T E C H N I C A L
R E P O R T

**0209**

# BOOTSTRAPPING THE CONDITIONAL SURVIVAL ESTIMATOR IN THE PARTIAL KOZIOL-GREEN MODEL

R. BRAEKERS and N. VERAVERBEKE

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

# Bootstrapping the conditional survival function estimator in the partial Koziol-Green model

**Roel Braekers and Noël Veraverbeke**
Limburgs Universitair Centrum, B-3590 Diepenbeek, Belgium

Keywords: Bootstrap; Fixed design; Informative censoring; Nonparametric regression; Proportional hazards; Right censoring; Weak convergence
Running head: bootstrap partial Koziol-Green model
AMS Subject classifications: Primary 62G09; secondary 62G05, 62G20

**ABSTRACT**

In the partial Koziol-Green regression model, the lifetime variable may be censored by two types of censoring variables. One is called informative because it satisfies the Koziol-Green assumption on proportionality of hazards, and the other one is general. [3] proposed a nonparametric estimator for the conditional lifetime distribution and obtained a Gaussian approximation for the corresponding process. In the present paper we propose an appropriate resampling scheme and show that this leads to a valid bootstrap approximation for the process.

# 1 Introduction

Consider a regression model with independent and non-negative responses $Y_1, \ldots, Y_n$ at fixed design points $0 \leq x_1 \leq \ldots \leq x_n \leq 1$. These responses are subject to random right censoring in the following way: each $Y_i$ can be censored by the minimum of two independent and non-negative variables $C_i$ and $D_i$. The $C_i$ are informative censoring times, satisfying the proportional hazards condition, while the $D_i$ are arbitrary non informative censoring times. The observed random variables at design point $x_i$ are $(Z_i, \delta_i)$ where $Z_i = Y_i \wedge C_i \wedge D_i$ and $\delta_i = 1$ if $Y_i \leq C_i \wedge D_i$, $\delta_i = 0$ if $C_i \leq Y_i \wedge D_i$ and $\delta_i = -1$ if $D_i \leq Y_i \wedge C_i$. For a given design value $x \in ]0, 1[$, we write $F_x, G_{1x}, G_{2x}, H_x$ for the distribution function of respectively, the response $Y_x$ at $x$, the informative censoring variable $C_x$, the non informative censoring variable $D_x$ and $Z_x = Y_x \wedge C_x \wedge D_x$. Also we will write $\delta_x$ for the indicator variable that takes the values $1, 0, -1$ according as $Z_x = Y_x, C_x, D_x$. (Note that for the design values $x_i$ we write $Y_i, C_i, D_i, Z_i, \delta_i$ instead of $Y_{x_i}, C_{x_i}, D_{x_i}, Z_{x_i}, \delta_{x_i}$). We will assume the usual conditions of independence of the $Y_i$, the $C_i$ and the $D_i$, and also between $Y_i$, $C_i$ and $D_i$. The proportional hazards assumption on the censoring variable $C_x$ can be expressed as follows: for $t \geq 0$,

$$1 - G_{1x}(t) = (1 - F_x(t))^{\beta_x} \tag{1.1}$$

for some $\beta_x > 0$.
We will assume that the distribution functions $F_x$ and $G_{2x}$ are continuous. It then follows by (1.1) that $G_{1x}$ is also continuous. We also assume that the probabilities $p_{xk} = P(\delta_x = k)$ (k = -1,0,1) are strictly between 0 and 1 for all $x$. An easy calculation in [3] shows that $\beta_x = \dfrac{p_{x0}}{p_{x1}}$.

In that same paper we introduced the estimator $\widehat{F}_{xh}(t)$ for $F_x(t)$ in the following way:

$$1 - \widehat{F}_{xh}(t) = (1 - L_{xh}(t))^{\gamma_{xh}} \tag{1.2}$$

where

$$L_{xh}(t) = 1 - \left\{ \prod_{Z_{(i)} \leq t} \left( 1 - \frac{w_{n(i)}(x; h_n)}{1 - \sum\limits_{j=1}^{i-1} w_{n(j)}(x; h_n)} \right)^{\varepsilon_{(i)}} \right\} I(t < Z_{(n)}). \tag{1.3}$$

and

$$\gamma_{xh} = \frac{p_{x1h}}{p_{x0h} + p_{x1h}}. \tag{1.4}$$

with $p_{xkh} = \sum\limits_{i=1}^{n} w_{ni}(x; h_n) I(\delta_i = k)$ (k = -1,0,1) and $\varepsilon_i = I(\delta_i \neq -1)$. Here $Z_{(1)} \leq \ldots \leq Z_{(n)}$ are the ordered $Z_1, \ldots, Z_n$ and the $\varepsilon_{(i)}$ and $w_{n(i)}(x; h_n)$ are the corresponding $\varepsilon_i$ and $w_{ni}(x; h_n)$ of the ordered $Z_i$.

These estimators involve a sequence of smoothing weights $\{w_{ni}(x; h_n)\}$, depending on a positive bandwidth sequence $\{h_n\}$, tending to zero as $n \to \infty$. In our present case of fixed design points it is customary to take the Gasser-Müller type weights, given by

$$\begin{aligned} w_{ni}(x; h_n) &= \frac{1}{c_n(x; h_n)} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x - z}{h_n}\right) dz \qquad (i = 1, \ldots, n) \\ c_n(x; h_n) &= \int_0^{x_n} \frac{1}{h_n} K\left(\frac{x - z}{h_n}\right) dz. \end{aligned}$$

Here $x_0 = 0$ and $K$ is a known probability density function (kernel).

For simplicity we have used the same kernel K for estimation of $L_{xh}(t)$ and $p_{xkh}$. The extension to different kernels is of course straightforward.

## 2 The bootstrap procedure

The purpose of this paper is to propose a bootstrap method for approximating the distribution function of $(nh_n)^{1/2}(\widehat{F}_{xh}(t) - F_x(t))$. This will provide an alternative for the normal approximation derived in Theorem 3 of [3]. The normal approximation suffers from complicated expressions for the asymptotic mean and variance (see (5.1) and (5.2) below). For the Kaplan Meier estimator the bootstrap has been introduced by [4] and for the bootstrapped Beran estimator we refer to [8].

Given the design points $x_i$, the minima $Z_i$ and indicators $\delta_i$ $(i = 1, \ldots, n)$, we propose the following

**Resampling procedure**

Resample $(Z_i^*, \delta_i^*)$ $(i = 1, \ldots, n)$ with replacement from $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$, giving probability $w_{nj}(x_i; g_n)$ to $(Z_j, \delta_j)$, $j = 1, \ldots, n$. The sequence $\{g_n\}$ is a pilot bandwidth sequence which is different from $\{h_n\}$.

It is easily verified that the above procedure is equivalent to two alternative procedures.

**Alternative procedure 1**

Define the random variables $Y_i^*$, $C_i^*$, $D_i^*$ (independently) as follows:

$$\begin{array}{lll} Y_1^*, \ldots, Y_n^* & \text{are independent;} & Y_i^* \sim \widehat{F}_{x_i g} \\ C_1^*, \ldots, C_n^* & \text{are independent;} & C_i^* \sim \widehat{G}_{1 x_i g} \\ D_1^*, \ldots, D_n^* & \text{are independent;} & D_i^* \sim G_{2 x_i g} \end{array}$$

Here $\widehat{F}_{x_i g}$ is an estimator for $F_{x_i}$ as defined in (1.2). The distribution $\widehat{G}_{1x_i g}$ is defined as $1 - (1 - \widehat{F}_{x_i g})^{\beta_{x_i g}}$ where $\beta_{x_i g} = (1 - \gamma_{x_i g})/\gamma_{x_i g}$. Finally $G_{2x_i g}$ is a Beran type estimator for the distribution $G_{2x_i}$. It is given by

$$G_{2x_ig}(t) = 1 - \left\{ \prod_{Z_{(j)} \le t} \left( 1 - \frac{w_{n(j)}(x_i; g_n)}{1 - \sum_{k=1}^{j-1} w_{n(k)}(x_i; g_n)} \right)^{I(\delta_{(j)}=-1)} \right\} I(t < Z_{(n)}).$$

Then define, for $i = 1, \dots, n$, $Z_i^* = Y_i^* \wedge C_i^* \wedge D_i^*$ and $\delta_i^* = 1$ if $Y_i^* \le C_i^* \wedge D_i^*$, $\delta_i^* = 0$ if $C_i^* \le Y_i^* \wedge D_i^*$, $\delta_i^* = -1$ if $D_i^* \le Y_i^* \wedge C_i^*$.

**Alternative procedure 2**

Another equivalent resampling procedure can be set up by exploiting the following characterization of the model: assumption (1.1) holds for some $\beta_x > 0$ if and only if $Z_x$ and $\delta_x$ are independent, conditionally on the event $\{\delta_x \neq -1\}$. (see [3]).
Now define the random variables $U_i^*$, $\varepsilon_i^*$, $D_i^*$ (independently) as follows:

$$\begin{array}{rll} U_1^*, \dots, U_n^* & \text{are independent;} & U_i^* \sim L_{x_ig} \\ \varepsilon_1^*, \dots, \varepsilon_n^* & \text{are independent;} & \varepsilon_i^* \sim \text{Bin}(1, \gamma_{x_ig}) \\ D_1^*, \dots, D_n^* & \text{are independent;} & D_i^* \sim G_{2x_ig} \end{array}$$

where $L_{x_ig}$ is the estimator of $L_{x_i}$, as defined in (1.3). $G_{2x_ig}$ has the same definition as in the previous procedure. The $\varepsilon_i^*$ come from a Bernoulli distribution with parameter $\gamma_{x_ig}$, which is defined as in (1.4) and is an estimator for $\gamma_{x_i}$. The bootstrap sample can now be defined, for $i = 1, \dots, n$, as: $Z_i^* = U_i^* \wedge D_i^*$ and $\delta_i^* = \varepsilon_i^*$ if $U_i^* \le D_i^*$ and $\delta_i^* = -1$ if $U_i^* > D_i^*$.

Based on the bootstrap sample $(Z_1^*, \delta_1^*), \dots, (Z_n^*, \delta_n^*)$, we define the bootstrap analogue of the partial Koziol-Green estimator given in (1.2) as:

$$1 - \widehat{F}_{xhg}^*(t) = (1 - L_{xhg}^*(t))^{\gamma_{xhg}^*}$$

where

$$L_{xhg}^*(t) = 1 - \left\{ \prod_{Z_{(i)}^* \le t} \left( 1 - \frac{w_{n(i)}(x; h_n)}{1 - \sum_{j=1}^{i-1} w_{n(j)}(x; h_n)} \right)^{\varepsilon_{(i)}^*} \right\} I(t < Z_{(n)}^*) \text{ and } \gamma_{xhg}^* = \frac{p_{x1hg}^*}{p_{x0hg}^* + p_{x1hg}^*}$$

with $p_{xkhg}^* = \sum_{i=1}^{n} w_{ni}(x; h_n) I(\delta_i^* = k)$ (k = -1,0,1) and $\varepsilon_i^* = I(\delta_i^* \neq -1)$. Here $Z_{(1)}^* \le \dots \le Z_{(n)}^*$, and the $\varepsilon_{(i)}^*$ and $w_{n(i)}(x; h_n)$ correspond to $Z_{(i)}^*$.

In the sequel we will use notations $P^*$, $E^*$, $\text{Var}^*$, ... for probability, expectation, variance, ... conditionally on the original observations.

The paper is organized as follows. After some notation and regularity conditions in Section 3, we obtain results on an almost sure asymptotic representation in Section 4 and on weak convergence of the bootstrapped process in Section 5. A simulated example will be presented in Section 6. In the appendix we will give the proofs of the theorems and some lemmas which we need in these proofs.

# 3 Regularity conditions

We will need the following additional notation. For the design points $x_1, \ldots, x_n$ we denote $\underline{\Delta}_n = \min_{1\le i\le n}(x_i - x_{i-1})$ and $\overline{\Delta}_n = \max_{1\le i\le n}(x_i - x_{i-1})$. For the kernel $K$ we use $\|K\| = \sup_{u\in I\!R} K(u)$ , $\|K\|_2^2 = \int_{-\infty}^{+\infty} K^2(u)du$, $\mu_1^K = \int_{-\infty}^{+\infty} uK(u)du$ and $\mu_2^K = \int_{-\infty}^{+\infty} u^2K(u)du$. For any distribution function $\tilde{L}$, we denote by $T_{\tilde{L}} = \inf\{t : \tilde{L}(t) = 1\}$ the right endpoint of its support. Note that $T_{F_x} = T_{G_{1x}}$ and $T_{H_x} = T_{F_x} \wedge T_{G_{2x}}$.
We will use the following assumptions on the design points and on the kernel. (Condition (A2') below is slightly stronger than condition (A2) in [3]).

(A1) $x_n \to 1$ , $\overline{\Delta}_n = O(n^{-1})$ , $\overline{\Delta}_n - \underline{\Delta}_n = o(n^{-1})$.

(A2') $K$ is a twice differentiable probability density function with finite support $[-M, M]$ for some $M > 0$ , $\mu_1^K = 0$, $K''$ is continuous and $K(-M) = K'(-M) = K(M) = K'(M) = 0$.

Note that $c_n(x; h_n) = 1$ for $n$ sufficiently large (depending on $x$) since $x_n \to 1$ and $K$ has finite support. This makes that in all proofs of asymptotic results we may take $c_n(x; h_n) = 1$.

Further we will need typical smoothness conditions on functions like $F_x(t), G_{2x}(t)$ (as functions of $x$ and $t$) and $\beta_x$ and $p_{xk}$ (as functions of $x$).

(A3) The second order partial derivatives $\ddot{F}_x(t) = \frac{\partial^2}{\partial x^2}F_x(t), \ddot{G}_{2x}(t) = \frac{\partial^2}{\partial x^2}G_{2x}(t)$, $F''_x(t) = \frac{\partial^2}{\partial t^2}F_x(t), G''_{2x}(t) = \frac{\partial^2}{\partial t^2}G_{2x}(t), \dot{F}'_x(t) = \frac{\partial^2}{\partial x \partial t}F_x(t), \dot{G}'_{2x}(t) = \frac{\partial^2}{\partial x \partial t}G_{2x}(t)$ exist and are continuous for $0 \le x \le 1$ and $0 \le t \le T$ with $T < T_{F_x} \wedge T_{G_{2x}}$.

(A4) The second order partial derivatives $\ddot{p}_{xk} = \frac{\partial^2}{\partial x^2} p_{xk}$ $(k = -1, 0, 1)$ exist and are continuous for $0 \leq x \leq 1$.

The first order partial derivatives will be denoted by $\dot{F}_x(t), F'_x(t), \dot{G}_{2x}(t), G'_{2x}(t), \dot{\beta}_x, \dot{p}_{xk}$. Further notations are $\|\dot{p}_{xk}\| = \sup_{0\leq x\leq 1} |\dot{p}_{xk}|$ , $\|\ddot{p}_{xk}\| = \sup_{0\leq x\leq 1} |\ddot{p}_{xk}|$.

The parameter $g_n$ which is used to construct the resampled values is an appropriate pilot bandwidth sequence which is typically asymptotically larger than $h_n$, i.e. $g_n/h_n \to \infty$ in a certain way. This technique of oversmoothing with the initial bandwidth has been successfully used in other resampling schemes in regression (e.g. [1], [5], [8]). Furthermore we will state our results for the fixed bandwidth sequence $h_n$ of optimal rate, i.e. $h_n = Cn^{-1/5}$ for some constant $C > 0$.

# 4 Almost sure asymptotic representation

Parallel to the result of [3] for the conditional estimator $\widehat{F}_{xh}(t)$, we now derive in Theorem 1 an almost sure asymptotic representation for the bootstrap estimator $\widehat{F}^*_{xhg}(t)$. It represents the process $\widehat{F}^*_{xhg}(t) - \widehat{F}_{xh}(t)$ as a weighted sum plus a remainder term with a nice almost sure behaviour. This result will be a key tool in the next section.

Let us introduce the functions $H_x(t) = \mathrm{P}(Z_x \leq t)$ and $H^{0,1}_x(t) = \mathrm{P}(Z_x \leq t, \varepsilon_x = 1) = \mathrm{P}(Z_x \leq t, \delta_x \neq -1)$.

**Theorem 1.** Assume (A1), (A2'), (A3), (A4) in $[0,T]$ with $T < T_{H_x}$, $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to +\infty$, $\frac{ng_n^5}{\log n} \cdot \frac{h_n}{g_n} = O(1)$. Then, for $t < T_{H_x}$,

$$\widehat{F}^*_{xhg}(t) - \widehat{F}_{xg}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n) g_{tx}(Z^*_i, \delta^*_i) - \sum_{i=1}^{n} w_{ni}(x; g_n) g_{tx}(Z_i, \delta_i) + R_n(x,t)$$

where

$$\begin{aligned} g_{tx}(Z_i,\delta_i) \quad = \quad & \gamma_x(1-F_x(t)) \left\{ \int_0^t \frac{I(Z_i \leq s) - H_x(s)}{(1-H_x(s))^2} dH^{0,1}_x(s) \right. \\ & + \frac{I(Z_i \leq t, \delta_i \neq -1) - H^{0,1}_x(t)}{1-H_x(t)} - \int_0^t \frac{I(Z_i \leq s, \delta_i \neq -1) - H^{0,1}_x(s)}{(1-H_x(s))^2} dH_x(s) \\ & \left. - \frac{1}{p^2_{x1}} \log(1-F_x(t))[p_{x0} I(\delta_i = 1) - p_{x1} I(\delta_i = 0)] \right\} \end{aligned} \quad (4.1)$$

and where, for $T < T_{Hx}$, $\sup_{0\le t\le T} |R_n(x,t)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4})$ *a.s.*

The proof will be given in Section 7.1.

# 5 Weak convergence

We establish weak convergence to a Gaussian process of the process $(nh_n)^{1/2}(\widehat{F}^*_{xhg}(\cdot) - \widehat{F}_{xg}(\cdot))$ in the space $D[0,T]$ of right continuous functions with left hand limits, endowed with the Skorokhod topology.

We introduce the notation $\tilde{W}_x(\cdot)$ for a Gaussian process with mean function $\beta_x(t)$ and covariance function $\sigma_{st}$ given by

$$\begin{aligned}\beta_x(t) &= \frac{1}{2}\mu_2^K \gamma_x (1 - F_x(t)) \left\{ \int_0^t \left[ \frac{\ddot{H}_x(s) dH_x^{0,1}(s)}{(1 - H_x(s))^2} + \frac{d\ddot{H}_x^{0,1}(s)}{1 - H_x(s)} \right] \right. \\ & \left. - \frac{1}{p_{x1}^2} \log(1 - F_x(t)) [p_{x0}\ddot{p}_{x1} - p_{x1}\ddot{p}_{x0}] \right\} C^{5/2} \qquad (5.1)\end{aligned}$$

where $C > 0$ is some constant and

$$\begin{aligned}\sigma_{st} &= \|K\|_2^2 \gamma_x^2 (1 - F_x(t))(1 - F_x(s)) \left\{ \int_0^{s\wedge t} \frac{dH_x^{0,1}(y)}{(1 - H_x(y))^2} \right. \\ & \left. + \log(1 - F_x(s)) \log(1 - F_x(t)) \frac{p_{x0}(p_{x0} + p_{x1})}{p_{x1}^3} \right\}. \qquad (5.2)\end{aligned}$$

**Theorem 2.** Assume (A1), (A2'), (A3), (A4) in $[0,T]$ with $T < T_{H_x}$, $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$, $\frac{ng_n^5}{\log n} \to +\infty$, $\frac{ng_n^5}{\log n} \cdot \frac{h_n}{g_n} = O(1)$. Then, for $T < T_{H_x}$, as $n \to \infty$,

$$(nh_n)^{1/2}(\widehat{F}^*_{xhg}(\cdot) - \widehat{F}_{xg}(\cdot)) \to \tilde{W}_x(\cdot) \qquad \text{in } D[0,T]$$

The proof of this theorem is given in Section 7.2.

Since $\tilde{W}_x(\cdot)$ above is also the limiting Gaussian process of $(nh_n)^{1/2}(\widehat{F}_{xh}(\cdot) - F_x(\cdot))$ (see Theorem 3 in [3]), we conclude that $\mathrm{P}^*((nh_n)^{1/2}(\widehat{F}^*_{xhg}(t) - \widehat{F}_{xg}(t)) \le y)$ is a strongly consistent estimator for $\mathrm{P}((nh_n)^{1/2}(\widehat{F}_{xh}(t) - F_x(t)) \le y)$.

# 6 Simulated example

The finite sample performance of the normal approximation ([3]) and the bootstrap approximation of the distribution of $(nh_n)^{1/2}(\widehat{F}_{xh}(t) - F_x(t))$ will now be compared through some simulations. Let us take fixed and equidistant design points $x_i = i/n$ $(i = 1, \ldots, n)$. We assume that the survival times $Y_i$ $(i = 1, \ldots, n)$ are independent random variables and that

$$Y_i \sim \text{Weibull}(a_1 + a_2 x_i, b),$$

i.e. $F_{x_i}(t) = 1 - \exp(-(a_1 + a_2 x_i)t^b)$, $(t \geq 0)$ for some constants $a_1$, $a_2$ and $b$ such that $a_1 > 0 \wedge (-a_2)$ and $b > 0$. Similarly the non informative censoring times $D_i$ $(i = 1, \ldots, n)$ are independent and for each i, $D_i \sim \text{Weibull}(c_1 + c_2 x_i, b)$ for some $c_1 > 0 \wedge (-c_2)$. Furthermore we assume that $\beta_{x_i} = \exp(a_3 + a_4 x_i)$ $(i = 1, \ldots, n)$. By relation (1.1), this gives informative censoring times $C_i$ $(i = 1, \ldots, n)$ with for each i, $C_i \sim \text{Weibull}((a_1 + a_2 x_i)\exp(a_3 + a_4 x_i), b)$. We also assume that $C_i$ $(i = 1, \ldots, n)$ are independent and that for each i, the survival time $Y_i$, the informative censoring time $C_i$ and the non informative censoring time $D_i$ are independent.

It is now easily verified that the probabilities of uncensored, informatively censored and non informatively censored observations in the data set are given by the following expressions

$$\begin{aligned}
p_{x1} &= \frac{a_1 + a_2 x}{(a_1 + a_2 x)\exp(a_3 + a_4 x) + c_1 + c_2 x}, \\
p_{x0} &= \frac{(a_1 + a_2 x)\exp(a_3 + a_4 x)}{(a_1 + a_2 x)\exp(a_3 + a_4 x) + c_1 + c_2 x}, \\
p_{x,-1} &= \frac{c_1 + c_2 x}{(a_1 + a_2 x)\exp(a_3 + a_4 x) + c_1 + c_2 x}.
\end{aligned}$$

We study two different combinations of the parameters $a_1$, $a_2$, $a_3$, $a_4$, $c_1$, $c_2$. Figures 1(a) and 2(a) show how the probabilities uncensored, informatively censored and non informatively censored observations change for the value of the covariate $x$. On Figure 1, we choose the parameters such that the probability of informative censoring is always smaller then the probability of uncensoring.

[place Figure 1 about here]

In the Koziol-Green model, it is also important to know how the estimates behave when the amount of informatively censored observations in the data set is larger than the amount of uncensored observations. The choices of parameters in Figure 2 is specially designed to obtain such a situation.

[place Figure 2 about here]

We carried out the simulations for samples of size $n = 50$, for a biquadratic kernel function $K(x) = (15/16)(1 - x^2)^2 \mathrm{I}(|x| \le 1)$ and for $x = 0.5$ and $t = 0.5$. The normal approximation and the true distribution were obtained using 10 000 samples. For the bootstrap approximation 200 samples were taken and for each sample, 200 resamples were drawn.

For the bandwidth $h_n$ we take the optimal choice $h_{n,opt} = Cn^{-1/5}$ for some specific $C > 0$. We do not have such an optimal choice for the pilot bandwidth $g_n$. The only assumption we made, was that the $g_n$ is asymptotically larger than $h_n$. In this situation we take $g_n = 0.5$ since we have only observations in a window of width 0.5 around $x = 0.5$.

Figure 1(b)-(d) and Figure 2(c)-(d) show that both the normal and the bootstrap approximation are very close to the true distribution. In both cases the censoring is quite heavy. We see in Figure 2(b) that the quality of the normal and the bootstrap approximation is not good. However when we consider a smaller time point ($t = 0.25$), the approximations will again be better (not shown here). In general the approximations are not good when we take a large or very small value for $t$. Also when the covariate value $x$ is close to the boundaries 0 or 1, the approximations will fail.

# 7 Appendix

### 7.1. Proof of almost sure asymptotic representation

First we give some lemmas which we need in the proof of Theorem 1. Let us introduce the empirical estimators $H_{xh}(t)$ and $H^{0,1}_{xh}(t)$ of $H_x(t)$ and $H^{0,1}_x(t)$:

$$H_{xh}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n)\mathrm{I}(Z_i \le t),\ \ H^{0,1}_{xh}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n)\mathrm{I}(Z_i \le t, \delta_i \ne -1)$$

as well as their bootstrap analogues $H^*_{xhg}(t)$ and $H^{*0,1}_{xhg}(t)$:

$$H^*_{xhg}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n)\mathrm{I}(Z^*_i \le t),\ \ H^{*0,1}_{xhg}(t) = \sum_{i=1}^{n} w_{ni}(x; h_n)\mathrm{I}(Z^*_i \le t, \delta^*_i \ne -1).$$

**Lemma 1.** Assume (A1), (A2') and (A4), $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$, $\dfrac{ng_n^5}{\log n} \to +\infty$. Then as $n \to +\infty$,

(a) $|\text{bias } p^*_{xkhg} - \text{bias } p_{xkh}| = o((nh_n)^{-1/2})$ *a.s.*
where bias $p^*_{xkhg} = E^*\ p^*_{xkhg} - p_{xkg}$ and bias $p_{xkh} = E\ p_{xkh} - p_{xk}$.

(b) $|p^*_{xkhg} - p_{xkg}| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(c) $|\gamma^*_{xhg} - \gamma_{xg}| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(d) If (A3) holds instead of (A4), then as $n \to +\infty$,
$\sup_{0 \le t \le T} |\text{bias } H^*_{xhg}(t) - \text{bias } H_{xh}(t)| = o((nh_n)^{-1/2})$ *a.s.*
where bias $H^*_{xhg}(t) = E^*\ H^*_{xhg}(t) - H_{xg}(t)$ and bias $H_{xh}(t) = E\ H_{xh}(t) - H_x(t)$.

**Proof.** (a) We have,

$$\text{bias } p^*_{xkhg} = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right)(p_{x_i kg} - p_{zkg})dz + \int_0^{x_n} \frac{1}{h_n} K\left(\frac{x-z}{h_n}\right)(p_{zkg} - p_{xkg})dz.$$

To the first term, we apply Lemma 8 of [1]. A second order Taylor series for the second term, gives that,

$$\text{bias } p^*_{xkhg} = \frac{\mu_2^K h_n^2 \ddot{p}_{xkg}}{2} + o(h_n^2) + O\left(\frac{1}{ng_n^2}\right)$$

where $\ddot{p}_{xkg} = \sum_{i=1}^{n} \ddot{w}_{ni}(x; g_n) I(\delta_i = k)$, $\ddot{w}_{nj}(x; g_n) = \dfrac{1}{g_n^3} \displaystyle\int_{x_{j-1}}^{x_j} K''\left(\frac{x-z}{h_n}\right) dz$.
The term bias $p_{xkh}$ can be split up as above and a first order Taylor series on the first and a second order Taylor series on the second, gives that

$$\text{bias } p_{xkh} = \frac{\mu_2^K h_n^2 \ddot{p}_{xk}}{2} + o(h_n^2).$$

This leads to,

$$(nh_n)^{1/2} |\text{bias } p^*_{xkhg} - \text{bias } p_{xkh}| = \frac{\mu_2^K C^{5/2}}{2} |\ddot{p}_{xkg} - \ddot{p}_{xk}| + o(1).$$

Now

$$|\ddot{p}_{xkg} - \ddot{p}_{xk}| \le |\ddot{p}_{xkg} - E\ \ddot{p}_{xkg}| + |E\ \ddot{p}_{xkg} - \ddot{p}_{xk}|.$$

The first term is $o(1)$, using Bernstein's theorem (see e.g. [6]). For the second term we have, for large n,

$$|E\ \ddot{p}_{xkg} - \ddot{p}_{xk}| \leq \left| \sum_{i=1}^{n} \ddot{w}_{ni}(x;g_n) p_{x_ik} - \frac{1}{g_n^3} \int_0^{x_n} K''\left(\frac{x-z}{h_n}\right) p_{zk} dz \right|$$

$$+ \left| \frac{1}{g_n^3} \int_0^{x_n} K''\left(\frac{x-z}{h_n}\right) p_{zk} dz - \ddot{p}_{xk} \right|$$

$$\leq \sum_{i=1}^{n} \frac{1}{g_n^3} \int_{x_{i-1}}^{x_i} \left| K''\left(\frac{x-z}{h_n}\right) \right| |p_{x_ik} - p_{zk}|\, dz + \left| \int_{-M}^{M} K(u)(\ddot{p}_{x-ug_n,k} - \ddot{p}_{xk}) dz \right|, \text{ which is } o(1).$$

(b) Now, $|p^*_{xkhg} - p_{xkg}| \leq |p^*_{xkhg} - E^*\ p^*_{xkhg}| + |\text{bias } p^*_{xkhg} - \text{bias } p_{xkh}| + |\text{bias } p_{xkh}|$. The second term is $o((nh_n)^{-1/2})$ *a.s.* by part (a) of this lemma. The third term is $O(h_n^2 + n^{-1})$ as was shown in the proof of Lemma A.1. in [3]. For the first term we apply Berstein's inequality where we take $\varepsilon = \varepsilon_n = D(nh_n)^{-1/2}(\log n)^{1/2}$ with $D$ a constant. By the Borel-Cantelli lemma we obtain the result.
(c) It is easily seen that

$$|\gamma^*_{xhg} - \gamma_{xg}| \leq \frac{1}{(p^*_{x1hg} + p^*_{x0hg})(p_{x1g} + p_{x0g})} \left\{ |p^*_{x1hg} - p_{x1g}| + |p^*_{x0hg} - p_{x0g}| \right\}.$$

By part (b) and Lemma A.1. of [3], we have that $p^*_{xkhg} \xrightarrow{P^*} p_{xk}$ and $p_{xkg} \xrightarrow{P} p_{xk}$. The result now follows from part (b) of this lemma.
(d) The proof of this result is along the same lines as the proof of Lemma 4.1. in [8] (although the resampling scheme is different). An analogous result holds for $H^{*0,1}_{xhg}(t)$.

**Lemma 2.** Assume (A1), (A2'), (A3), (A4), $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$ and $\dfrac{ng_n^5}{\log n} \cdot \dfrac{h_n}{g_n} = O(1)$, $T < T_{H_x}$. Then as $n \to +\infty$,

(a) $|p_{xkg} - p_{xk}| = O((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(b) $|\gamma_{xg} - \gamma_x| = O((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(c) $\sup\limits_{0 \leq t \leq T} |L_{xg}(t) - L_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(d) $\sup\limits_{0 \leq t \leq T} |\widehat{F}_{xg}(t) - F_x(t)| = O((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

If moreover $\dfrac{ng_n^5}{\log n} \to +\infty$, then as $n \to +\infty$,

(e) $\sup\limits_{0 \leq t \leq T} |-\log(1 - L^*_{xhg}(t)) + \log(1 - L_{xg}(t))| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

(f) $\sup\limits_{0 \leq t \leq T} |L^*_{xhg}(t) - L_{xg}(t)| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*

**Proof.** (a), (b), (d). In Lemma 1 (resp. Lemma 2 or Theorem 1) of [3], we put $\varepsilon_n = D(nh_n)^{-1/2}(\log n)^{1/2}$ and apply the Borel-Cantelli lemma.
(c), (e). We prove these results along the same lines as in [8].
(f) By a first order Taylor series, we find,

$$L^*_{xhg}(t) - L_{xg}(t) = (1 - L_{xg}(t))[-e^{\theta_n}(\log(1 - L^*_{xhg}(t)) - \log(1 - L_{xg}(t))]$$

where $\theta_n$ lies between 0 and $\log(1-L^*_{xhg}(t))-\log(1-L_{xg}(t))$. Since $e^{\theta_n} \leq \dfrac{1}{1 - L_{xg}(t)}$, we now obtain the result by parts (c) and (e).

**Proof of Theorem 1.** Application of a one term Taylor expansion gives

$$\begin{aligned} \widehat{F}^*_{xhg}(t) - \widehat{F}_{xg}(t) &= \gamma_{xg}(1 - L_{xg}(t))^{\gamma_{xg}-1}[L^*_{xhg}(t) - L_{xg}(t)] \\ &\quad -(1 - \widehat{F}_{xg}(t))\log(1 - L_{xg}(t))[\gamma^*_{xhg} - \gamma_{xg}] + R_{n1}(x,t) \quad (7.1) \end{aligned}$$

where

$$\begin{aligned} |R_{n1}(x,t)| &\leq (1 - \tilde{L}_{xg}(t))^{-2}(L^*_{xhg}(t) - L_{xg}(t))^2 + \tilde{\gamma}_{xg}^{-2}(\gamma^*_{xhg} - \gamma_{xg})^2 \\ &\quad +2(1 - \tilde{L}_{xg}(t))^{-1}|L^*_{xhg}(t) - L_{xg}(t)||\gamma^*_{xhg} - \gamma_{xg}| \end{aligned}$$

with $\tilde{L}_{xg}(t)$ between $L_{xg}(t)$ and $L^*_{xhg}(t)$ and $\tilde{\gamma}_{xg}$ between $\gamma_{xg}$ and $\gamma^*_{xhg}$.

For the remainder term, we find the following upperbound:

$$\begin{aligned} |R_{n1}(x,t)| &\leq \left[\frac{1}{(1 - L_x(t))^2} + \frac{(\tilde{L}_{xg}(t) - L_x(t))(2 - \tilde{L}_{xg}(t) - L_x(t))}{(1 - \tilde{L}_{xg}(t))^2(1 - L_x(t))^2}\right](L^*_{xhg}(t) - L_{xg}(t))^2 \\ &+ \left[\frac{2}{(1 - L_x(t))^2} + \frac{2(\tilde{L}_{xg}(t) - L_x(t))(2 - \tilde{L}_{xg}(t) - L_x(t))}{(1 - \tilde{L}_{xg}(t))^2(1 - L_x(t))^2}\right](L^*_{xhg}(t) - L_{xg}(t))(\gamma^*_{xhg} - \gamma_{xg}) \\ &+ \left[\frac{1}{\gamma_x^2} - \frac{\tilde{\gamma}_{xg}^2 - \gamma_x^2}{\tilde{\gamma}_{xg}^2\gamma_x^2}\right](\gamma^*_{xhg} - \gamma_{xg})^2. \end{aligned}$$

By Lemma 1 and 2, we find that $\tilde{L}_{xg}(T) \xrightarrow{P^*} L_x(T)$ and $\sup_{0\leq t\leq T} |\tilde{L}_{xg}(t) - L_x(t)| = O_{P^*}((nh_n)^{-1/2}(\log n)^{1/2})$ *a.s.*, so this results in $\sup_{0\leq t\leq T} |R_{n1}(x,t)| = O_{P^*}((nh_n)^{-1}\log n)$ *a.s.*

To obtain the asymptotic representation of $L^*_{xhg}(t) - L_{xg}(t)$, we use Theorem 4.1 of [8]:

$$\begin{aligned} L^*_{xhg}(t) - L_{xg}(t) &= (1 - L_{xg}(t))\left\{\int_0^t \frac{H^*_{xhg} - H_{xg}}{(1 - H_x)^2} dH^{0,1}_x \right. \\ &+ \left. \frac{H^{*0,1}_{xhg}(t) - H^{0,1}_{xg}(t)}{1 - H_x(t)} - \int_0^t \frac{H^{*0,1}_{xhg} - H^{0,1}_{xg}}{(1 - H_x)^2} dH_x \right\} + R_{n2}(x,t) \quad (7.2) \end{aligned}$$

with $\sup\limits_{0\le t\le T} |R_{n2}(x,t)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4}\ a.s.$

To find an asymptotic representation for $\gamma^*_{xhg} - \gamma_{xg}$ we apply a one term Taylor expansion and obtain

$$\begin{aligned}
\gamma^*_{xhg} - \gamma_{xg} &= \frac{p^*_{x1hg}}{p^*_{x1hg} + p^*_{x0hg}} - \frac{p_{x1g}}{p_{x1g} + p_{x0g}} \\
&= \frac{p_{x0g}}{(p_{x1g} + p_{x0g})^2}[p^*_{x1hg} - p_{x1g}] - \frac{p_{x1g}}{(p_{x1g} + p_{x0g})^2}[p^*_{x0hg} - p_{x0g}] + R_{n3}(x) \quad (7.3)
\end{aligned}$$

where

$$|R_{n3}(x)| \le \frac{2}{(\tilde{p}_{x1g} + \tilde{p}_{x0g})^2}\left[(p^*_{x1hg} - p_{x1g})^2 + (p^*_{x0hg} - p_{x0g})^2\right]$$

with $\tilde{p}_{x0g}$ between $p^*_{x0hg}$ and $p_{x0g}$ and $\tilde{p}_{x1g}$ between $p^*_{x1hg}$ and $p_{x1g}$.

By adding and subtracting the same terms, we find that:

$$\begin{aligned}
|R_{n3}(x)| &\le \frac{2}{(p_{x1} + p_{x0})^2}\left[(p^*_{x1hg} - p_{x1g})^2 + (p^*_{x0hg} - p_{x0g})^2\right] \\
&- \frac{2(\tilde{p}_{x1g} - p_{x1g} + p_{x1g} - p_{x1} + \tilde{p}_{x0g} - p_{x0g} + p_{x0g} - p_{x0})(\tilde{p}_{x1g} + \tilde{p}_{x0g} + p_{x1} + p_{x0})}{(\tilde{p}_{x1g} + \tilde{p}_{x0g})^2(p_{x1} + p_{x0})^2} \times \\
&\left[(p^*_{x1hg} - p_{x1g})^2 + (p^*_{x0hg} - p_{x0g})^2\right]
\end{aligned}$$

Lemma 1 and 2 imply that $|R_{n3}(x)| = O_{P^*}((nh_n)^{-1}\log n)$.

Combining the results (7.2) and (7.3) into (7.1) gives

$$\begin{aligned}
&\widehat{F}^*_{xhg}(t) - \widehat{F}_{xg}(t) \\
&= \gamma_{xg}(1-\widehat{F}_{xg}(t))\left\{\left(\int_0^t \frac{H^*_{xhg} - H_{xg}}{(1-H_x)^2}dH^{0,1}_x + \frac{H^{*0,1}_{xhg}(t) - H^{0,1}_{xg}(t)}{1-H_x(t)} - \int_0^t \frac{H^{*0,1}_{xhg} - H^{0,1}_{xg}}{(1-H_x)^2}dH_x\right)\right. \\
&\left. - \frac{1}{p^2_{x1g}}\log(1-\widehat{F}_{xg}(t))\left[p_{x0g}(p^*_{x1hg} - p_{x1g}) - p_{x1g}(p^*_{x0hg} - p_{x0g})\right]\right\} + R_n(x,t) \\
&= \sum_{i=1}^n w_{ni}(x;h_n)g_{tx}(Z^*_i,\delta^*_i) - \sum_{i=1}^n w_{ni}(x;g_n)g_{tx}(Z_i,\delta_i) + R_n(x,t)
\end{aligned}$$

where $g_{tx}$ is defined in (4.1) and $\sup\limits_{0\le t\le T} |R_n(x,t)| = O_{P^*}((nh_n)^{-3/4}(\log n)^{3/4})\ a.s.$

In the last equality we replaced $\gamma_{xg}$, $\widehat{F}_{xg}$, $p_{x0g}$, $p_{x1g}$ by $\gamma_x$, $F_x$, $p_{x0}$, $p_{x1}$ respectively. This is allowed by Lemma 2.

### 7.2. Proof of weak convergence

We first establish a lemma on the bootstrap bias which will be needed in the proof of Theorem 2.

**Lemma 3.** Assume (A1), (A2'), (A3) and (A4), $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$, $\dfrac{ng_n^5}{\log n} \to +\infty$ and $T < T_{H_x}$. Then as $n \to +\infty$,

$$(nh_n)^{1/2} \sup_{0 \le t \le T} |b_n^*(t|x) - b_n(t|x)| = o(1) \quad a.s.$$

where

$$\begin{aligned} b_n^*(t|x) &= \sum_{i=1}^n w_{ni}(x, h_n) E^* g_{tx}(Z_i^*, \delta_i^*) - \sum_{i=1}^n w_{ni}(x, g_n) g_{tx}(Z_i, \delta_i) \\ b_n(t|x) &= \sum_{i=1}^n w_{ni}(x, h_n) E g_{tx}(Z_i, \delta_i). \end{aligned}$$

**Proof.** A straightforward calculation gives

$$\begin{aligned} b_n^*(t|x) - b_n(t|x) = \gamma_x(1 - F_x(t)) \Bigg\{ &\int_0^t \frac{\text{bias } H_{xhg}^*(s) - \text{bias } H_{xh}(s)}{(1 - H_x(s))^2} dH_x^{0,1}(s) \\ &+ \frac{\text{bias } H_{xhg}^{*0,1}(t) - \text{bias } H_{xh}^{0,1}(t)}{1 - H_x(t)} - \int_0^t \frac{\text{bias } H_{xhg}^{*0,1}(s) - \text{bias } H_{xh}^{0,1}(s)}{(1 - H_x(s))^2} dH_x(s) \\ &- \frac{\log(1 - F_x(t))}{p_{x1}^2} \Big[ p_{x0}(\text{bias } p_{x1hg}^* - \text{bias } p_{x1h}) - p_{x1}(\text{bias } p_{x0hg}^* - \text{bias } p_{x0h}) \Big] \Bigg\}. \end{aligned}$$

The result now follows from Lemma 1. We also use the inequality $(1 - F_x(t)) \log(1 - F_x(t)) \le 1$.

**Lemma 4.** Assume (A1), (A2'), (A3) and (A4) in $[0, T]$ with $T < T_{H_x}$, $h_n = Cn^{-1/5}$ for some $C > 0$, $g_n \to 0$, $\dfrac{ng_n^5}{\log n} \to +\infty$ and $\dfrac{ng_n^5}{\log n} \cdot \dfrac{h_n}{g_n} = O(1)$. Then, as $n \to +\infty$,

(a) $(nh_n)^{1/2} \max\limits_{1 \le i \le n} \sup\limits_{0 \le t \le T} |H_{x_i g}(t) - H_{x_i}(t)| \cdot \sum\limits_{j=1}^n w_{nj}^2(x, h_n) = o(1) \quad a.s.$
The same holds for the subdistribution $H_{x_i}^{0,1}(t)$.

(b) $(nh_n)^{1/2} \max\limits_{1 \le i \le n} |p_{x_i gk} - p_{x_i k}| \cdot \sum\limits_{j=1}^n w_{nj}^2(x, h_n) = o(1) \quad a.s. \ (k = -1, 0, 1)$

**Proof.** We have

$$
\begin{aligned}
&P\left((nh_n)^{1/2}\max_{1\le i\le n}\sup_{0\le t\le T}|H_{x_ig}(t)-H_{x_i}(t)|\cdot\sum_{j=1}^n w_{nj}^2(x,h_n)>\varepsilon\right)\\
&\le P\left(\|K\|_2^2\max_{1\le i\le n}\sup_{0\le t\le T}|H_{x_ig}(t)-H_{x_i}(t)|>\varepsilon\right)\\
&\le \sum_{i=1}^n P\left(\sup_{0\le t\le T}|H_{x_ig}(t)-H_{x_i}(t)|>\frac{\varepsilon}{\|K\|_2^2}\right).
\end{aligned}
$$

In the same way, we find

$$
P\left((nh_n)^{1/2}\max_{1\le i\le n}|p_{x_igk}-p_{x_ik}|\cdot\sum_{j=1}^n w_{nj}^2(x,h_n)>\varepsilon\right)\le\sum_{i=1}^n P\left(|p_{x_igk}-p_{x_ik}|>\frac{\varepsilon}{\|K\|_2^2}\right).
$$

By Lemma A3 in [8] and Lemma A1 in [3], we get the results.

**Proof of Theorem 2.** By the asymptotic representation of Theorem 1, we can write

$$
\begin{aligned}
(nh_n)^{1/2}(\widehat{F}^*_{xhg}(t)-\widehat{F}_{xg}(t)) &= W^*_{hg}(t|x)+(nh_n)^{1/2}(b^*_n(t|x)-b_n(t|x))\\
&\quad+((nh_n)^{1/2}b_n(t|x)-\beta_x(t))+\beta_x(t)+R_n(x,t)
\end{aligned}
$$

where $W^*_{hg}(t|x)=(nh_n)^{1/2}\sum_{i=1}^n w_{ni}(x,h_n)\left[g_{tx}(Z_i^*,\delta_i^*)-E^*g_{tx}(Z_i^*,\delta_i^*)\right]$ and $\beta_x(t)$ is defined in (5.2).

In [3], we already found that $\sup_{0\le t\le T}|(nh_n)^{1/2}b_n(t|x)-\beta_x(t)|=o(1)$ *a.s.*. So taking Lemma 3 into account, we only need to prove that $W^*_{hg}(\cdot|x)$ converges weakly to $\tilde{W}_x(\cdot)$ in $D[0,T]$. This is done in two steps. First we will show convergence of the finite dimensional distributions and afterwards we will prove tightness. More in detail, this means that first we show that $(W^*_{hg}(t_1|x),\ldots,W^*_{hg}(t_q|x))\xrightarrow{D}N(0;\sigma_{t_jt_k})$ for any $q=1,2,\ldots$ and any $0\le t_1\le\ldots\le t_q\le T$. This is done by verifying the following two conditions of [2] on the summands $W_{nik}=(nh_n)^{1/2}w_{ni}(x,h_n)\left[g_{t_kx}(Z_i^*,\delta_i^*)-E^*g_{t_kx}(Z_i^*,\delta_i^*)\right]$:

$$
\lim_{n\to\infty}\sum_{i=1}^n E^*(W_{nij}W_{nik})=\sigma_{t_jt_k}\qquad(1\le j,k\le q)\tag{7.4}
$$

$$
\lim_{n\to\infty}\sum_{i=1}^n\int_{\{|W_{ni}|>\varepsilon\}}|W_{ni}|^2dP^*=0\tag{7.5}
$$

for every $\varepsilon>0$, where $|W_{ni}|^2=\sum_{k=1}^q W_{nik}^2$.

For (7.4) we have:

$$
\sum_{i=1}^n E^*(W_{nij}W_{nik})=nh_n\sum_{i=1}^n w_{ni}^2(x;h_n)\mathrm{Cov}^*(g_{t_jx}(Z_i^*,\delta_i^*),g_{t_kx}(Z_i^*,\delta_i^*))\tag{7.6}
$$

The right hand side of this expression can be calculated as in [3]. We get the same result, but with $H_{x_i}$, $H_{x_i}^{0,1}$, $p_{x_ik}$ replaced by $H_{xig}$, $H_{xig}^{0,1}$, $p_{xigk}$. With this and Lemma 4, we can write (7.6) as:

$$\sum_{i=1}^{n} E^*(W_{nij}W_{nik}) = \sigma_{t_jt_k} + o(1) \qquad (1 \le j,k \le q).$$

For (7.5) we first note that the functions $g_{t_kx}(Z_i^*,\delta_i^*) - E^* g_{t_kx}(Z_i^*,\delta_i^*)$ are bounded uniformly on $[0,T]$, so that $\max\limits_{1\le i\le n} |W_{ni}| = O_{P^*}((nh_n)^{-1/2})$, $\sum\limits_{i=1}^{n} |W_{ni}|^2 = O_{P^*}(1)$ and hence

$$\begin{aligned}\sum_{i=1}^{n} \int_{\{|W_{ni}|>\varepsilon\}} |W_{ni}|^2 dP^* &\le \int_{\{\max\limits_{1\le i\le n}|W_{ni}|>\varepsilon\}} \sum_{i=1}^{n} |W_{ni}|^2 dP^* \\ &\le O_{P^*}(1) P^*(\max_{1\le i\le n} |W_{ni}| > \varepsilon) = o_{P^*}(1).\end{aligned}$$

This proves the convergence of the finite dimensional distributions.

To prove tightness, we will check the conditions of Theorem 2.11.9 in [7] (the bracketing central limit theorem). Let us first denote
$W_{ni}(t) = (nh_n)^{1/2} w_{ni}(x,h_n) g_{tx}(Z_i^*,\delta_i^*)$ and $\mathcal{F} = [0,T]$. On $\mathcal{F}$ we define the semimetric $\rho(t,t')$ as

$$\rho(t,t') = \max\left\{ |H_x^{0,1}(t) - H_x^{0,1}(t')|,\ |F_x(t) - F_x(t')|,\ \sup_{x_i\in[0,1]} \sup_{n\in I\!N} |H_{xig}^{0,1}(t) - H_{xig}^{0,1}(t)|^{1/2} \right\}.$$

Before we can verify the conditions of Theorem 2.11.9, we need to divide $\mathcal{F}$ for every $n$ and $\varepsilon$, in a partition of subintervals $\{\mathcal{F}_{\varepsilon j}^n\}$ such that

$$\sum_{i=1}^{n} E^* \sup_{t,t'\in\mathcal{F}_{\varepsilon j}^n} |W_{ni}(t) - W_{ni}(t')|^2 \le \varepsilon^2. \tag{7.7}$$

The smallest number of subintervals in this partition such that (7.7) holds, is called the bracketing number $N_{[]}(\varepsilon,\mathcal{F},L_2^n)$. In our situation we will divide $\mathcal{F}$ in a partition $\{\mathcal{F}_{\varepsilon j}^n\}$ where $\rho(t,t') \le C\cdot\varepsilon$ for every $t,t' \in \mathcal{F}_{\varepsilon j}^n$ and $C$ a constant which we will specify further on. We first prove that (7.7) holds for this partition. In this way the bracketing number is $N_{[]}(\varepsilon,\mathcal{F},L_2^n) = O\left(\dfrac{1}{\varepsilon}\right)$.

Some easy calculation shows that

$$|W_{ni}(t) - W_{ni}(t')| = (nh_n)^{1/2} w_{ni}(x;h_n) |g_{tx}(Z_i^*,\delta_i^*) - g_{t'x}(Z_i^*,\delta_i^*)|$$

$$\leq (nh_n)^{1/2} w_{ni}(x;h_n)\gamma_x \left\{ \frac{|I(Z_i^* \leq t, \delta_i^* \neq -1) - I(Z_i^* \leq t', \delta_i^* \neq -1)|}{1 - H_x(T)} \right.$$

$$\left. + \left( \frac{2}{(1 - H_x(T))^2} + \frac{|-1 - \log(1 - F_x(T))|}{p_{x1}^2} \right) |F_x(t) - F_x(t')| + \frac{|H_x^{0,1}(t) - H_x^{0,1}(t')|}{(1 - H_x(T))^2} \right\}$$

and by the inequality of Cauchy-Schwarz, we get

$$|W_{ni}(t) - W_{ni}(t')|^2$$

$$\leq 3(nh_n) w_{ni}^2(x;h_n)\gamma_x^2 \left\{ \frac{|I(Z_i^* \leq t, \delta_i^* \neq -1) - I(Z_i^* \leq t', \delta_i^* \neq -1)|^2}{(1 - H_x(T))^2} \right.$$

$$\left. + \left( \frac{2}{(1 - H_x(T))^2} + \frac{|-1 - \log(1 - F_x(T))|}{p_{x1}^2} \right)^2 |F_x(t) - F_x(t')|^2 + \frac{|H_x^{0,1}(t) - H_x^{0,1}(t')|^2}{(1 - H_x(T))^4} \right\}.$$

Hence,

$$\sum_{i=1}^n E^* \sup_{t,t' \in \mathcal{F}_{\varepsilon j}^n} |W_{ni}(t) - W_{ni}(t')|^2 \leq 3(nh_n) \sum_{i=1}^n w_{ni}^2(x;h_n)\gamma_x^2 \times$$

$$\left\{ \frac{2}{(1 - H_x(T))^4} + \left( \frac{2}{(1 - H_x(T))^2} + \frac{|-1 - \log(1 - F_x(T))|}{p_{x1}^2} \right)^2 \right\} C^2 \varepsilon^2$$

$$\leq 3\|K\|_2^2 \gamma_x^2 \left\{ \frac{2}{(1 - H_x(T))^4} + \left( \frac{2}{(1 - H_x(T))^2} + \frac{|-1 - \log(1 - F_x(T))|}{p_{x1}^2} \right)^2 \right\} C^2 \varepsilon^2.$$

which is $\varepsilon^2$ for an appropriate choice of $C$. At this point it is now easy to verify the third condition of the bracketing central limit theorem, since we have that,

$$\int_0^{\delta_n} \sqrt{\log N_{[]}(\varepsilon, \mathcal{F}, L_2^n)} d\varepsilon = \int_0^{\delta_n} \sqrt{O\left(\frac{1}{\varepsilon}\right)} d\varepsilon$$

which tends to zero if $\delta_n \downarrow 0$. The inequalities above also give

$$\sum_{i=1}^n E^*(W_{ni}(t) - W_{ni}(t'))^2$$

$$\leq 3\|K\|_2^2 \gamma_x^2 \left\{ \frac{2}{(1 - H_x(T))^4} + \left( \frac{2}{(1 - H_x(T))^2} + \frac{|-1 - \log(1 - F_x(T))|}{p_{x1}^2} \right)^2 \right\} \rho(t,t')^2$$

so that $\sup_{\rho(t,t') < \delta_n} \sum_{i=1}^n E^*(W_{ni}(t) - W_{ni}(t'))^2 \to 0$ for every $\delta_n \downarrow 0$. This is the second condition of the bracketing central limit theorem. Since $W_{ni}(t)$ is bounded uniformly on $[0,T]$, we see that the first condition is easily satisfied. So by the bracketing central limit theorem, we have that the process $W_{hg}^*(\cdot|x)$ is asymptotically tight. Together with the convergence of the finite dimensional distributions, this concludes our proof.

## Acknowledgement

The authors acknowledge support from the Interuniversity Attration Poles Programme P5/24 – Belgian State – Federal Office for Scientific, Technical and Cultural Affairs.

## References

[1] M. Aerts, P. Janssen and N. Veraverbeke, Bootstrapping regression quantiles, *J. Nonparametric Statistics.* **4**, (1994), 1-20.

[2] A. Araujo and E. Giné, *The central limit theorem for real and Banach valued random variables*, Wiley, New York, 1980.

[3] R. Braekers and N. Veraverbeke, The partial Koziol-Green model with covariates, *J. Statist. Planning Inf.* **92**, (2001), 55-71.

[4] B. Efron, Censored data and the bootstrap, *J. Amer. Statist. Assoc.* **76**, (1981), 312-319.

[5] W. Härdle and E. Mammen, Bootstrap methods in nonparametric regression, *Nonparametric Functional Estimation and Related Topics* (ed. G. Roussas), Kluwer, Dordrecht, 1991, pp. 111-123.

[6] R. J. Serfling, *Approximation theorems of mathematical statistics*, Wiley, New York, 1980.

[7] A. van der Vaart and J. A. Wellner, *Weak convergence and empirical processes*, Springer, New York, 1996.

[8] I. Van Keilegom and N. Veraverbeke, Estimation and bootstrap with censored data in fixed design nonparametric regression, *Ann. Inst. Statist. Math.* **49**, (1997), 467-491.

Roel Braekers (roel.braekers@luc.ac.be)
Noël Veraverbeke (noel.veraverbeke@luc.ac.be)
Limburgs Universitair Centrum
Universitaire Campus
B-3590 Diepenbeek (Belgium)

Figure 1: (a) Proportion of uncensored, informatively censored and non informatively censored observations for $a_1 = 1,\ a_2 = 0.5,\ a_3 = -0.5,\ a_4 = 0.5,\ c_1 = 0.6,\ c_2 = 2.$ (b)-(d) Normal approximation (small dashed curve) and bootstrap approximation (dashed curve) of $P((nh_n)^{1/2}(\widehat{F}_{xh}(t) - F_x(t)) \leq y)$ (solid curve) for $a_1 = 1,\ a_2 = 0.5,\ a_3 = -0.5,\ a_4 = 0.5,\ c_1 = 0.6,\ c_2 = 2,\ x = 0.5,\ t = 0.5.$

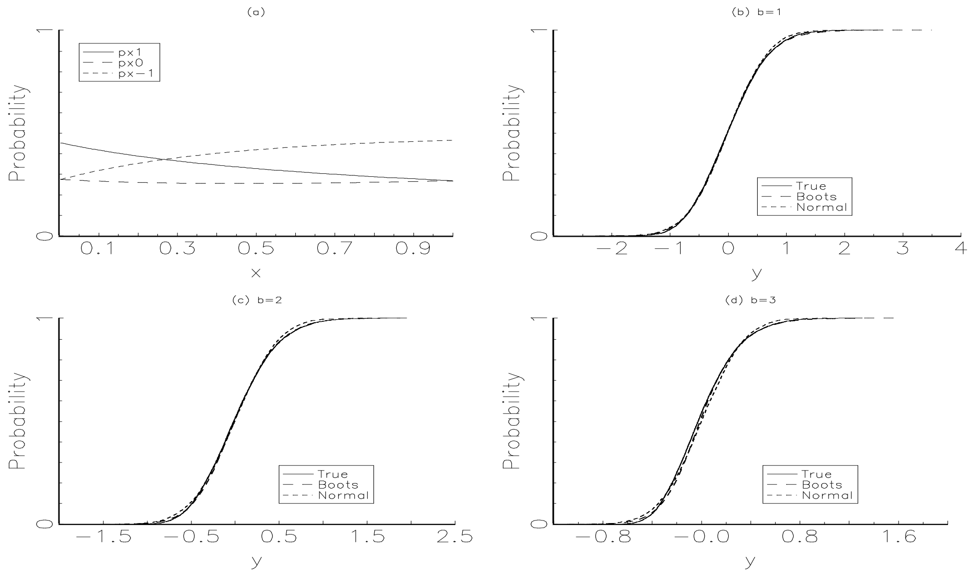

Figure 2: (a) Proportion of uncensored, informatively censored and non informatively censored observations for $a_1 = 1,\ a_2 = 2,\ a_3 = 0.5,\ a_4 = 0.5,\ c_1 = 1,\ c_2 = 1$.
(b)-(d) Normal approximation (small dashed curve) and bootstrap approximation (dashed curve) of $P((nh_n)^{1/2}(\widehat{F}_{xh}(t) - F_x(t)) \leq y)$ (solid curve) for $a_1 = 1,\ a_2 = 2,\ a_3 = 0.5,\ a_4 = 0.5,\ c_1 = 1,\ c_2 = 1,\ x = 0.5,\ t = 0.5$.