1928
		67	-0.1043	0.3392	0.1258	-0.0997	0.3423	0.1270	24	-0.1343	0.4573	0.2270
100	62	33	0.0098	0.3271	0.1070	0.0103	0.3346	0.1119	25	0.0149	0.4780	0.2285
		67	-0.0001	0.3711	0.1376	-0.0016	0.3757	0.1410	20	-0.0335	0.4814	0.2327
500	10	33	-0.0421	0.1202	0.0162	-0.0415	0.1205	0.0162	21	-0.0434	0.1323	0.0194
		67	-0.0462	0.1275	0.0184	-0.0459	0.1277	0.0184	18	-0.0470	0.1434	0.0228
500	28	33	-0.0744	0.1300	0.0224	-0.0739	0.1318	0.0228	23	-0.0785	0.1709	0.0353
		67	-0.0736	0.1404	0.0251	-0.0727	0.1418	0.0254	21	-0.0799	0.1822	0.0396
500	45	33	-0.0369	0.1470	0.0229	-0.0374	0.1493	0.0237	17	-0.0468	0.2380	0.0588
		67	-0.0359	0.1618	0.0274	-0.0366	0.1661	0.0289	25	-0.0434	0.2479	0.0633
500	62	33	0.1011	0.1814	0.0431	0.1004	0.1861	0.0447	25	0.0847	0.2968	0.0952
		67	0.0962	0.1932	0.0465	0.0925	0.1994	0.0483	25	0.0961	0.3107	0.1057

Table 14: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (c) $\pi(x,z) = \pi(z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	-0.1398	0.3226	0.1235	-0.1385	0.3228	0.1233	24	-0.1422	0.3476	0.1409
		67	-0.0697	0.3200	0.1072	-0.0688	0.3195	0.1067	11	-0.0762	0.3449	0.1246
100	28	33	-0.1304	0.2831	0.0971	-0.1302	0.2850	0.0981	24	-0.1401	0.3633	0.1515
		67	-0.1902	0.3556	0.1625	-0.1903	0.3560	0.1628	25	-0.2044	0.4222	0.2199
100	45	33	-0.0923	0.3011	0.0991	-0.0908	0.3063	0.1020	15	-0.1150	0.4355	0.2027
		67	-0.1084	0.3450	0.1307	-0.1063	0.3481	0.1324	22	-0.1319	0.4443	0.2146
100	62	33	0.0062	0.3279	0.1074	0.0068	0.3325	0.1105	23	-0.0146	0.4763	0.2269
		67	-0.0163	0.3893	0.1517	-0.0126	0.3955	0.1564	25	-0.0399	0.5188	0.2705
500	10	33	-0.0502	0.1309	0.0196	-0.0507	0.1310	0.0197	19	-0.0486	0.1449	0.0233
		67	-0.0482	0.1301	0.0192	-0.0481	0.1299	0.0192	18	-0.0477	0.1448	0.0232
500	28	33	-0.0778	0.1315	0.0233	-0.0778	0.1328	0.0237	25	-0.0838	0.1725	0.0367
		67	-0.0760	0.1379	0.0248	-0.0756	0.1391	0.0250	25	-0.0779	0.1849	0.0402
500	45	33	-0.0434	0.1469	0.0235	-0.0416	0.1507	0.0244	25	-0.0430	0.2498	0.0642
		67	-0.0220	0.1529	0.0238	-0.0223	0.1562	0.0249	24	-0.0262	0.2373	0.0570
500	62	33	0.0883	0.1789	0.0398	0.0868	0.1842	0.0414	25	0.0760	0.2940	0.0921
		67	0.1161	0.1938	0.0510	0.1120	0.2017	0.0532	11	0.1101	0.2901	0.0962

Table 15: Bias,SD and MSE for the estimators of the intercept θ_{10}. Case (d) $\pi(x,z) = \pi(x,z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	-0.0812	0.2851	0.0878	-0.0800	0.2850	0.0876	25	-0.0805	0.3196	0.1085
		67	-0.1417	0.3542	0.1454	-0.1417	0.3536	0.1450	17	-0.1463	0.3786	0.1646
100	28	33	-0.1337	0.2840	0.0985	-0.1345	0.2862	0.0999	21	-0.1458	0.3589	0.1500
		67	-0.1649	0.3463	0.1470	-0.1652	0.3479	0.1482	23	-0.1576	0.3963	0.1817
100	45	33	-0.0942	0.3009	0.0993	-0.0957	0.3041	0.1015	23	-0.1272	0.4113	0.1852
		67	-0.1133	0.3427	0.1302	-0.1165	0.3456	0.1329	15	-0.1176	0.4422	0.2092
100	62	33	0.0051	0.3290	0.1082	0.0063	0.3342	0.1116	23	-0.0264	0.4714	0.2227
		67	0.0002	0.3596	0.1292	0.0024	0.3620	0.1309	20	-0.0157	0.4871	0.2373
500	10	33	-0.0424	0.1220	0.0167	-0.0424	0.1221	0.0167	22	-0.0403	0.1359	0.0201
		67	-0.0465	0.1364	0.0208	-0.0463	0.1362	0.0207	16	-0.0469	0.1491	0.0244
500	28	33	-0.0768	0.1302	0.0228	-0.0769	0.1317	0.0232	18	-0.0763	0.1688	0.0343
		67	-0.0766	0.1417	0.0259	-0.0775	0.1425	0.0263	23	-0.0812	0.1840	0.0404
500	45	33	-0.0417	0.1469	0.0233	-0.0418	0.1477	0.0235	24	-0.0515	0.2436	0.0619
		67	-0.0297	0.1556	0.0251	-0.0313	0.1578	0.0258	25	-0.0381	0.2330	0.0557
500	62	33	0.0914	0.1782	0.0401	0.0910	0.1829	0.0417	24	0.0866	0.2978	0.0961
		67	0.1107	0.1960	0.0506	0.1118	0.1993	0.0522	24	0.0965	0.2973	0.0976

Table 16: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (a) $\pi(x,z) = \pi$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP (%)	MP (%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	0	0.1234	0.4551	0.2222	0.1231	0.4550	0.2220	25	0.1173	0.4929	0.2565
		33	0.1346	0.4639	0.2331	0.1358	0.4629	0.2325	12	0.1399	0.5064	0.2758
		67	0.1348	0.5299	0.2986	0.1322	0.5293	0.2973	4	0.1423	0.5685	0.3431
100	28	0	0.0647	0.4686	0.2235	0.0698	0.4720	0.2275	24	0.0968	0.5878	0.3546
		33	0.0679	0.4725	0.2276	0.0660	0.4753	0.2300	24	0.0685	0.5974	0.3612
		67	0.0655	0.5065	0.2605	0.0580	0.5093	0.2625	22	0.0632	0.6628	0.4428
100	45	0	-0.3962	0.4905	0.3973	-0.3962	0.5032	0.4099	17	-0.4330	0.7861	0.8048
		33	-0.4057	0.5038	0.4182	-0.4068	0.5161	0.4316	24	-0.4287	0.7826	0.7956
		67	-0.4196	0.5500	0.4783	-0.4111	0.5664	0.4895	23	-0.4643	0.8069	0.8660
100	62	0	-1.4086	0.5479	2.2839	-1.4125	0.5798	2.3309	20	-1.5351	0.9455	3.2497
		33	-1.4173	0.5754	2.3395	-1.4209	0.5986	2.3769	25	-1.5508	0.9594	3.3245
		67	-1.4772	0.6802	2.6442	-1.4729	0.6978	2.6559	22	-1.6346	1.0519	3.7773
500	10	0	0.0516	0.2030	0.0438	0.0515	0.2028	0.0438	21	0.0580	0.2161	0.0500
		33	0.0519	0.2082	0.0460	0.0517	0.2082	0.0460	18	0.0498	0.2268	0.0539
		67	0.0594	0.2188	0.0514	0.0590	0.2175	0.0507	13	0.0604	0.2388	0.0606
500	28	0	0.0363	0.2288	0.0536	0.0354	0.2306	0.0544	23	0.0370	0.3007	0.0917
		33	0.0363	0.2293	0.0539	0.0320	0.2316	0.0546	10	0.0387	0.3030	0.0932
		67	0.0367	0.2336	0.0559	0.0358	0.2357	0.0568	18	0.0400	0.3134	0.0997
500	45	0	-0.3370	0.2447	0.1734	-0.3348	0.2579	0.1785	25	-0.3655	0.4291	0.3175
		33	-0.3375	0.2473	0.1750	-0.3356	0.2550	0.1776	24	-0.3485	0.4227	0.3000
		67	-0.3400	0.2589	0.1825	-0.3371	0.2658	0.1842	21	-0.3477	0.4522	0.3251
500	62	0	-1.3436	0.3018	1.8962	-1.3452	0.3242	1.9144	24	-1.4350	0.5839	2.4000
		33	-1.3489	0.3021	1.9108	-1.3475	0.3253	1.9216	21	-1.4392	0.5668	2.3923
		67	-1.3596	0.3163	1.9485	-1.3579	0.3426	1.9610	17	-1.4135	0.5776	2.3313

Table 17: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (b) $\pi(x,z) = \pi(x)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	0.1382	0.4810	0.2502	0.1376	0.4801	0.2492	24	0.1407	0.5218	0.2918
		67	0.1371	0.4930	0.2616	0.1357	0.4929	0.2611	19	0.1385	0.5208	0.2902
100	28	33	0.0677	0.4812	0.2359	0.0697	0.4820	0.2370	21	0.0750	0.6252	0.3961
		67	0.0683	0.5116	0.2661	0.0705	0.5160	0.2710	22	0.0697	0.6460	0.4217
100	45	33	-0.4032	0.4990	0.4113	-0.4032	0.5084	0.4207	25	-0.4313	0.7706	0.7792
		67	-0.4330	0.5542	0.4944	-0.4313	0.5675	0.5077	23	-0.4340	0.8171	0.8554
100	62	33	-1.4246	0.5550	2.3373	-1.4287	0.5726	2.3686	24	-1.6308	0.9778	3.6148
		67	-1.4615	0.7502	2.6982	-1.4697	0.7624	2.7406	22	-1.5611	1.0960	3.6371
500	10	33	0.0533	0.2070	0.0457	0.0527	0.2071	0.0456	21	0.0549	0.2206	0.0516
		67	0.0593	0.2130	0.0489	0.0591	0.2131	0.0488	9	0.0597	0.2319	0.0573
500	28	33	0.0362	0.2302	0.0543	0.0363	0.2329	0.0555	21	0.0385	0.2980	0.0902
		67	0.0359	0.2320	0.0551	0.0363	0.2346	0.0563	16	0.0444	0.3088	0.0972
500	45	33	-0.3396	0.2474	0.1765	-0.3378	0.2570	0.1801	21	-0.3527	0.4198	0.3004
		67	-0.3408	0.2580	0.1827	-0.3378	0.2707	0.1873	23	-0.3507	0.4451	0.3209
500	62	33	-1.3473	0.3036	1.9072	-1.3506	0.3166	1.9242	25	-1.3943	0.5781	2.2780
		67	-1.3567	0.3168	1.9410	-1.3532	0.3357	1.9437	25	-1.4634	0.6070	2.5097

Table 18: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (c) $\pi(x,z) = \pi(z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP(%)	MP(%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	0.1939	0.5100	0.2974	0.1952	0.5080	0.2959	21	0.1968	0.5376	0.3274
		67	0.1026	0.5000	0.2603	0.1028	0.4992	0.2595	13	0.1109	0.5360	0.2993
100	28	33	0.0670	0.4756	0.2305	0.0669	0.4785	0.2332	24	0.0622	0.6150	0.3817
		67	0.0924	0.5159	0.2744	0.0912	0.5175	0.2759	24	0.0995	0.6449	0.4254
100	45	33	-0.4100	0.5017	0.4196	-0.4057	0.5088	0.4233	23	-0.4288	0.7657	0.7696
		67	-0.3968	0.5330	0.4412	-0.3892	0.5485	0.4521	19	-0.4112	0.7543	0.7376
100	62	33	-1.4359	0.5714	2.3880	-1.4427	0.6063	2.4486	21	-1.5924	1.0107	3.5562
		67	-1.4574	0.6654	2.5664	-1.4587	0.6836	2.5945	21	-1.5686	1.0114	3.4823
500	10	33	0.0655	0.2174	0.0515	0.0649	0.2168	0.0512	19	0.0630	0.2332	0.0583
		67	0.0606	0.2168	0.0506	0.0606	0.2163	0.0504	18	0.0586	0.2356	0.0589
500	28	33	0.0375	0.2298	0.0542	0.0377	0.2318	0.0551	23	0.0437	0.3073	0.0963
		67	0.0399	0.2298	0.0544	0.0387	0.2324	0.0554	24	0.0341	0.3118	0.0983
500	45	33	-0.3434	0.2469	0.1788	-0.3410	0.2562	0.1819	22	-0.3865	0.4575	0.3585
		67	-0.3321	0.2469	0.1712	-0.3295	0.2577	0.1749	22	-0.3441	0.4186	0.2934
500	62	33	-1.3588	0.3042	1.9389	-1.3584	0.3177	1.9461	25	-1.4413	0.5850	2.4192
		67	-1.3325	0.3032	1.8675	-1.3355	0.3216	1.8868	25	-1.4029	0.5485	2.2686

Table 19: Bias,SD and MSE for the estimators of the slope θ_{20}. Case (d) $\pi(x,z) = \pi(x,z)$

			Presmoothed estimator			**Best Multiple Imputation**			**Best K**	**Worst Multiple Imputation**		
n	CP (%)	MP (%)	BIAS	SD	MSE	BIAS	SD	MSE		BIAS	SD	MSE
100	10	33	0.1160	0.4700	0.2342	0.1146	0.4695	0.2334	25	0.1125	0.5140	0.2766
		67	0.1979	0.5350	0.3251	0.1980	0.5342	0.3243	16	0.2052	0.5663	0.3625
100	28	33	0.0687	0.4768	0.2318	0.0669	0.4797	0.2343	20	0.0714	0.6066	0.3727
		67	0.0731	0.5143	0.2696	0.0759	0.5175	0.2734	12	0.0441	0.6188	0.3845
100	45	33	-0.4083	0.4982	0.4146	-0.4005	0.5137	0.4240	18	-0.3957	0.7610	0.7351
		67	-0.4137	0.5388	0.4612	-0.4067	0.5456	0.4628	24	-0.4601	0.7786	0.8173
100	62	33	-1.4348	0.5731	2.3869	-1.4429	0.6034	2.4457	21	-1.5418	0.9938	3.3639
		67	-1.4502	0.6717	2.5538	-1.4516	0.7038	2.6019	18	-1.5593	1.0540	3.5411
500	10	33	0.0537	0.2088	0.0464	0.0539	0.2085	0.0463	20	0.0508	0.2244	0.0529
		67	0.0611	0.2223	0.0531	0.0608	0.2213	0.0526	16	0.0610	0.2389	0.0607
500	28	33	0.0369	0.2307	0.0546	0.0365	0.2333	0.0557	18	0.0319	0.3057	0.0944
		67	0.0405	0.2342	0.0564	0.0417	0.2344	0.0566	23	0.0432	0.3022	0.0931
500	45	33	-0.3424	0.2467	0.1780	-0.3438	0.2500	0.1806	24	-0.3558	0.4397	0.3198
		67	-0.3323	0.2487	0.1722	-0.3302	0.2532	0.1731	25	-0.3436	0.4291	0.3020
500	62	33	-1.3563	0.3065	1.9335	-1.3518	0.3416	1.9440	13	-1.4383	0.5768	2.4011
		67	-1.3443	0.3105	1.9034	-1.3395	0.3288	1.9022	19	-1.3936	0.5518	2.2461

Table 20: Bias,SD and MSE for Kaplan-Meier based estimator of the intercept θ_{10} and the slope θ_{20}

		θ_{10}			θ_{20}		
n	CP(%)	BIAS	SD	MSE	BIAS	SD	MSE
100	10	-0.0823	0.2756	0.0826	0.1178	0.4698	0.2344
	28	-0.1479	0.3294	0.1303	0.0901	0.5849	0.3499
	45	-0.1150	0.3566	0.1403	-0.4552	0.6567	0.6380
	62	-0.0201	0.3658	0.1341	-1.6104	0.7707	3.1868
500	10	-0.0466	0.1220	0.0171	0.0610	0.2132	0.0491
	28	-0.1077	0.1560	0.0359	0.0561	0.2765	0.0795
	45	-0.0854	0.1741	0.0376	-0.4539	0.3275	0.3132
	62	0.0122	0.1921	0.0370	-1.5688	0.3982	2.6196