Tests of exogeneity in duration models with censored data

Gilles Crommen[1]
gilles.crommen@kuleuven.be

Jean-Pierre Florens[2]
jean-pierre.florens@tse-fr.eu

Ingrid Van Keilegom[1]
ingrid.vankeilegom@kuleuven.be

[1]ORSTAT, KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium
[2]Toulouse School of Economics, Esplanade de l'Université 1, 31080 Toulouse, France

Abstract

Consider the setting in which a researcher is interested in the causal effect of a treatment Z on a duration time T, which is subject to right censoring. We assume that $T = \varphi(X, Z, U)$, where X is a vector of baseline covariates, $\varphi(X, Z, U)$ is strictly increasing in the error term U for each (X, Z) and $U \sim \mathcal{U}[0, 1]$. Therefore, the model is nonparametric and nonseparable. We propose nonparametric tests for the hypothesis that Z is exogenous, meaning that Z is independent of U given X. The test statistics rely on an instrumental variable W that is independent of U given X. We assume that X, W and Z are all categorical. Test statistics are constructed for the hypothesis that the conditional rank $V_T = F_{T|X,Z}(T \mid X, Z)$ is independent of (X, W) jointly. Under an identifiability condition on φ, this hypothesis is equivalent to Z being exogenous. However, note that V_T is censored by $V_C = F_{T|X,Z}(C \mid X, Z)$, which complicates the construction of the test statistics significantly. We derive the limiting distributions of the proposed tests and prove that our estimator of the distribution of V_T converges to the uniform distribution at a rate faster than the usual parametric $n^{-1/2}$-rate. We demonstrate that the test statistics and bootstrap approximations for the critical values have a good finite sample performance in various Monte Carlo settings. Finally, we illustrate the tests with an empirical application to the National Job Training Partnership Act (JTPA) Study.

1 Introduction

We consider the setting in which a researcher is interested in identifying the causal effect of a treatment variable Z on a duration of interest T. Note that Z could be a policy intervention, a new medical treatment or other forms of exposure. A central challenge in this context arises from the problem of endogeneity, which refers to Z not being independent of the error term in the structural model for T. Endogeneity can result from a variety of sources such as unobserved confounders, sample selection, measurement error and noncompliance (Rubin, 1974; Heckman, 1979; Angrist et al., 1996). When Z is endogenous, standard estimation methods can be severely biased. Because much applied work relies on observational data, addressing endogeneity should be a necessary step in any empirical analysis.

When the treatment Z is exogenous, meaning independent of the error term in the structural model for the outcome T, estimating its causal effect is a well-posed problem for which standard estimation approaches are well-suited. By contrast, when Z is endogenous, estimation of the treatment effect becomes more challenging. A common approach to deal with endogeneity is the use of instrumental variable (IV) methods. This approach relies on external sources of variation

(instruments) that influence Z, but are independent of the unobserved heterogeneity affecting both Z and T. However, with nonparametric IV methods, the estimation problem often takes the form of an ill-posed inverse problem (Newey and Powell, 2003; Chernozhukov and Hansen, 2005; Darolles et al., 2011; Blundell et al., 2013). This means that nonparametric estimation with instruments is more convoluted and unstable compared to standard estimation methods such as ordinary least squares or a general quantile estimator. As a result, researchers often focus on simpler estimands, such as averages, rather than attempting to recover full functional relationships. These issues highlight the importance of having a statistical test for the hypothesis that Z is exogenous.

We consider the following nonparametric nonseparable model

$$T = \varphi(X, Z, U), \tag{1}$$

where the map $u \mapsto \varphi(X, Z, u)$ is strictly increasing for each (X, Z) and $U \sim \mathcal{U}[0, 1]$. The variable T is a duration of interest, Z is a possibly endogenous treatment variable and X represents a vector of baseline covariates. Note that if $\varphi(X, Z, \cdot)$ is strictly increasing for each (X, Z), the rank invariance assumption as described by Chernozhukov and Hansen (2005) is implied (Dong and Shen, 2018). The rank invariance assumption is a condition on the potential outcomes and implies that individuals have the same unobserved rank when their treatment changes. Further, we will suppose that there exists an instrument W that is independent of U given X. We also allow for the duration time T to be censored by a right censoring time C. Therefore, we only observe the follow-up time $Y = \min\{T, C\}$ and the censoring indicator $\Delta = \mathbb{1}(Y = T)$ alongside X, W and Z. We will assume that X, W and Z are categorical variables with finite support to facilitate the construction of the test statistics. Moreover, it is assumed that C is independent of T and W jointly given X and Z. The goal of this paper is to develop nonparametric tests for the hypothesis that Z is exogenous, meaning that Z is independent of U conditional on X, without having to estimate the function φ.

There are two main approaches to test for exogeneity in a nonparametric nonseparable model. The first is based on estimating the function φ under the identification assumptions of the nonparametric IV model, and comparing the estimate to general regression or quantile estimates. The second approach is to estimate the residuals under the exogeneity assumption and verify if the IV condition is satisfied. We will follow the second approach to avoid the difficulties associated with nonparametric IV estimation. This puts us in a similar framework to Fève et al. (2018), who developed a nonparametric test for exogeneity based on an estimator of the distribution of the conditional rank $F_{T|Z}(T \mid Z)$. Contrary to Fève et al. (2018), we allow for T to be subject to right censoring. Moreover, we also allow for the inclusion of baseline covariates in model (1). These generalizations introduce significant challenges during the construction of the test statistics. Note that if T is right censored by C, this implies that $V_T = F_{T|X,Z}(T \mid X, Z)$ is right censored by $V_C = F_{T|X,Z}(C \mid X, Z)$. Therefore, conditional on $F_{T|X,Z}$, we only observe $V = \min\{V_T, V_C\}$ and $\Delta = \mathbb{1}(Y = T)$. Because of this, also the estimated $\widehat{V}_T = \widehat{F}_{T|X,Z}(T \mid X, Z)$ is subject to right censoring, which complicates the estimation of the distribution of V_T significantly. Whereas Fève et al. (2018) permits W and Z to be continuous, we restrict X, W and Z to be categorical. This restriction has the advantage that we do not need to smooth over the covariates, treatment and duration time when estimating $F_{T|X,Z}$. Under suitable identification conditions on φ, we show that the independence between V_T and (X, W) is equivalent to Z being independent of U given X. Other tests of exogeneity for censored duration outcomes have been proposed in the literature by Smith and Blundell (1986) and Rivers and Vuong (1988) among others. However, these tests assume fixed and fully observed censoring times in Tobit models. To the best of our knowledge, no other test for exogeneity in a nonparametric nonseparable model that allows for random right censoring is available in the literature.

The nonparametric nonseparable model is a common framework for analyzing the effects of endogenous treatments on censored duration outcomes. Several nonparametric contributions that focus on estimating local average treatment effects under a monotonicity assumption include Frandsen (2015), Sant'Anna (2016) and Blanco et al. (2020), while Chernozhukov et al. (2015) and Beyhum et al. (2022) address identification and estimation of population-level treatment effects under a type of rank invariance assumption. See Wüthrich (2020) for a comparison of rank invariance to monotonicity for the identification of quantile treatment effects. Another strand of the literature achieves identification under separability assumptions (Abbring and Van den Berg, 2003, 2005; Bijwaard and Ridder, 2005). Some semiparametric approaches include Tchetgen et al. (2015), Li et al. (2015), Chan (2016) Kianian et al. (2021), Beyhum et al. (2024) and Tedesco et al. (2025) among others.

The remainder of the paper is organized as follows. Section 2 discusses under which conditions exogeneity of Z is equivalent to the independence between V_T and (X, W), how the test statistics are constructed and explains the estimation procedure. The test statistics' limiting distributions and two bootstrap approaches to approximate the critical values are given in Section 3. The finite-sample performance is investigated in Section 4 through Monte Carlo simulations and Section 5 provides an empirical application to the National Job Training Partnership Act (JTPA) Study. Section 6 discusses possible extensions and practical considerations. The technical proofs can be found in the Appendix.

2 The test statistics

In this paper, we develop nonparametric tests for the null hypothesis that

$$H_0 : V_T \perp\!\!\!\perp (X, W),$$

where $V_T = F_{T|X,Z}(T \mid X, Z)$ with $F_{T|X,Z}(t \mid x, z) = \mathbb{P}(T \leq t \mid X = x, Z = z)$, X is a vector of baseline covariates and W is an instrumental variable such that $W \perp\!\!\!\perp U \mid X$ with $\perp\!\!\!\perp$ denoting statistical independence. Because T is right censored by C, V_T is right censored by $V_C = F_{T|X,Z}(C \mid X, Z)$. Therefore, conditionally on $F_{T|X,Z}$, we only observe $V = F_{T|X,Z}(Y \mid X, Z) = \min\{V_T, V_C\}$ and $\Delta = \mathbb{1}(Y = T)$. Note that T and V_T both being subject to right censoring introduces significant challenges in constructing the test statistics. We now introduce the following model condition:

(A0) Model (1) is identified. Moreover, $\varphi(X, Z, U)$ is strictly increasing in U for each (X, Z), $U \sim \mathcal{U}[0, 1]$ and W is independent of U conditionally on X.

The model being identified means that if

$$T = \varphi_1(X, Z, U_1) = \varphi_2(X, Z, U_2),$$

and U_1 and U_2 are both independent of W conditionally on X and uniform on $[0, 1]$, then

$$U_1 = U_2 \text{ a.s. and } \varphi_1 = \varphi_2.$$

Note that $U \sim \mathcal{U}[0, 1]$ can be assumed without loss of generality as long as U is continuously distributed and has positive density on its support. The equivalence between $V_T \perp\!\!\!\perp (X, W)$ and Z being exogenous, meaning that $Z \perp\!\!\!\perp U \mid X$, is established by the following proposition.

Proposition 1. *Assume (A0), then* $Z \perp\!\!\!\perp U \mid X \iff W \perp\!\!\!\perp V_T \mid X \iff V_T \perp\!\!\!\perp (X, W)$.

Proof. If $Z \perp\!\!\!\perp U \mid X$, clearly $U = V_T$ a.s. and hence $W \perp\!\!\!\perp V_T \mid X$. Since by construction $V_T \perp\!\!\!\perp X$, we have that $V_T \perp\!\!\!\perp (X, W)$. On the other hand, if $V_T \perp\!\!\!\perp (X, W)$, it is clear that $W \perp\!\!\!\perp V_T \mid X$ and by noticing that $T = F_{T|X,Z}^{-1}(V_T \mid X, Z)$ we can conclude that $U = V_T$ a.s. by the identification of model (1). Since V_T is independent of Z by construction, clearly $Z \perp\!\!\!\perp U \mid X$. $\qquad\square$

Proposition 1 shows that while our independence test between V_T and (X, W) does not rely on Assumption (A0), the independence property can be interpreted as an exogeneity property under this assumption. The identification of model (1), which is nonparametric and nonseparable, using instrumental variables is a complex issue that has been studied by Chesher (2003), Chernozhukov and Hansen (2005), Chernozhukov et al. (2007) and Chen et al. (2014) among others. A sufficient condition for the identification of model (1) can be found in Appendix A of Fève et al. (2018), which is a completeness type condition that has been proven to be nontestable (Canay et al., 2013).

Under H_0, it is clear that

$$\mathbb{P}(V_T \leq v, X = x, W = w) - \mathbb{P}(V_T \leq v)\mathbb{P}(X = x, W = w), \tag{2}$$

is equal to zero for all (v, x, w). Equivalently, we have that

$$\mathbb{P}(V_T \leq v \mid X = x, W = w) - \mathbb{P}(V_T \leq v), \tag{3}$$

is equal to zero for all v and all (x, w) for which $\mathbb{P}(X = x, W = w) > 0$. While it is possible to construct our test statistics based on an estimator of (2), V_T being subject to right censoring by V_C substantially complicates estimation of $\mathbb{P}(V_T \leq v, X = x, W = w)$. Nonetheless, possible estimators for this joint distribution have been proposed by Stute (1993) and Akritas (1994). The estimator proposed by Stute (1993) is simple to implement, but it would require the additional assumption that $\mathbb{P}(V_T \leq V_C \mid X, W, V_T) = \mathbb{P}(V_T \leq V_C \mid V_T)$. On the other hand, the approach proposed by Akritas (1994) requires the estimation of $\mathbb{P}(V_T \leq v \mid X = x, W = w)$ as a first step to estimate $\mathbb{P}(V_T \leq v, X = x, W = w)$. Therefore, we will construct our test statistics based on a nonparametric estimator of (3). To facilitate the construction of the test statistics, we will make the following key assumptions:

(A1) X, W and Z take values in the finite sets $\mathcal{R}_X = \{x_1, \ldots, x_{d_x}\}, \mathcal{R}_W = \{w_1, \ldots, w_{d_w}\}$ and $\mathcal{R}_Z = \{z_1, \ldots, z_{d_z}\}$ respectively.

(A2) $C \perp\!\!\!\perp (T, W) \mid X, Z$.

Even though assumption (A1) might be restrictive in some settings, it will allow us to develop a tractable, fully nonparametric approach to test H_0. Moreover, it has the advantage that we do not need to smooth over the covariates, treatment and duration time when estimating $F_{T|X,Z}$. Note that under Assumption (A1), a necessary condition for Assumption (A0) to hold is that $d_w \geq d_z$. Possible extensions to allow for continuous covariates, instruments and treatments are discussed in Section 6. Further, it follows that Assumption (A2) is equivalent to assuming that (i) $C \perp\!\!\!\perp W \mid X, Z$ and (ii) $T \perp\!\!\!\perp C \mid X, W, Z$. Clearly, (i) implies that after conditioning on (X, Z), the instrument should not affect the censoring time C. In particular, if W indicates being randomized to a treatment or control group, this means that, conditional on the baseline covariates X and treatment participation Z, being assigned to the treatment or control group should not influence the censoring time. Practitioners can check for violations of (i) by comparing the censoring rates for different levels of W conditional on (X, Z). On the other hand, (ii) implies that the duration and censoring times are independent given the baseline covariates, treatment assignment and treatment participation. Even though this assumption is widely adopted in the literature on

censored duration outcomes, it is, in general, nontestable (Tsiatis, 1975) and its violation can lead to biased estimates (Moeschberger and Klein, 1984). Note that when C is observed in addition to Y and Δ, Frandsen (2019) proposed a test for censoring point independence. In Section 6, we outline a possible extension to our approach that allows for some forms of dependence between T and C after conditioning on (X, W, Z).

2.1 Estimation

For all $i = 1, \ldots, n$, define

$$\widehat{V}_i = \widehat{F}_{T|X,Z}(Y_i \mid X_i, Z_i),$$

where $\widehat{F}_{T|X,Z}$ is a conditional Kaplan-Meier estimator, that is,

$$\widehat{F}_{T|X,Z}(t \mid x, z) = 1 - \prod_{i:Y_{(i)} \leq t} \left(1 - \frac{d_{(i)}(x,z)}{r_{(i)}(x,z)} \right),$$

with $Y_{(1)}, \ldots, Y_{(m)}$ the m ordered distinct follow-up times, the number of events $d_{(i)}(x,z) = \sum_{j=1}^n \mathbb{1}(Y_j = Y_{(i)}, \Delta_j = 1, X_j = x, Z_j = z)$ and the risk set $r_{(i)}(x,z) = \sum_{j=1}^n \mathbb{1}(Y_j \geq Y_{(i)}, X_j = x, Z_j = z)$. Because $\widehat{F}_{T|X,Z}(\cdot \mid x, z)$ is a monotonically increasing function, we have that

$$\widehat{V}_i = \widehat{F}_{T|X,Z}(\min\{T_i, C_i\} \mid X_i, Z_i) = \min\left\{ \widehat{F}_{T|X,Z}(T_i \mid X_i, Z_i), \widehat{F}_{T|X,Z}(C_i \mid X_i, Z_i) \right\},$$

such that $\widehat{V}_{T_i} = \widehat{F}_{T|X,Z}(T_i \mid X_i, Z_i)$ is censored by $\widehat{V}_{C_i} = \widehat{F}_{T|X,Z}(C_i \mid X_i, Z_i)$ with the same censoring indicator $\Delta_i = \mathbb{1}(Y_i = T_i)$. It is important to note that, since $V_T \perp\!\!\!\perp X, Z$ by construction, Assumption (A2) implies that $V_T \perp\!\!\!\perp V_C$. Therefore, we can estimate the distribution function of $V_T = F_{T|X,Z}(T \mid X, Z)$ by a Kaplan-Meier estimator, plugging-in $\{\widehat{V}_i\}_{i=1,\ldots,n}$ as the observed follow-up times. Specifically, let

$$\widehat{F}_{\widehat{V}_T}(v) = 1 - \prod_{i:\widehat{V}_{(i)} \leq v} \left(1 - \frac{\widehat{d}_{(i)}}{\widehat{r}_{(i)}} \right),$$

with $\widehat{V}_{(1)}, \ldots, \widehat{V}_{(k)}$ the k ordered distinct estimated conditional ranks, $\widehat{d}_{(i)} = \sum_{j=1}^n \mathbb{1}(\widehat{V}_j = \widehat{V}_{(i)}, \Delta_j = 1)$ and $\widehat{r}_{(i)} = \sum_{j=1}^n \mathbb{1}(\widehat{V}_j \geq \widehat{V}_{(i)})$. Lastly, it is important to note that we cannot estimate $F_{V_T|X,W}(v \mid x, w) = \mathbb{P}(V_T \leq v \mid X = x, W = w)$ by a conditional Kaplan-Meier estimator, since our assumptions do not imply $V_T \perp\!\!\!\perp V_C \mid X, W$. However, Assumption (A2) does imply that $V_T \perp\!\!\!\perp V_C \mid X, Z, W$. Therefore, we can estimate $F_{V_T|X,W,Z}(v \mid x, w, z) = \mathbb{P}(V_T \leq v \mid X = x, W = w, Z = z)$ by a conditional Kaplan-Meier estimator and marginalize out Z, that is,

$$\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) = \sum_{z \in \mathcal{R}_Z} \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z)\widehat{p}_{Z|X,W}(z \mid x, w),$$

where

$$\widehat{p}_{Z|X,W}(z \mid x, w) = \frac{\sum_{i=1}^n \mathbb{1}(X_i = x, W_i = w, Z_i = z)}{\sum_{i=1}^n \mathbb{1}(X_i = x, W_i = w)},$$

is an estimator of $p_{Z|X,W}(z \mid x, w) = \mathbb{P}(Z = z \mid X = x, W = w)$ and

$$\widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z) = 1 - \prod_{i:\widehat{V}_{(i)} \leq v} \left(1 - \frac{\widehat{d}_{(i)}(x,w,z)}{\widehat{r}_{(i)}(x,w,z)} \right),$$

with $\widehat{d}_{(i)}(x, w, z) = \sum_{j=1}^n \mathbb{1}(\widehat{V}_j = \widehat{V}_{(i)}, \Delta_j = 1, X_j = x, W_j = w, Z_j = z)$ and $\widehat{r}_{(i)}(x, w, z) = \sum_{j=1}^n \mathbb{1}(\widehat{V}_j \geq \widehat{V}_{(i)}, X_j = x, W_j = w, Z_j = z)$. Finally, let

$$\widehat{D}(v, x, w) = \widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - \widehat{F}_{\widehat{V}_T}(v),$$

be an estimator for (3). To check the null hypothesis, we can compute the following Kolmogorov-Smirnov statistic:

$$T_n^{KS} = \sqrt{n} \sup_{v \in \mathcal{I}, (x,w) \in \mathcal{R}_{X,W}} |\widehat{D}(v, x, w)|, \tag{4}$$

or the following weighted Cramér-von Mises statistic:

$$T_n^{CM} = n \sum_{(x,w) \in \mathcal{R}_{X,W}} \widehat{\pi}(x, w) \int_{\mathcal{I}} \widehat{D}^2(v, x, w) \, d\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w), \tag{5}$$

where $\widehat{\pi}(x, w)$ is a specified weight function, $\mathcal{R}_{X,W} = \{(x, w) \in \mathcal{R}_X \times \mathcal{R}_W : \mathbb{P}(X = x, W = w) > 0\}$ and $\mathcal{I} \subseteq [0, 1 - \gamma]$ for some small $\gamma > 0$ that will be defined by Assumption (A3). For example, one may take the weight function $\widehat{\pi}(x, w)$ of the Cramér-von Mises statistic to be constant or to equal the empirical proportion $n^{-1} \sum_{i=1}^n \mathbb{1}(X_i = x, W_i = w)$. Note that γ being strictly positive is only necessary for the asymptotic theory. In Section 4, we will set $\gamma = 0$ and show that the test still performs well using Monte Carlo simulations.

3   Distribution of test statistics

Before stating the main theorems regarding the test statistics' limiting distributions, we need to introduce some more notation and assumptions. Firstly, let

$$p_{X,W}(x, w) = \mathbb{P}(X = x, W = w), \quad p_{W|V_T,X,Z}(w \mid v, x, z) = \mathbb{P}(W = w \mid V_T = v, X = x, Z = z),$$
$$S_{V_T}(v) = 1 - F_{V_T}(v), \quad S_{V_T|X,W,Z}(v \mid x, w, z) = 1 - F_{V_T|X,W,Z}(v \mid x, w, z),$$
$$S_{V,1|X,W,Z}(v \mid x, w, z) = \mathbb{P}(F_{T|X,Z}(Y \mid X, Z) > v, \Delta = 1 \mid X = x, W = w, Z = z),$$
$$S_{V|X,W,Z}(v \mid x, w, z) = \mathbb{P}(F_{T|X,Z}(Y \mid X, Z) > v \mid X = x, W = w, Z = z),$$

and similarly for $S_{V,1|X,Z}$ and $S_{V|X,Z}$. Moreover, let

$$\zeta_i(v, x, w, z) = \int_0^v \frac{N_{i,x,w,z}(u) dS_{V,1|X,W,Z}(u \mid x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)^2} - \int_0^v \frac{dN_{i,x,w,z}^1(u)}{S_{V|X,W,Z}(u \mid x, w, z)},$$

and

$$\xi_i(u, x, z) = \int_0^u \frac{N_{i,x,z}(s) dS_{V,1|X,Z}(s \mid x, z)}{S_{V|X,Z}(s \mid x, z)^2} - \int_0^u \frac{dN_{i,x,z}^1(s)}{S_{V|X,Z}(s \mid x, z)},$$

with $N_{i,x,w,z}(v) = \mathbb{1}(V_i \geq v, X_i = x, W_i = w, Z_i = z)$, $N_{i,x,w,z}^1(v) = \mathbb{1}(V_i \geq v, \Delta_i = 1, X_i = x, W_i = w, Z_i = z)$ and similarly for $N_{i,x,z}(v)$ and $N_{i,x,z}^1(v)$. Lastly, we will need the following regularity assumptions:

(A3) For all $(x, w, z) \in \mathcal{R}_X \times \mathcal{R}_W \times \mathcal{R}_Z$, the conditional distribution function $F_{Y|X,W,Z}(y \mid x, w, z) = \mathbb{P}(Y \leq y \mid X = x, W = w, Z = z)$, with $f_{Y|X,W,Z}(y \mid x, w, z)$ the respective conditional density function, is twice continuously differentiable with respect to y, such that

$$\sup_y \left| \frac{\partial^k}{\partial y^k} F_{Y|X,W,Z}(y \mid x, w, z) \right| < \infty,$$

6

for $k = 1, 2$. Moreover, for all $(x, w, z) \in \mathcal{R}_X \times \mathcal{R}_W \times \mathcal{R}_Z$, we have that

$$\sup_{0 \leq F_{T|X,Z}(y|x,z) \leq 1-\gamma} \frac{f_{Y|X,W,Z}(y \mid x, w, z)}{f_{T|X,Z}(y \mid x, z)} < \infty,$$

where $\gamma > 0$ is such that $F_{Y|X,W,Z}(F_{T|X,Z}^{-1}(1 - \gamma \mid x, z) \mid x, w, z) < 1$ for all $(x, w, z) \in \mathcal{R}_X \times \mathcal{R}_W \times \mathcal{R}_Z$ and $f_{T|X,Z}(y \mid x, z)$ is the conditional density function of T given $X = x$ and $Z = z$.

(A4) We have that

$$\sup_{(x,w)\in\mathcal{R}_{X,W}} |\widehat{\pi}(x, w) - \pi(x, w)| \xrightarrow{a.s.} 0,$$

for some function $\pi(x, w)$.

We now have the following result.

3.1   Limiting distributions

Theorem 1. *Under Assumptions (A1) - (A3), we have that*

$$\widehat{D}(v, x, w) = n^{-1} \sum_{i=1}^{n} \frac{\omega_i(v, x, w)}{p_{X,W}(x, w)} + F_{V_T|X,W}(v \mid x, w) - v + o_p(n^{-1/2}),$$

uniformly in $(v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W}$, where

$$\omega_i(v, x, w) = S_{V_T|X,W,Z}(v \mid x, w, Z_i)\zeta_i(v, x, w, Z_i) - p_{W|V_T,X,Z}(w \mid v, x, Z_i)S_{V_T}(v)\xi_i(v, x, Z_i)$$
$$+ \mathbb{1}(X_i = x, W_i = w) \left[F_{V_T|X,W,Z}(v \mid x, w, Z_i) - F_{V_T|X,W}(v \mid x, w) \right].$$

Note that $F_{V_T|X,W}(v \mid x, w) = v$ for all $(v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W}$ under H_0. During the proof of this theorem, we also show in Appendix A that

$$\sup_{v\in\mathcal{I}} \left| \widehat{F}_{\widehat{V}_T}(v) - v \right| = o_p(n^{-1/2}),$$

meaning that $\widehat{F}_{\widehat{V}_T}(v)$ converges to v at a rate faster than the usual parametric $n^{-1/2}$-rate. While this was already shown by Fève et al. (2018) for uncensored conditional ranks, it is not obvious that the result would still hold in the presence of right censoring. Further, let $\ell^\infty(\mathcal{I} \times \mathcal{R}_{X,W})$ be the set of all uniformly bounded real-valued functions equipped with the supremum norm. We are now ready to give the main result regarding the limiting distributions of the proposed test statistics.

Theorem 2. *Under Assumptions (A1) - (A4) and H_0, the empirical process*

$$\left\{ n^{1/2} \left[\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - \widehat{F}_{\widehat{V}_T}(v) \right] : (v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W} \right\},$$

converges weakly in $\ell^\infty(\mathcal{I} \times \mathcal{R}_{X,W})$ to a Gaussian process $\mathbb{G}(v, x, w)$ with mean zero and covariance function

$$Cov \left(\frac{\omega_1(v_1, x_1, w_1)}{p_{X,W}(x_1, w_1)}, \frac{\omega_1(v_2, x_2, w_2)}{p_{X,W}(x_2, w_2)} \right),$$

where $\omega_1(\cdot)$ is defined in Theorem 1. Moreover, we have that

$$T_n^{KS} \xrightarrow{d} \sup_{v\in\mathcal{I},(x,w)\in\mathcal{R}_{X,W}} |\mathbb{G}(v, x, w)|,$$

and

$$T_n^{CM} \xrightarrow{d} \sum_{(x,w)\in\mathcal{R}_{X,W}} \pi(x,w) \int_{\mathcal{I}} \mathbb{G}^2(v,x,w)\, dF_{V_T}(v),$$

where $\pi(x,w)$ is defined by Assumption (A4).

An explicit expression for the covariance function is not provided, as it would be excessively long and almost infeasible to estimate in practice. Instead, the limiting distributions are approximated using two possible bootstrap approaches that are detailed in Section 3.2. Further, consider the local alternative

$$H_a : F_{V_T|X,W}(v \mid x, w) = v + n^{-1/2} H(v, x, w) \text{ for all } (v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W},$$

where for all $(x, w) \in \mathcal{R}_{X,W}$ the function H is such that $F_{V_T|X,W}$ remains a valid conditional distribution function under H_a. It follows that, under H_a, the test statistics converge in distribution to the same limiting distribution as under H_0, except for the additive bias $H(v, x, w)$, that is

$$T_n^{KS} \xrightarrow{d} \sup_{v\in\mathcal{I},(x,w)\in\mathcal{R}_{X,W}} |\mathbb{G}(v, x, w) + H(v, x, w)|,$$

and

$$T_n^{CM} \xrightarrow{d} \sum_{(x,w)\in\mathcal{R}_{X,W}} \pi(x,w) \int_{\mathcal{I}} [\mathbb{G}(v, x, w) + H(v, x, w)]^2 \, dF_{V_T}(v).$$

Therefore, the proposed tests can detect local alternatives that converge to H_0 at the parametric $n^{-1/2}$-rate.

3.2 Bootstrap approximations

Due to the complicated covariance structure of the process $\mathbb{G}(v, x, w)$, we use bootstrap approximations for the limiting distributions of T_n^{KS} and T_n^{CM}. For the bootstrap procedure to be valid, we follow the approach of Fève et al. (2018) and impose the slightly stronger null hypothesis

$$H_0^* : V_T \perp\!\!\!\perp (X, W, Z),$$

instead of $H_0 : V_T \perp\!\!\!\perp (X, W)$. Note that H_0^* is only slightly stronger than H_0 since $V_T \perp\!\!\!\perp (X, Z)$ by construction. We propose two bootstrap procedures that only differ in the construction of the bootstrap censoring times. Procedure A is the standard bootstrap for right censored data (Efron, 1981), in which both the duration and censoring times are generated from (conditional) Kaplan-Meier fits. Procedure B is based on the conditional random censoring algorithm (Robinson, 1983). In this approach, the known censoring times from the original sample are kept for the censored observations. For the uncensored observations, censoring times are generated from a conditional Kaplan-Meier fit, conditionally on the information that the unknown censoring time is greater than the observed duration time. In Section 4, we investigate whether there are scenarios under which one of the bootstrapping approaches performs better. The two procedures (Type A and B) are implemented as follows.

For each $b \in \{1, \ldots, B\}$:

    1. For $i = 1, \ldots, n$, let $X_i^* = X_i$, $W_i^* = W_i$ and $Z_i^* = Z_i$.

2. Let $U_{T,1}^*, \ldots, U_{T,n}^*$ be i.i.d. variables randomly drawn from $\mathcal{U}[0,1]$ and $\widehat{F}_{C|X,Z}^{-1}(u \mid X_i, Z_i) = \inf\{C_i : \widehat{F}_{C|X,Z}(C_i \mid X_i, Z_i) \geq u\}$, where $\widehat{F}_{C|X,Z}$ is calculated with a conditional Kaplan-Meier estimator that uses $(1 - \Delta)$ as the censoring indicator.

   A. Let $U_{C,1}^*, \ldots, U_{C,n}^*$ be i.i.d. variables randomly drawn from $\mathcal{U}[0,1]$. For $i = 1, \ldots, n$, let $C_i^* = \widehat{F}_{C|X,Z}^{-1}(U_{C,i}^* \mid X_i, Z_i)$.

   B. Let $U_{C,1}^*, \ldots, U_{C,n}^*$ be i.i.d. variables randomly drawn from $\mathcal{U}[\widehat{F}_{C|X,Z}(Y_i \mid X_i, Z_i), 1]$. For $i = 1, \ldots, n$, let $C_i^* = \Delta_i \times \widehat{F}_{C|X,Z}^{-1}(U_{C,i}^* \mid X_i, Z_i) + (1 - \Delta_i) \times C_i$.

3. Next, let $Y_i^* = \min\left\{\widehat{F}_{T|X,Z}^{-1}(U_{T,i}^* \mid X_i, Z_i), C_i^*\right\}$ and $\Delta_i^* = \mathbb{1}\left\{Y_i^* = \widehat{F}_{T|X,Z}^{-1}(U_{T,i}^* \mid X_i, Z_i)\right\}$.

4. Using $\{Y_i^*, \Delta_i^*, X_i, W_i, Z_i\}_{i=1}^n$ calculate the test statistic $T_{n,b}^*$.

Finally, compute the $p-$value as $B^{-1} \sum_{b=1}^B \mathbb{1}\left\{T_{n,b}^* > T_n\right\}$.

4 Finite sample study

In this section, we investigate the finite-sample performance of the proposed tests described in Section 2 and both of the bootstrap approximations described in Section 3.2 through Monte Carlo simulations. To estimate the Cramér-von Mises test statistic, we set $\widehat{\pi}(x, w) = 1$ for all simulation settings. Moreover, we implement the warp-speed method of Giacomini et al. (2013) to obtain the critical values. The procedure calculates a single bootstrap test statistic for each Monte Carlo sample and aggregates these statistics across simulations to approximate the critical value, which significantly reduces computation time. We will begin by describing the data-generating process, followed by a discussion of the proposed tests' performance under different degrees of endogeneity, censoring and instrument strength.

4.1 Data-generating process

For $i = 1, \ldots, n$, let
$$U_{T,i} \sim \mathcal{U}[0,1], \ U_{C,i} \sim \mathcal{U}[0,1] \text{ and } X_i \sim \text{Bin}(0.45),$$

such that
$$\pi_{W,i} = \frac{\exp(0.9 - 0.3X_i)}{1 + \exp(0.9 - 0.3X_i)},$$

and $W_i \sim \text{Bin}(\pi_{W,i})$. Moreover, let

$$\pi_{Z,i} = \frac{\exp(-2 + 0.2X_i + \eta W_i + \alpha \bar{U}_{T,i})}{1 + \exp(-2 + 0.2X_i + \eta W_i + \alpha \bar{U}_{T,i})},$$

with $\bar{U}_{T,i} = U_{T,i} - 0.5$ and $Z_i \sim \text{Bin}(\pi_{Z,i})$. The reason for centering $U_{T,i}$ is to keep the censoring rate somewhat stable when we vary α, which determines the endogeneity of Z. This will allow us to compare the power of the test statistics under different degrees of endogeneity. Moreover, Z is exogenous when $\alpha = 0$ and η controls the instrument strength. Further, let

$$T_i = \exp\left\{4 - 0.5X_i - Z_i + \Phi^{-1}(U_{T,i})\right\},$$

with Φ^{-1} the inverse of the standard normal distribution. Furthermore, let

$$C_i = -\frac{\log(1 - U_{C,i})}{\exp\{\lambda + 0.9X_i + 0.8Z_i\}}.$$

Note that λ controls the censoring rate. Lastly, we generate $Y_i = \min(T_i, C_i)$ and $\Delta_i = \mathbb{1}(Y_i = T_i)$ to get the simulated data $\{Y_i, \Delta_i, X_i, W_i, Z_i\}_{i=1,\dots,n}$. Even though we only have one covariate X, the parameter values were chosen such that there are very few data points for which $W_i = 0$ and $Z_i = 1$. On average, 1.9% of the simulated data have $(X_i, W_i, Z_i) = (0,0,1)$ and 2.3% have $(X_i, W_i, Z_i) = (1,0,1)$. This is a similar situation to the empirical application described in Section 5. For each of the following simulation settings, we performed 1000 Monte Carlo replications.

4.2 Endogeneity and sample size

To generate the data used in this part of the simulation study, we let $(\eta, \lambda) = (2.4, -5.7)$. By choosing these parameter values, we have an instrument strength of around $\tau_{WZ} = 0.45$ (measured by Kendall's Tau between W and Z) and approximately 25% censoring. We use Kendall's tau as a proxy for instrument strength because it is simple to interpret and report. Note that these values were chosen such that the instrument strength and censoring rate are comparable to the empirical strata examined in Section 5.

Looking at Figure 1, we see that both the Kolmogorov-Smirnov (KS) and Cramér-von Mises (CM) test statistics are able to discriminate the null from the alternative hypotheses. It is also clear that the CM test statistic performs better, where for a sample size of $n = 1000$, the distributions for $\alpha = 0$ and $\alpha = 5$ are almost completely separated. Figure 2 shows the power of the two test statistics as a function of the degree of endogeneity τ_{ZU} (measured by Kendall's Tau between Z and U_T) for different sample sizes and both of the proposed bootstrap approximations. Again, we see that the CM test statistic outperforms the KS test statistic. As the endogeneity and/or the sample size increase, so does the power of the test statistics. Under the null ($\tau_{ZU} = 0$), both test statistics reject at the nominal level of 5%, independent of the sample size. Comparing the two bootstrap approximations, there seems to be no noticeable difference. Looking at Figure 3, we see that the p-values obtained from Monte Carlo are well approximated by both bootstrapping procedures. However, approach B seems to be slightly more conservative. Overall, we find that our test statistics have good power, even when there are almost no observations with $W = 0$ and $Z = 1$. Moreover, the CM test statistic consistently outperforms the KS test statistic regardless of sample size or the degree of endogeneity.

4.3 Instrument strength and censoring rate

In this subsection, we will separately vary η and λ alongside α to assess the influence of the instrument strength and censoring rate, respectively, on the power of our proposed test statistics. All simulations use a sample size of $n = 1000$ and the other parameters have the same value as described in Section 4.1. We start by varying the instrument strength $\eta \in \{0.4, 0.9, 1.4, 1.9, 2.4, 2.9, 3.4\}$ and fixing $\lambda = -5.7$, such that there is around 25% censoring. Figure 4 shows us that the power increases when the instrument strength increases, as would be expected. Moreover, when there is weak endogeneity ($\alpha = 2$), the increase in power becomes less as the instrument strength increases. The CM test statistic again consistently outperforms the KS test statistic. Concerning the bootstrap approximations, it can be seen that approach B performs slightly worse than approach A for the KM test statistic when there is weak endogeneity. Under the null, the rejection rate remains at the nominal 5% level for both approximations, independent of the instrument strength.
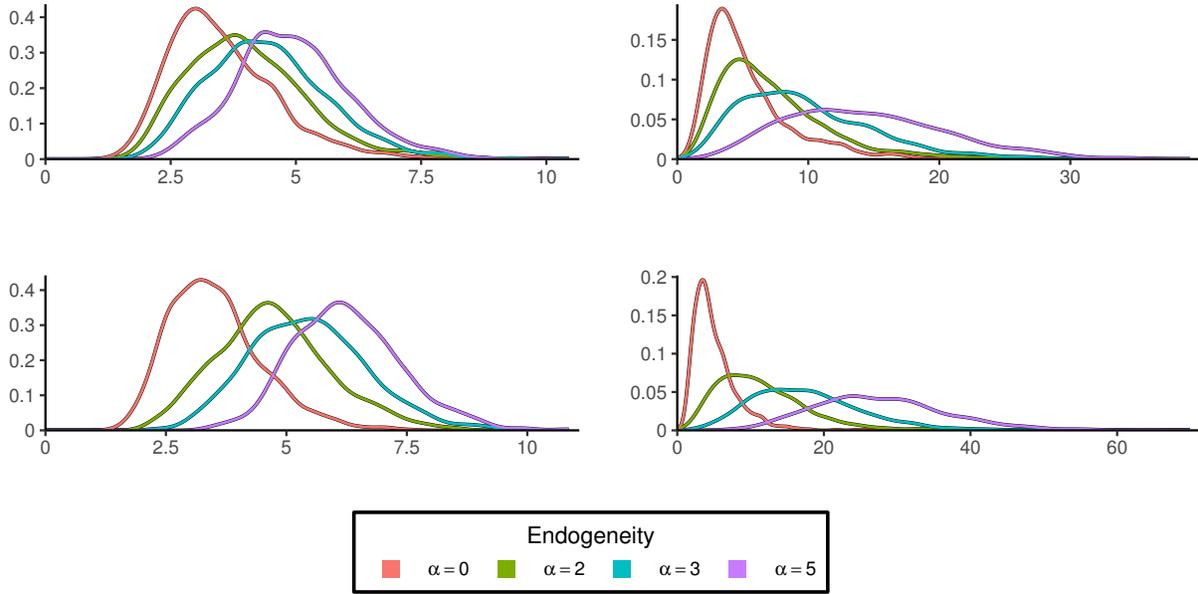
Figure 1: The distribution of T_n^{KS} (left) and T_n^{CM} (right) under the null and various alternative hypotheses (increasing degree of endogeneity α). The first row shows the results for $n = 500$ and the second row for $n = 1000$.
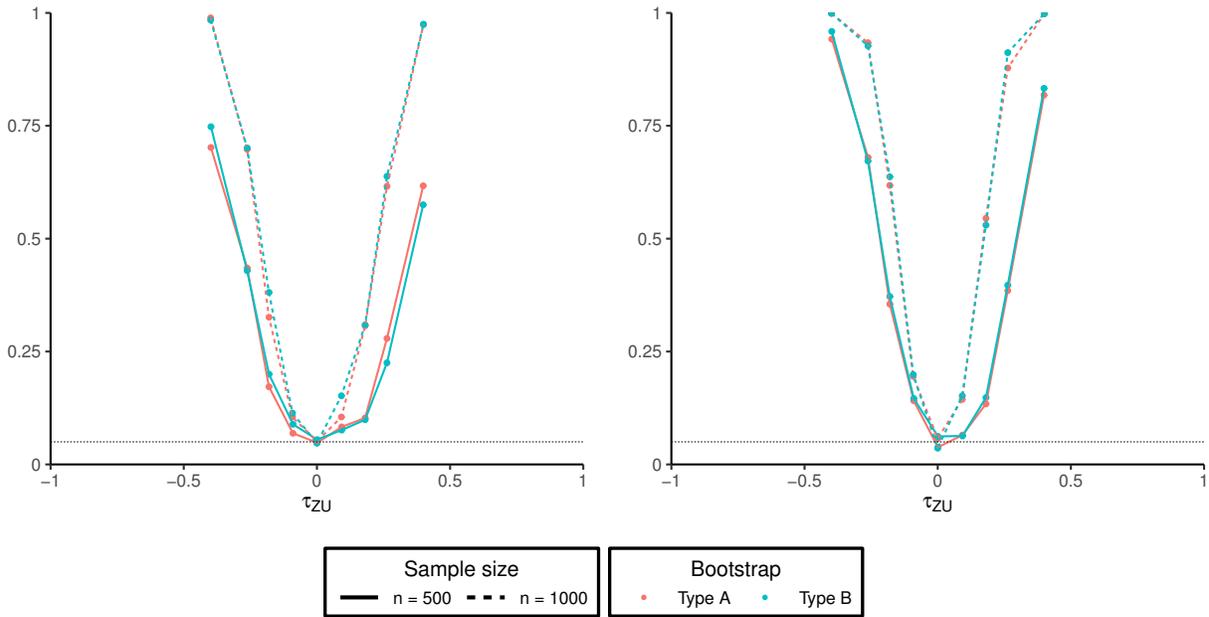

Figure 2: The power of T_n^{KS} (left) and T_n^{CM} (right) with respect to the degree of endogeneity τ_{ZU} for $n = 500$, $n = 1000$ and both bootstrap approximations (Type A and B).
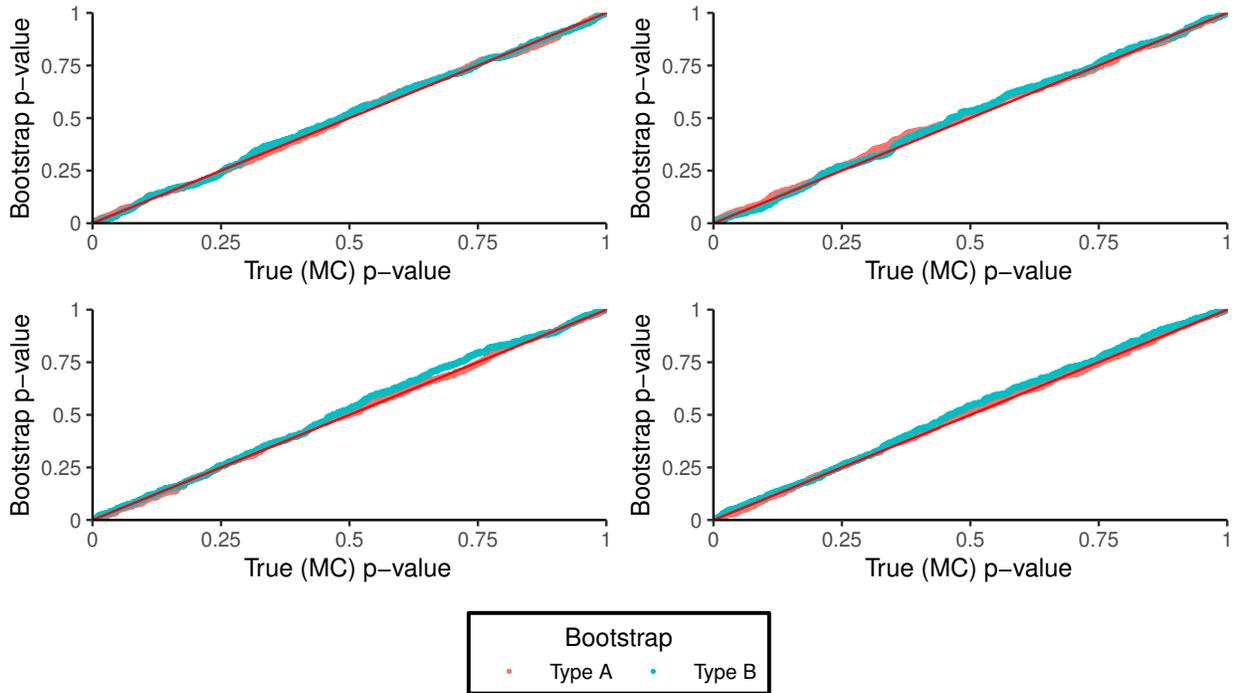
Figure 3: The Monte Carlo and bootstrap p-values for T_n^{KS} (left) and T_n^{CM} (right) under the null for both bootstrap approximations (Type A and B). The first row shows the results for $n = 500$ and the second row for $n = 1000$.
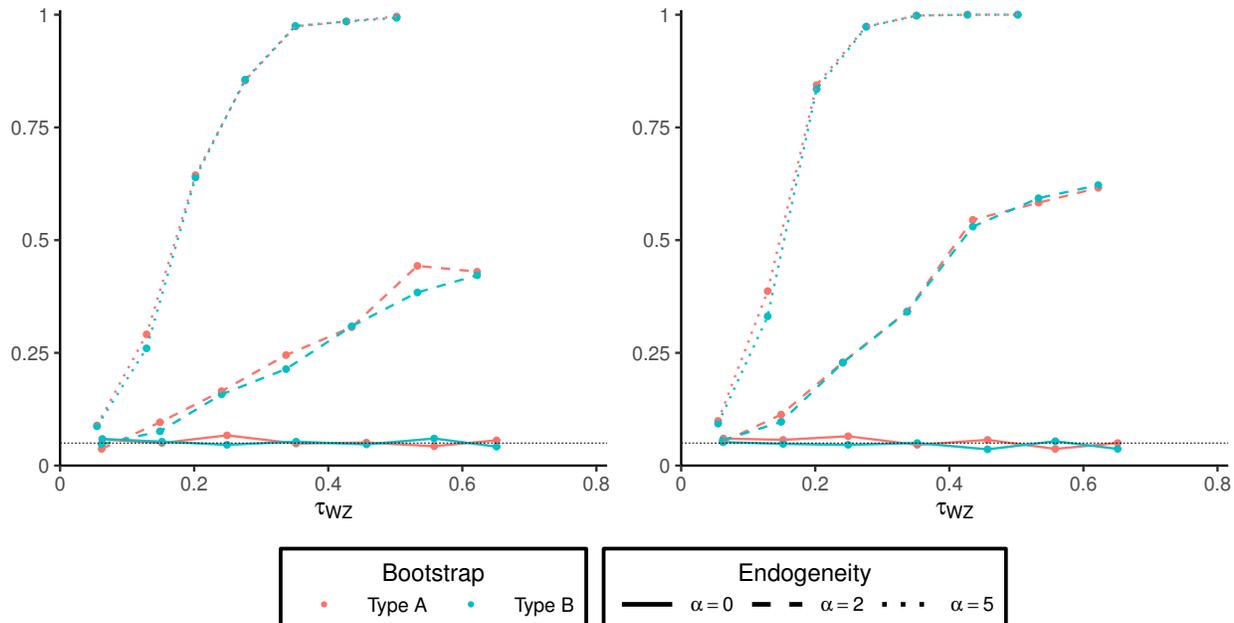

Figure 4: The power of T_{1000}^{KS} (left) and T_{1000}^{CM} (right) with respect to the instrument strength τ_{WZ} for both bootstrap approximations (Type A and B) and increasing degree of endogeneity (α).

To examine the impact of censoring, we vary the parameter $\lambda \in \{-5.7, -5.1, -4.6, -4.2, -3.8\}$ and fix $\eta = 2.4$ such that there is an instrument strength of around $\tau_{WZ} = 0.45$. As we would expect, Figure 5 indicates that the power of both test statistics decreases when there is more censoring. However, until approximately 50% censoring, the power of the CM test statistic remains relatively stable. When there is weak endogeneity, we see that bootstrap approach B performs slightly better when the censoring rate is above 50%. Interestingly, the CM test statistic performs a little worse than the KM test statistic when there is weak endogeneity and high censoring. The rejection rate under the null remains somewhat stable as the censoring rate increases, but falls slightly below 5% when the censoring rate is around 75%. Overall, the test statistics still perform well despite reduced instrument strength and high censoring rates.
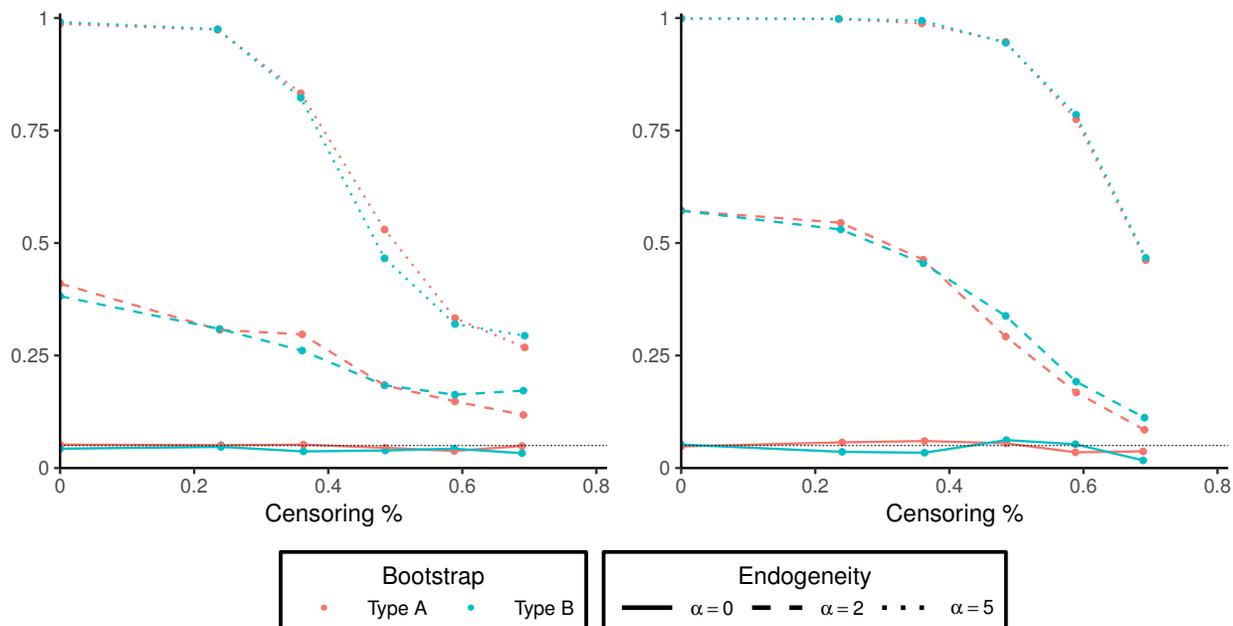

Figure 5: The power of T_{1000}^{KS} (left) and T_{1000}^{CM} (right) with respect to the censoring rate for both bootstrap approximations (Type A and B) and increasing degree of endogeneity (α).

5 Empirical application

The data come from the National Job Training Partnership Act (JTPA) Study, a large randomized evaluation of more than 600 federally funded programs designed to increase the employability of eligible adults and out-of-school youths. The study was conducted between 1987 and 1989, enrolled over 20000 applicants and collected follow-up information on employment, earnings, and program participation. Random assignment placed unemployed individuals into either a treatment group (eligible for JTPA services) or a control group (ineligible for 18 months). Nevertheless, roughly 3% of control-group members received JTPA services despite their ineligibility. Because individuals may self-select into treatment in a nonrandom way, noncompliance can induce endogeneity of the treatment (Angrist et al., 1996). All participants were surveyed between 12 and 36 months after randomization (average of 21 months). A second follow-up survey was administered to a subsample of 5,468 respondents, focusing on the interval between the two interviews, and took place

between 23 and 48 months after the initial randomization. The data can be downloaded at .

Because the study combines a clear policy intervention with extensive follow-up on earnings and labor-market outcomes, the data have become a standard for evaluating causal estimands in both cross-sectional and duration contexts. In particular, Bloom et al. (1997), Abadie et al. (2002) and Wüthrich (2020) investigate the impact of JTPA services on the sum of earnings after treatment. The effect of JTPA services on unemployment duration has been investigated by Frandsen (2015) and Beyhum et al. (2024) under the (conditional) independent censoring assumption, while Crommen et al. (2024) and Crommen et al. (2025a) allow for dependent censoring. To make the independent censoring assumption more plausible, Frandsen (2015) and Beyhum et al. (2024) rely exclusively on data from the first follow-up survey. Because almost all participants were surveyed, it is reasonable to assume that the censoring is administrative. By contrast, Crommen et al. (2024) and Crommen et al. (2025a) incorporate data from the second follow-up survey. They argue that participants who were invited to a second survey but did not participate could introduce endogenous censoring. Because the validity of our test statistics relies on the conditional independent censoring assumption, we restrict our empirical analysis to data from the first follow-up interview. Our goal is to test whether instrumental variable methods are required to identify the causal effect of JTPA services on unemployment duration, i.e., if JTPA services are an endogenous treatment.

Due to the size of the data set, we will focus our attention on two strata. The first stratum consists of 1127 single white men without children who are 30 years or younger and reported having no job at the time of treatment assignment. The other stratum is similar to the first one, except that the 1017 men are non-white. We will refer to these strata as white and non-white men, respectively. The covariate $HSGED$ equals 1 if an individual held a high school diploma or GED at the time of treatment assignment and 0 otherwise. Approximately 46% of white men and 40% of non-white men had a high school diploma or obtained a GED at the time of treatment assignment. Moreover, treatment assignment will be our instrument W ($W = 1$ if assigned to the treatment group, 0 for the control group). We deem treatment assignment to be a valid instrument as it is randomly assigned, correlated with JTPA participation (Kendall's Tau of 0.457 for white men and 0.475 for non-white men) and influences time to employment only through treatment participation. Note that $Z = 1$ if they participated in JTPA services, and 0 otherwise. Around 68% of white men and 71% of non-white men were assigned to the treatment group. In both strata, only 47% actually participated in JTPA services.

Figure 6 displays histograms of the observed follow-up times for each stratum, where a darker shading indicates a higher censoring rate. Consistent with Frandsen (2015), we find that most observations are censored at approximately 600 days, around which time most of the follow-up interviews took place. Beyond 600 days, nearly all observations are censored. Moreover, Table 1 shows the censoring rates for each level of $HSGED$, W and Z. keeping $HSGED$ and Z fixed, we find that the censoring rate remains stable for different levels of W. Therefore, we do not find any major violations of Assumption (A2). Table 1 also shows the number of observations for each level of $HSGED$, W and Z. As expected, there are very few observations from the control group who participated in JTPA services, i.e., $W = 0$ and $Z = 1$. However, as shown in Section 4, the test still performs well when there are only a few observations with $W = 0$ and $Z = 1$.

We compute the p-values of both test statistics using the two bootstrap approximations described in Section 3.2, based on 1000 resamples. Similarly to Section 4, we set $\widehat{\pi}(x, w) = 1$ for the Cramér–von Mises test statistic. The results are reported in Table 2 and Figure 7 displays the Kaplan–Meier curves, conditional on high school diploma/GED status and treatment participation, for both the white and non-white men. For the stratum of white men, we fail to reject the null hypothesis that Z is exogenous. Therefore, we can simply compare the estimated Kaplan-Meier
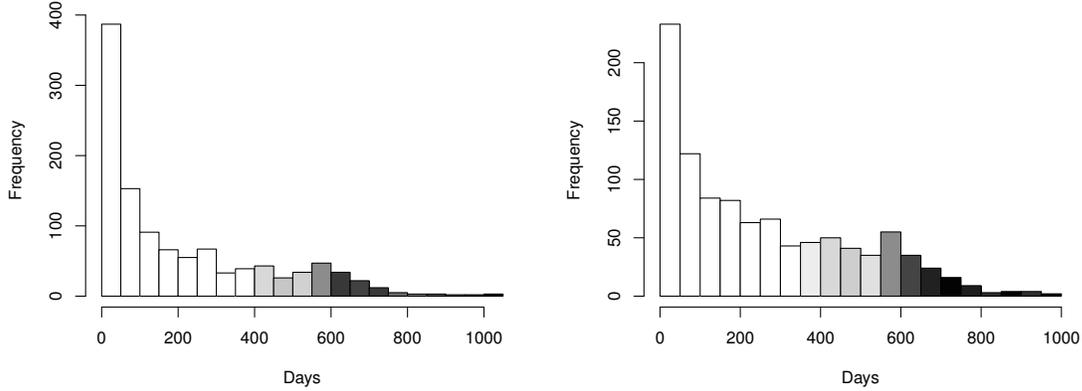
Figure 6: Histograms of the observed follow-up times Y for white men (left) and non-white men (right), in days, starting from random assignment. The darker the shade, the higher the censoring rate.

$W = 0$	$Z = 0$	$Z = 1$
HSGED = 0	$(171, 89.5\%)$	$(32, 90.6\%)$
HSGED = 1	$(141, 90.8\%)$	$(22, 90.9\%)$

$W = 1$	$Z = 0$	$Z = 1$
HSGED = 0	$(152, 87.5\%)$	$(255, 90.6\%)$
HSGED = 1	$(126, 87.3\%)$	$(228, 89.9\%)$

$W = 0$	$Z = 0$	$Z = 1$
HSGED = 0	$(154, 79.2\%)$	$(12, 91.7\%)$
HSGED = 1	$(112, 86.6\%)$	$(19, 84.2\%)$

$W = 1$	$Z = 0$	$Z = 1$
HSGED = 0	$(173, 79.2\%)$	$(272, 90.1\%)$
HSGED = 1	$(96, 87.5\%)$	$(179, 87.7\%)$

Table 1: Number of observations and censoring rate, respectively, for each level of high school diploma/GED status ($HSGED$), JTPA participation (Z) and treatment assignment (W) for white men (top row) and non-white men (bottom row).

curves for the treated and untreated. A log-rank test indicates no statistically significant difference between the survival curves (p-value of 0.330). Even after conditioning on having a high school diploma or GED, no significant difference is found (p-value of 0.503 for both levels of $HSGED$). We therefore conclude that, within this stratum, there is no evidence of a significant effect of JTPA services on unemployment duration.

By contrast, for the stratum of non-white men, the null hypothesis that Z is exogenous is rejected. This suggests the presence of unobserved heterogeneity influencing both treatment participation and unemployment duration. A log-rank test further reveals a significant difference between the survival functions for the treated and untreated (p-value $= 0.0003$). We find that, for non-white men without a high school diploma or GED, the treated have a significantly lower probability of being unemployed compared to the untreated (p-value of 8.87×10^{-5}). However, for non-white men with a high school diploma or GED, there seems to be no significant difference (p-value of 0.403). It is important to emphasize that because Z is endogenous in this stratum, we cannot interpret the observed survival differences as causal effects of JTPA participation. The effect of JTPA services among non-white men without a high school diploma or GED may, for

white men	Type A	Type B
T_{1127}^{KS}	0.595	0.572
T_{1127}^{CM}	0.399	0.392

non-white men	Type A	Type B
T_{1017}^{KS}	0.003	0.002
T_{1017}^{CM}	0.030	0.017

Table 2: Bootstrap p-values of the Kolmogorov–Smirnov (KS) and Cramér–von Mises (CM) test statistics for both bootstrap approximations (Type A and B), for white men (left) and non-white men (right).
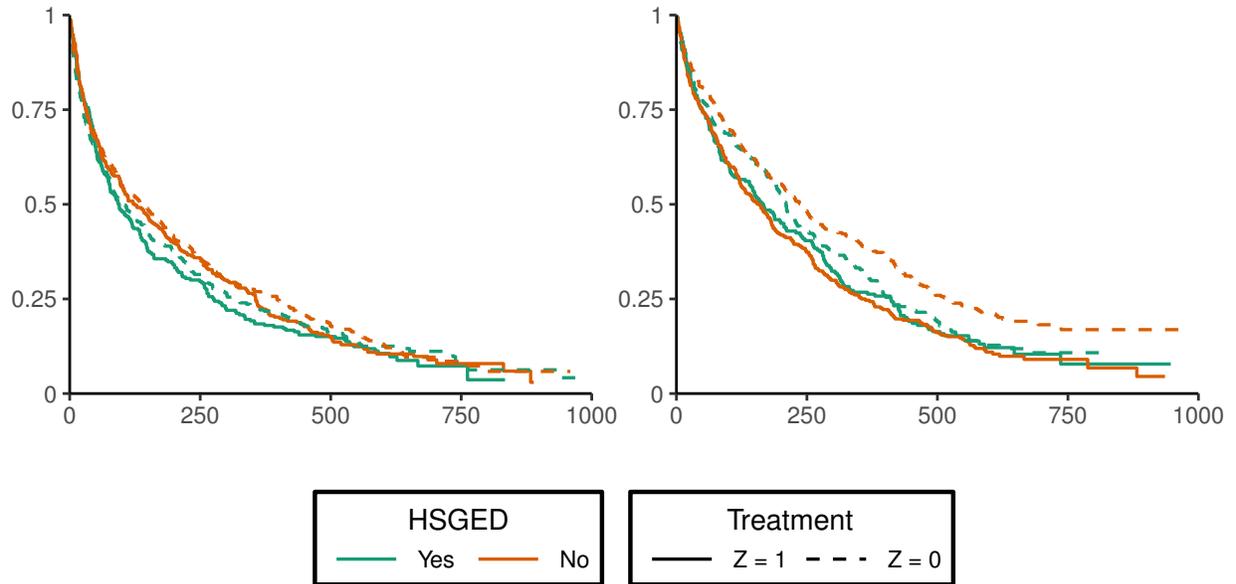

Figure 7: Kaplan-Meier curves, conditional on high school diploma/GED status (*HSGED*) and treatment participation (Z) for white men (left) and non-white men (right).

example, reflect self-selection of more motivated or higher-ability individuals into treatment. To actually estimate the causal effect of JTPA participation under endogeneity, one would need to use instrumental variable methods that allow for right censoring. Possible approaches include the nonparametric local average treatment effect framework of Frandsen (2015), as well as the population-level treatment effect methods developed by Beyhum et al. (2022, 2024), which are nonparametric and semiparametric, respectively.

6   Conclusion and future research

This paper develops nonparametric tests for exogeneity in nonparametric nonseparable duration models subject to right censoring. The tests are based on the independence between the conditional ranks, which are also right censored, and the instrumental variable. The proposed tests do not require estimation of the underlying structural model using nonparametric instrumental variable methods, which is even more challenging in the presence of right censoring. We derive the asymptotic properties of the proposed tests and illustrate their power under various scenarios using Monte Carlo simulations. The validity of the two bootstrap approaches to approximate the critical

values is also analyzed via simulations. We find that the Cramér–von Mises statistic consistently outperforms the Kolmogorov–Smirnov statistic, and that both bootstrapping approaches perform equally well in almost all scenarios. An empirical application to the National Job Training Partnership Act (JTPA) Study illustrates the usefulness of the tests in applied settings. In particular, it provides practitioners with a test to decide if standard survival comparisons are appropriate or whether instrumental variable methods that allow for right censoring are needed to correctly identify treatment effects.

The tests can be extended in multiple directions. Firstly, extending the methodology to accommodate continuous covariates, instruments and treatments would substantially broaden its applicability. The most straightforward way to achieve this is by replacing the conditional Kaplan–Meier estimator with a smooth conditional estimator, such as Beran (1981)'s estimator. Although conceptually straightforward, this extension introduces both theoretical and practical complications, such as bandwidth selection and the curse of dimensionality. A way of dealing with this could be to instead use semiparametric models, such as the Cox (1972) proportional hazards model. Another important extension is to relax the conditional independent censoring assumption and allow for some forms of dependence between T and C after conditioning on (X, W, Z). Copula-based methods, such as the copula-graphic estimator (Rivest and Wells, 2001; Zheng and Klein, 1995), offer a natural approach to modeling such dependencies. However, other copula-based methods for dependent censoring could also be considered (see Crommen et al. (2025b) for a recent review).

Acknowledgments and funding

References

Abadie, A., Angrist, J., and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica*, 70(1):91–117.

Abbring, J. H. and Van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517.

Abbring, J. H. and Van den Berg, G. J. (2005). Social experiments and instrumental variables with duration outcomes. Discussion Paper TI 05-047/3, Tinbergen Institute.

Akritas, M. G. (1994). Nearest neighbor estimation of a bivariate distribution under random censoring. *The Annals of Statistics*, 22(3):1299–1327.

Akritas, M. G. and Van Keilegom, I. (2003). Estimation of bivariate and marginal distributions with censored data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 65(2):457–471.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical report, Department of Statistics, University of California, Berkeley.

Beyhum, J., Florens, J.-P., and Van Keilegom, I. (2022). Nonparametric instrumental regression with right censored duration outcomes. *Journal of Business & Economic Statistics*, 40(3):1034–1045.

Beyhum, J., Tedesco, L., and Van Keilegom, I. (2024). Instrumental variable quantile regression under random right censoring. *The Econometrics Journal*, 27(1):21–36.

Bijwaard, G. E. and Ridder, G. (2005). Correcting for selective compliance in a re-employment bonus experiment. *Journal of Econometrics*, 125(1-2):77–111.

Blanco, G., Chen, X., Flores, C. A., and Flores-Lagunes, A. (2020). Bounds on average and quantile treatment effects on duration outcomes under censoring, selection, and noncompliance. *Journal of Business & Economic Statistics*, 38(4):901–920.

Bloom, H. S., Orr, L. L., Bell, S. H., Cave, G., Doolittle, F., Lin, W., and Bos, J. M. (1997). The benefits and costs of JTPA Title II-A programs: key findings from the National Job Training Partnership Act Study. *The Journal of Human Resources*, 32(3):549–576.

Blundell, R., Horowitz, J., and Parey, M. (2013). Nonparametric estimation of a heterogeneous demand function under the Slutsky inequality restriction. Technical Report CWP54/13, Institute for Fiscal Studies.

Canay, I. A., Santos, A., and Shaikh, A. M. (2013). On the testability of identification in some nonparametric models with endogeneity. *Econometrica*, 81(6):2535–2559.

Chan, K. C. G. (2016). Reader reaction: instrumental variable additive hazards models with exposure-dependent censoring. *Biometrics*, 72(3):1003–1005.

Chen, X., Chernozhukov, V., Lee, S., and Newey, W. K. (2014). Local identification of nonparametric and semiparametric models. *Econometrica*, 82(2):785–809.

Chernozhukov, V., Fernández-Val, I., and Kowalski, A. E. (2015). Quantile regression with censoring and endogeneity. *Journal of Econometrics*, 186(1):201–221.

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chernozhukov, V., Imbens, G. W., and Newey, W. K. (2007). Instrumental variable estimation of nonseparable models. *Journal of Econometrics*, 139(1):4–14.

Chesher, A. (2003). Identification in nonseparable models. *Econometrica*, 71(5):1405–1441.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 34(2):187–202.

Crommen, G., Beyhum, J., and Van Keilegom, I. (2024). An instrumental variable approach under dependent censoring. *TEST*, 33(2):473–495.

Crommen, G., Beyhum, J., and Van Keilegom, I. (2025a). Estimation of the complier causal hazard ratio under dependent censoring. *arXiv preprint arXiv:2504.02096.* (Under revision).

Crommen, G., Deresa, N. W., D'Haen, M., Ding, J., Willems, I., and Van Keilegom, I. (2025b). Recent advances in copula-based methods for dependent censoring. *SORT*, 49(1):3–42.

Darolles, S., Fan, Y., Florens, J.-P., and Renault, E. (2011). Nonparametric instrumental regression. *Econometrica*, 79(5):1541–1565.

Dong, Y. and Shen, S. (2018). Testing for rank invariance or similarity in program evaluation. *Review of Economics and Statistics*, 100(1):78–85.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319.

Fève, F., Florens, J.-P., and Van Keilegom, I. (2018). Estimation of conditional ranks and tests of exogeneity in nonparametric nonseparable models. *Journal of Business & Economic Statistics*, 36(2):334–345.

Frandsen, B. R. (2015). Treatment effects with censoring and endogeneity. *Journal of the American Statistical Association*, 110(512):1745–1752.

Frandsen, B. R. (2019). Testing censoring point independence. *Journal of Business & Economic Statistics*, 37(3):496–505.

Giacomini, R., Politis, D. N., and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, 29(3):567–589.

Gill, R. (1983). Large sample behaviour of the product-limit estimator on the whole line. *The Annals of Statistics*, 11(1):49–58.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Kianian, B., Kim, J. I., Fine, J. P., and Peng, L. (2021). Causal proportional hazards estimation with a binary instrumental variable. *Statistica Sinica*, 31(2):673.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference.* Springer.

Li, J., Fine, J., and Brookhart, A. (2015). Instrumental variable additive hazards models. *Biometrics*, 71(1):122–130.

Lo, S.-H. and Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representations. *Probability Theory and Related Fields*, 71(3):455–465.

Moeschberger, M. L. and Klein, J. P. (1984). Consequences of departures from independence in exponential series systems. *Technometrics*, 26(3):277–284.

Newey, W. K. and Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

Rivers, D. and Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics*, 39(3):347–366.

Rivest, L.-P. and Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *Journal of Multivariate Analysis*, 79(1):138–155.

Robinson, J. (1983). Bootstrap confidence intervals in location-scale models with progressive censoring. *Technometrics*, 25(2):179–187.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688.

Sant'Anna, P. H. C. (2016). Program evaluation with right-censored data. *arXiv preprint arXiv:1604.02642*.

Smith, R. J. and Blundell, R. W. (1986). An exogeneity test for a simultaneous equation Tobit model with an application to labor supply. *Econometrica*, 54(3):679–685.

Stute, W. (1993). Consistent estimation under random censorship when covariables are present. *Journal of Multivariate Analysis*, 45(1):89–103.

Tchetgen, E. J. T., Walter, S., Vansteelandt, S., Martinussen, T., and Glymour, M. (2015). Instrumental variable estimation in a survival context. *Epidemiology*, 26(3):402–410.

Tedesco, L., Beyhum, J., and Van Keilegom, I. (2025). Instrumental variable estimation of the proportional hazards model by presmoothing. *Electronic Journal of Statistics*, 19(1):656–717.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22.

Van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes: with Applications to Statistics*. Springer.

Wüthrich, K. (2020). A comparison of two quantile models with endogeneity. *Journal of Business & Economic Statistics*, 38(2):443–456.

Zheng, M. and Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82(1):127–138.

Appendix

Throughout the Appendix, let

$$\widehat{S}_{\widehat{V}}(v) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(\widehat{F}_{T|X,Z}(Y_i \mid X_i, Z_i) > v), \qquad\qquad S_{\widehat{V}}(v) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v),$$

$$\widehat{S}_{V}(v) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(F_{T|X,Z}(Y_i \mid X_i, Z_i) > v), \qquad\qquad S_{V}(v) = \mathbb{P}(F_{T|X,Z}(Y \mid X, Z) > v),$$

$$\widehat{S}_{\widehat{V},1}(v) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(\widehat{F}_{T|X,Z}(Y_i \mid X_i, Z_i) > v, \Delta_i = 1), \quad S_{\widehat{V},1}(v) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v, \Delta = 1),$$

$$\widehat{S}_{V,1}(v) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(F_{T|X,Z}(Y_i \mid X_i, Z_i) > v, \Delta_i = 1), \quad S_{V,1}(v) = \mathbb{P}(F_{T|X,Z}(Y \mid X, Z) > v, \Delta = 1),$$

where $S_{\widehat{V}}(v) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v)$ is calculated with respect to the law of (Y, Z, X) conditional on $\widehat{F}_{T|X,Z}$ with (Y, X, Z) independent of the data $\{Y_i, X_i, Z_i\}_{i=1}^{n}$ and $S_{\widehat{V},1}(v) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v, \Delta = 1)$ is calculated with respect to the law of (Y, Δ, X, Z) conditional on $\widehat{F}_{T|X,Z}$ with (Y, Δ, X, Z) independent of the data $\{Y_i, \Delta_i, X_i, Z_i\}_{i=1}^{n}$. Moreover, let

$$\widehat{S}_{\widehat{V},1|X,W,Z}(v \mid x, w, z) = \frac{\widehat{S}_{\widehat{V},X,W,Z,1}(v, x, w, z)}{\widehat{p}_{X,W,Z}(x, w, z)}, \quad S_{\widehat{V},1|X,W,Z}(v \mid x, w, z) = \frac{S_{\widehat{V},X,W,Z,1}(v, x, w, z)}{p_{X,W,Z}(x, w, z)},$$

$$\widehat{S}_{V,1|X,W,Z}(v \mid x, w, z) = \frac{\widehat{S}_{V,X,W,Z,1}(v, x, w, z)}{\widehat{p}_{X,W,Z}(x, w, z)}, \quad S_{V,1|X,W,Z}(v \mid x, w, z) = \frac{S_{V,X,W,Z,1}(v, x, w, z)}{p_{X,W,Z}(x, w, z)},$$

with

$$\widehat{S}_{\widehat{V},X,W,Z,1}(v, x, w, z) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(\widehat{F}_{T|X,Z}(Y_i \mid X_i, Z_i) > v, X_i = x, W_i = w, Z_i = z, \Delta_i = 1),$$

$$S_{\widehat{V},X,W,Z,1}(v, x, w, z) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v, X = x, W = w, Z = z, \Delta = 1),$$

$$\widehat{S}_{V,X,W,Z,1}(v, x, w, z) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(F_{T|X,Z}(Y_i \mid X_i, Z_i) > v, X_i = x, W_i = w, Z_i = z, \Delta_i = 1),$$

$$S_{V,X,W,Z,1}(v, x, w, z) = \mathbb{P}(F_{T|X,Z}(Y \mid X, Z) > v, X = x, W = w, Z = z, \Delta = 1),$$

$$\widehat{p}_{X,W,Z}(x, w, z) = n^{-1} \sum_{i=1}^{n} \mathbb{1}(X_i = x, W_i = w, Z_i = z),$$

$$p_{X,W,Z}(x, w, z) = \mathbb{P}(X = x, W = w, Z = z),$$

where $S_{\widehat{V},X,W,Z,1}(v, x, w, z) = \mathbb{P}(\widehat{F}_{T|X,Z}(Y \mid X, Z) > v, X = x, W = w, Z = z, \Delta = 1)$ is calculated with respect to the law of (Y, Δ, X, W, Z) conditional on $\widehat{F}_{T|X,Z}$ with (Y, Δ, X, W, Z) independent of the data $\{Y_i, \Delta_i, X_i, W_i, Z_i\}_{i=1}^{n}$. Further, let $\mathcal{R}_{X,W,Z} = \{(x, w, z) \in \mathcal{R}_X \times \mathcal{R}_W \times \mathcal{R}_Z : \mathbb{P}(X = x, W = w, Z = z) > 0\}$ and similarly for $\mathcal{R}_{X,Z}$. Note that F always denotes the distribution function, $S = 1 - F$ the survival function and f the density function. Moreover, we omit definitions when conditioning V or $\widehat{V}$ on, or forming the joint distribution with, any subset of (X, W, Z, Δ), as these are implied by earlier definitions. Before proving the main results of Section 3.1, we provide the following technical lemmas.

A Technical lemmas

Lemma 1. *Under Assumptions (A1) - (A3), we have that*

(i) $\sup_{v \in \mathcal{I}} \left| \widehat{S}_{\widehat{V}}(v) - \widehat{S}_V(v) - S_{\widehat{V}}(v) + S_V(v) \right| = o_p(n^{-1/2})$,

(ii) $\sup_{v \in \mathcal{I}} \left| \widehat{S}_{\widehat{V},1}(v) - \widehat{S}_{V,1}(v) - S_{\widehat{V},1}(v) + S_{V,1}(v) \right| = o_p(n^{-1/2})$,

(iii)
$$
\sup_{v \in \mathcal{I},(x,w,z) \in \mathcal{R}_{X,W,Z}} \left| \widehat{S}_{\widehat{V},X,W,Z}(v,x,w,z) - \widehat{S}_{V,X,W,Z}(v,x,w,z) \right.
$$
$$
\left. - S_{\widehat{V},X,W,Z}(v,x,w,z) + S_{V,X,W,Z}(v,x,w,z) \right| = o_p(n^{-1/2}),
$$

(iv)
$$
\sup_{v \in \mathcal{I},(x,w,z) \in \mathcal{R}_{X,W,Z}} \left| \widehat{S}_{\widehat{V},X,W,Z,1}(v,x,w,z) - \widehat{S}_{V,X,W,Z,1}(v,x,w,z) \right.
$$
$$
\left. - S_{\widehat{V},X,W,Z,1}(v,x,w,z) + S_{V,X,W,Z,1}(v,x,w,z) \right| = o_p(n^{-1/2}).
$$

Proof. We will start by giving the proof of (i). Define the class

$$
\mathcal{F} = \Big\{ (y,x,z) \to \mathbb{1}\left(F(y \mid x,z) \leq v\right) : v \in \mathcal{I},
$$
$$
F(\cdot \mid x,z) \text{ is monotone onto } [0,1] \text{ for all } (x,z) \in \mathcal{R}_{X,Z} \Big\}.
$$

We will now show that $\mathcal{F}$ is Donsker by proving that

$$
\int_0^\infty \sqrt{\log N_{[\,]}(\epsilon, \mathcal{F}, \|\cdot\|_{L_2})} \, d\epsilon < \infty,
$$

where $N_{[\,]}(\epsilon, \mathcal{F}, \|\cdot\|_{L_2})$ is the bracketing number and $\|g\|_{L_j} = \left(\mathbb{E}[g^j(Y,X,Z)]\right)^{1/j}$ for any j and function g (Van der Vaart and Wellner, 1996, Theorem 2.5.6). Since X and Z only take on a finite number of values by Assumption (A1), we can rewrite the function class

$$
\mathcal{F} = \Big\{ (y,x,z) \to \sum_{(\tilde{x},\tilde{z}) \in \mathcal{R}_{X,Z}} \mathbb{1}\left(x = \tilde{x}, z = \tilde{z}\right) \mathbb{1}\left(F(y \mid \tilde{x}, \tilde{z}) \leq v\right) : v \in \mathcal{I},
$$
$$
F(\cdot \mid \tilde{x}, \tilde{z}) \text{ is monotone onto } [0,1] \text{ for all } (\tilde{x},\tilde{z}) \in \mathcal{R}_{X,Z} \Big\},
$$

which is the sum of the classes

$$
\mathcal{F}_{\tilde{x},\tilde{z}} = \Big\{ (y,x,z) \to \mathbb{1}\left(x = \tilde{x}, z = \tilde{z}\right) \mathbb{1}\left(F(y \mid \tilde{x}, \tilde{z}) \leq v\right) : v \in \mathcal{I}, F(\cdot \mid \tilde{x}, \tilde{z}) \text{ is monotone onto } [0,1] \Big\},
$$

over all $(\tilde{x},\tilde{z}) \in \mathcal{R}_{X,Z}$. Note that we can rewrite $\mathcal{F}_{\tilde{x},\tilde{z}}$ as the product of the classes

$$
\mathcal{G}_{\tilde{x},\tilde{z}} = \Big\{ (x,z) \to \mathbb{1}\left(x = \tilde{x}, z = \tilde{z}\right) \Big\},
$$

and

$$
\mathcal{H}_{\tilde{x},\tilde{z}} = \Big\{ y \to \mathbb{1}\left(F(y \mid \tilde{x}, \tilde{z}) \leq v\right) : v \in \mathcal{I}, F(\cdot \mid \tilde{x}, \tilde{z}) \text{ is monotone onto } [0,1] \Big\}.
$$

It is clear that the bracketing number of $\mathcal{G}_{\tilde{x},\tilde{z}}$ is 1 and that $\mathcal{H}_{\tilde{x},\tilde{z}}$ has VC-dimension 2 for all $(\tilde{x},\tilde{z}) \in \mathcal{R}_{X,Z}$ (Van der Vaart and Wellner, 1996, Example 2.6.1), meaning they are both Donsker. Since both of these classes are uniformly bounded by 1, we know from Lemma 9.25 by Kosorok (2008) that $N_{[]}(\epsilon, \mathcal{F}_{\tilde{x},\tilde{z}}, \|\cdot\|_{L_2}) = O(\epsilon^{-2})$. Moreover, this also implies that

$$N_{[]}(\epsilon, \mathcal{F}, \|\cdot\|_{L_2}) = O(\epsilon^{-2d_{x,z}}),$$

where $d_{x,z} = |\mathcal{R}_{X,Z}|$. From this it follows that $\mathcal{F}$ is Donsker since

$$\int_0^1 \sqrt{\log(1/\epsilon)}\, d\epsilon < \infty,$$

and for $\epsilon > 1$, it is clear that 1 bracket suffices. We continue by using Corollary 2.3.12 in Van der Vaart and Wellner (1996), from which we know that for every $\epsilon > 0$:

$$\lim_{\kappa \downarrow 0} \lim_{n \to \infty} \mathbb{P}\left(\sup_{f,g \in \mathcal{F}, \rho(f-g) < \kappa} \left| n^{-1/2} \sum_{i=1}^n \{f(Y_i, X_i, Z_i) - g(Y_i, X_i, Z_i) \right.\right.$$
$$\left.\left. - \mathbb{E}[f(Y,X,Z)] + \mathbb{E}[g(Y,X,Z)]\} \right| > \epsilon \right) = 0,$$

where $\mathcal{F}$ is a Donsker class and

$$\rho^2(f-g) = \mathbb{E}\left[\{f(Y,X,Z) - g(Y,X,Z) - \mathbb{E}\left[f(Y,X,Z) - g(Y,X,Z)\right]\}^2 \right].$$

Therefore, we calculate

$$\rho^2 \left(\mathbb{1}(\widehat{F}_{T|X,Z}(Y \mid X, Z) \leq v) - \mathbb{1}(F_{T|X,Z}(Y \mid X, Z) \leq v) \right)$$

$$\leq \mathbb{E}\left[\left(\mathbb{1}(\widehat{F}_{T|X,Z}(Y \mid X, Z) \leq v) - \mathbb{1}(F_{T|X,Z}(Y \mid X, Z) \leq v) \right)^2 \right]$$

$$= \int \left[F_{Y|X,Z} \left(\widehat{F}_{T|X,Z}^{-1}(v \mid x, z) \mid x, z \right) dF_{X,Z}(x,z) \right.$$

$$+ \int F_{Y|X,Z} \left(F_{T|X,Z}^{-1}(v \mid x, z) \mid x, z \right) dF_{X,Z}(x,z)$$

$$- 2 \int F_{Y|X,Z} \left(\widehat{F}_{T|X,Z}^{-1}(v \mid x, z) \wedge F_{T|X,Z}^{-1}(v \mid x, z) \mid x, z \right) dF_{X,Z}(x,z)$$

$$= \int \left| F_{Y|X,Z} \left(\widehat{F}_{T|X,Z}^{-1}(v \mid x, z) \mid x, z \right) - F_{Y|X,Z} \left(F_{T|X,Z}^{-1}(v \mid x, z) \mid x, z \right) \right| dF_{X,Z}(x,z)$$

$$\leq M \int \left| \widehat{F}_{T|X,Z}^{-1}(v \mid x, z) - F_{T|X,Z}^{-1}(v \mid x, z) \right| dF_{X,Z}(x,z)$$

$$\leq M \max_{(x,z) \in \mathcal{R}_{X,Z}} \sup_{v \in \mathcal{I}} \left| \widehat{F}_{T|X,Z}^{-1}(v \mid x, z) - F_{T|X,Z}^{-1}(v \mid x, z) \right|,$$

with M some constant and where $\wedge$ indicates the minimum. It now follows from Theorem 2 by Lo and Singh (1986) that

$$\sup_{v \in \mathcal{I}} \left| \widehat{F}_{T|X,Z}^{-1}(v \mid x, z) - F_{T|X,Z}^{-1}(v \mid x, z) \right| = o(1) \quad a.s.,$$

for all $(x,z) \in \mathcal{R}_{X,Z}$, which completes the proof. The proofs of (ii), (iii) and (iv) are similar to the proof of (i) and therefore omitted. $\qquad\square$

Lemma 2. *Under Assumptions (A1) - (A3), we have that*

(i) $\sup_{v \in \mathcal{I}} \left| \widehat{S}_{\widehat{V}}(v) - S_V(v) \right| = O_p(n^{-1/2})$,

(ii) $\sup_{v \in \mathcal{I}, (x,w,z) \in \mathcal{R}_{X,W,Z}} \left| \widehat{S}_{\widehat{V}|X,W,Z}(v \mid x, w, z) - S_{V|X,W,Z}(v \mid x, w, z) \right| = O_p(n^{-1/2})$.

Proof. We will start by giving the proof of (i). Using Lemma 1, it is clear that

$$\widehat{S}_{\widehat{V}}(v) - S_V(v) = \widehat{S}_V(v) - S_V(v) + S_{\widehat{V}}(v) - S_V(v) + o_p(n^{-1/2}),$$

uniformly in $v \in \mathcal{I}$. Note that by using a Taylor expansion, with $S_{Y|X,Z}(y \mid x, z) = \mathbb{P}(Y > y \mid X = x, Z = z)$ and $f_{Y|X,Z}$ the corresponding density, we have that

$$S_{\widehat{V}}(v) = \sum_{(x,z) \in \mathcal{R}_{X,Z}} S_{Y|X,Z}(\widehat{F}^{-1}_{T|X,Z}(v \mid x, z) \mid x, z) p_{X,Z}(x, z)$$

$$= \sum_{(x,z) \in \mathcal{R}_{X,Z}} \left[S_{Y|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z) - f_{Y|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z) \right.$$

$$\left. \times [\widehat{F}^{-1}_{T|X,Z}(v \mid x, z) - F^{-1}_{T|X,Z}(v \mid x, z)] \right] p_{X,Z}(x, z) + o_p(n^{-1/2}),$$

such that

$$S_{\widehat{V}}(v) - S_V(v)$$

$$= \sum_{(x,z) \in \mathcal{R}_{X,Z}} f_{Y|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z) \left[F^{-1}_{T|X,Z}(v \mid x, z) - \widehat{F}^{-1}_{T|X,Z}(v \mid x, z) \right] p_{X,Z}(x, z)$$

$$+ o_p(n^{-1/2})$$

$$= \sum_{(x,z) \in \mathcal{R}_{X,Z}} \frac{f_{Y|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z)}{f_{T|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z)} \left[\widehat{F}_{V_T|X,Z}(v \mid x, z) - F_{V_T|X,Z}(v \mid x, z) \right] p_{X,Z}(x, z)$$

$$+ o_p(n^{-1/2}),$$

uniformly in $v \in \mathcal{I}$ and where the second equality follows from Theorem 2 by Lo and Singh (1986). Using Theorem 1.2 by Gill (1983), it now follows that $\sup_{v \in \mathcal{I}} \left| S_{\widehat{V}}(v) - S_V(v) \right| = O_p(n^{-1/2})$. Moreover, using standard empirical process results, it is clear that $\sup_{v \in \mathcal{I}} \left| \widehat{S}_V(v) - S_V(v) \right| = O_p(n^{-1/2})$, from which the result follows. We continue with the proof of (ii). Applying Lemma 1, we have that

$$\widehat{S}_{\widehat{V}|X,W,Z}(v \mid x, w, z) - S_{V|X,W,Z}(v \mid x, w, z)$$

$$= \widehat{S}_{V|X,W,Z}(v \mid x, w, z) - S_{V|X,W,Z}(v \mid x, w, z)$$

$$+ \left[S_{\widehat{V}|X,W,Z}(v \mid x, w, z) - S_{V|X,W,Z}(v \mid x, w, z) \right] \times (1 + o_p(1)) + o_p(n^{-1/2}),$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. Using a similar Taylor expansion as in the proof of (i), we have that

$$S_{\widehat{V}|X,W,Z}(v \mid x, w, z) - S_{V|X,W,Z}(v \mid x, w, z)$$

$$= \frac{f_{Y|X,W,Z}(F^{-1}_{T|X,Z}(v \mid x, w, z) \mid x, w, z)}{f_{T|X,Z}(F^{-1}_{T|X,Z}(v \mid x, z) \mid x, z)} \left[\widehat{F}_{V_T|X,Z}(v \mid x, z) - F_{V_T|X,Z}(v \mid x, z) \right] + o_p(n^{-1/2}),$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. The result now follows from using the same empirical process arguments as those used in the proof of (i). $\qquad\square$

Lemma 3. *Under Assumptions (A1) - (A3), we have that*

(i) $\sup_{v \in \mathcal{I}} \left| \int_0^v \left(\widehat{S}_{\widehat{V}}^{-1}(u) - S_V^{-1}(u) \right) d \left(\widehat{S}_{\widehat{V},1}(u) - S_{V,1}(u) \right) \right| = o_p(n^{-1/2}),$

(ii) $\displaystyle\sup_{v \in \mathcal{I}, (x,w,z) \in \mathcal{R}_{X,W,Z}} \left| \int_0^v \left(\widehat{S}_{\widehat{V}|X,W,Z}^{-1}(u \mid x, w, z) - S_{V|X,W,Z}^{-1}(u \mid x, w, z) \right) \right.$

$$\left. d \left(\widehat{S}_{\widehat{V},1|X,W,Z}(u \mid x, w, z) - S_{V,1|X,W,Z}(u \mid x, w, z) \right) \right| = o_p(n^{-1/2}).$$

Proof. We will start by giving the proof of (i). Following the proof of Lemma 2 by Lo and Singh (1986), we start by dividing $\mathcal{I}$ into subintervals $[l_i, l_{i+1}]$, $i = 1, 2, \ldots, k_n$ where $k_n = O(n^{1/2} \log(n)^{-1/2})$ and $0 = l_1 < l_2 < \cdots < l_{k_n+1} = 1 - \gamma$ are such that $S_V(l_i) - S_V(l_{i+1}) \le M n^{-1/2} \log(n)^{1/2}$, where M is some constant whose value may change from line to line. We then have that

$$\left| \int_0^v \left(\widehat{S}_{\widehat{V}}^{-1}(u) - S_V^{-1}(u) \right) d \left(\widehat{S}_{\widehat{V},1}(u) - S_{V,1}(u) \right) \right|$$

$$\le k_n \sup_{u \in \mathcal{I}} \left| \widehat{S}_{\widehat{V}}^{-1}(u) - S_V^{-1}(u) \right| \max_{1 \le i \le k_n} \left| \widehat{S}_{\widehat{V},1}(l_{i+1}) - \widehat{S}_{\widehat{V},1}(l_i) - S_{V,1}(l_{i+1}) + S_{V,1}(l_i) \right| \qquad (6)$$

$$+ 2 \max_{1 \le i \le k_n} \sup_{u \in [l_i, l_{i+1}]} \left| \widehat{S}_{\widehat{V}}^{-1}(u) - \widehat{S}_{\widehat{V}}^{-1}(l_i) - S_V^{-1}(u) + S_V^{-1}(l_i) \right|. \qquad (7)$$

We first investigate (7), and start by further dividing each $[l_i, l_{i+1}]$ into the subintervals $[l_{ij}, l_{i(j+1)}]$, $j = 1, \ldots, a_n$ such that $\left| S_V(l_{i(j+1)}) - S_V(l_{ij}) \right| = O(n^{-3/4} \log(n)^{3/4})$ uniformly in i, j, and $a_n = O(n^{1/4} \log(n)^{-1/4})$. Thanks to our Lemma 2, it follows that

$$\sup_{u \in [l_i, l_{i+1}]} \left| \widehat{S}_{\widehat{V}}^{-1}(u) - \widehat{S}_{\widehat{V}}^{-1}(l_i) - S_V^{-1}(u) + S_V^{-1}(l_i) \right|$$

$$\le \sup_{u \in [l_i, l_{i+1}]} \left| \frac{\widehat{S}_{\widehat{V}}(u) - S_V(u)}{S_V(u)^2} - \frac{\widehat{S}_{\widehat{V}}(l_i) - S_V(l_i)}{S_V(l_i)^2} \right| + O_p(n^{-1} \log(n))$$

$$\le \sup_{u \in [l_i, l_{i+1}]} \frac{\left| \widehat{S}_{\widehat{V}}(u) - \widehat{S}_{\widehat{V}}(l_i) - S_V(u) + S_V(l_i) \right|}{S_V(l_{i+1})^2} + O_p(n^{-1} \log(n))$$

$$\le M \max_{1 \le j \le a_n} \left| \widehat{S}_{\widehat{V}}(l_{ij}) - \widehat{S}_{\widehat{V}}(l_i) - S_V(l_{ij}) + S_V(l_i) \right| + O_p(n^{-3/4} \log(n)^{3/4}).$$

From Lemma 1, it follows that

$$M \max_{1 \le j \le a_n} \left| \widehat{S}_{\widehat{V}}(l_{ij}) - \widehat{S}_{\widehat{V}}(l_i) - S_V(l_{ij}) + S_V(l_i) \right|$$

$$\le M \max_{1 \le j \le a_n} \left| \widehat{S}_V(l_{ij}) - \widehat{S}_V(l_i) - S_V(l_{ij}) + S_V(l_i) \right| \qquad (8)$$

$$+ M \max_{1 \le j \le a_n} \left| S_{\widehat{V}}(l_{ij}) - S_{\widehat{V}}(l_i) - S_V(l_{ij}) + S_V(l_i) \right| + o_p(n^{-1/2}). \qquad (9)$$

25

Since (8) is the same term as the one in Lemma 2 by Lo and Singh (1986), we know that it is $O_p(n^{-3/4}\log(n)^{3/4})$. Using the same Taylor expansion from the beginning of the proof of (i) from our Lemma 2, we have that

$$M \max_{1 \leq j \leq a_n} \left| S_{\widehat{V}}(l_{ij}) - S_{\widehat{V}}(l_i) - S_V(l_{ij}) + S_V(l_i) \right|$$

$$\leq M d_{x,z} \max_{(x,z) \in \mathcal{R}_{X,Z}} \max_{1 \leq j \leq a_n} \left| \Psi_{x,z}(l_{ij})\widehat{F}_{V_T|X,Z}(l_{ij} \mid x,z) - \Psi_{x,z}(l_i)\widehat{F}_{V_T|X,Z}(l_i \mid x,z) \right.$$

$$\left. - \Psi_{x,z}(l_{ij})F_{V_T|X,Z}(l_{ij} \mid x,z) + \Psi_{x,z}(l_i)F_{V_T|X,Z}(l_i \mid x,z) \right| + o_p(n^{-1/2}),$$

where $d_{x,z} = |\mathcal{R}_{X,Z}|$ and $\Psi_{x,z}(u) = f_{Y|X,Z}(F_{T|X,Z}^{-1}(u \mid x,z) \mid x,z)/f_{T|X,Z}(F_{T|X,Z}^{-1}(u \mid x,z) \mid x,z)$. Similar to the steps of the proof of Theorem 2 by Lo and Singh (1986), we find that

$$\max_{1 \leq j \leq a_n} \left| \Psi_{x,z}(l_{ij})\widehat{F}_{V_T|X,Z}(l_{ij} \mid x,z) - \Psi_{x,z}(l_i)\widehat{F}_{V_T|X,Z}(l_i \mid x,z) \right.$$

$$\left. - \Psi_{x,z}(l_{ij})F_{V_T|X,Z}(l_{ij} \mid x,z) + \Psi_{x,z}(l_i)F_{V_T|X,Z}(l_i \mid x,z) \right| = O_p(n^{-3/4}\log(n)^{3/4}).$$

Since the proof of (6) is similar, it is omitted. The proof of (ii) is almost identical to the proof of (i), and therefore also omitted. $\qquad\square$

Lemma 4. *Under Assumptions (A1) - (A3), we have that*

(i) $\sup_{v \in \mathcal{I}} \left| \widehat{F}_{\widehat{V}_T}(v) - v \right| = o_p(n^{-1/2})$,

(ii) $\sup_{v \in \mathcal{I},(x,w,z) \in \mathcal{R}_{X,W,Z}} \left| \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x,w,z) - F_{V_T|X,W,Z}(v \mid x,w,z) \right| = O_p(n^{-1/2})$.

Proof. We will start by giving the proof of (i). Using Lemma 2, we can follow the same steps from the proof of Theorem 1 by Lo and Singh (1986) to show that

$$\log \left(1 - \widehat{F}_{\widehat{V}_T}(v) \right) - \log(1 - v)$$

$$= -\int_0^v \frac{\widehat{S}_{\widehat{V}}(u)}{S_V(u)^2} dS_{V,1}(u) + \int_0^v \frac{d\widehat{S}_{\widehat{V},1}(u)}{S_V(u)} \tag{10}$$

$$+ \int_0^v \left(\widehat{S}_{\widehat{V}}^{-1}(u) - S_V^{-1}(u) \right) d \left(\widehat{S}_{\widehat{V},1}(u) - S_{V,1}(u) \right) + o_p(n^{-1/2}), \tag{11}$$

uniformly in $v \in \mathcal{I}$. In view of Lemma 3, it is clear that (11) is $o_p(n^{-1/2})$ uniformly in $v \in \mathcal{I}$. Further, using Lemma 1, it follows that (10) can be written as

$$-\int_0^v \frac{\widehat{S}_V(u)}{S_V(u)^2} dS_{V,1}(u) + \int_0^v \frac{d\widehat{S}_{V,1}(u)}{S_V(u)} - \int_0^v \frac{S_{\widehat{V}}(u)}{S_V(u)^2} dS_{V,1}(u) + \int_0^v \frac{dS_{\widehat{V},1}(u)}{S_V(u)} + o_p(n^{-1/2}), \tag{12}$$

uniformly in $v \in \mathcal{I}$. Note that the first two terms of (12) are the same terms from Theorem 1 by Lo and Singh (1986). Using the same Taylor expansion as in the proof of Lemma 2, we get that

$$S_{\widehat{V}}(v) = \sum_{(x,z) \in \mathcal{R}_{X,Z}} S_{Y|X,Z}(\widehat{F}_{T|X,Z}^{-1}(v \mid x,z) \mid x,z)p_{X,Z}(x,z)$$

$$= \sum_{(x,z) \in \mathcal{R}_{X,Z}} \left[S_{Y|X,Z}(F_{T|X,Z}^{-1}(v \mid x,z) \mid x,z) - f_{Y|X,Z}(F_{T|X,Z}^{-1}(v \mid x,z) \mid x,z) \right.$$

$$\left. \times [\widehat{F}_{T|X,Z}^{-1}(v \mid x,z) - F_{T|X,Z}^{-1}(v \mid x,z)] \right] p_{X,Z}(x,z) + o_p(n^{-1/2}),$$

uniformly in $v \in \mathcal{I}$, and similarly for $S_{\widehat{V},1}(v)$. Therefore, we get that the last two terms of (12) are equal to

$$
\sum_{(x,z)\in\mathcal{R}_{X,Z}} p_{X,Z}(x,z) \left\{ \int_0^v \frac{f_{V|X,Z}(u \mid x,z)[\widehat{F}_{T|X,Z}^{-1}(u \mid x,z) - F_{T|X,Z}^{-1}(u \mid x,z)]}{S_V(u)^2} dS_{V,1}(u) \right.
$$
$$
\left. - \int_0^v \frac{d\left(f_{V,1|X,Z}(u \mid x,z)[\widehat{F}_{T|X,Z}^{-1}(u \mid x,z) - F_{T|X,Z}^{-1}(u \mid x,z)] \right)}{S_V(u)} \right\} + o_p(n^{-1/2}),
$$

uniformly in $v \in \mathcal{I}$ with $f_{V|X,Z}$ and $f_{V,1|X,Z}$ the corresponding densities of $S_{V|X,Z}$ and $S_{V,1|X,Z}$ respectively. Let $S_{V_T}(v) = \mathbb{P}(V_T > v)$ and $S_{V_C}(v) = \mathbb{P}(V_C > v)$. Using Theorem 2 by Lo and Singh (1986) and the fact that $V_T \perp\!\!\!\perp X, Z$, this can be rewritten as

$$
\sum_{(x,z)\in\mathcal{R}_{X,Z}} p_{X,Z}(x,z) \left\{ \int_0^v \frac{f_{V|X,Z}(u \mid x,z)[\widehat{F}_{V_T|X,Z}(u \mid x,z) - F_{V_T|X,Z}(u \mid x,z)]}{S_V(u)S_{V_T}(u)} du \right.
$$
$$
\left. + \int_0^v \frac{d\left(S_{V_C|X,Z}(u \mid x,z)[\widehat{F}_{V_T|X,Z}(u \mid x,z) - F_{V_T|X,Z}(u \mid x,z)] \right)}{S_V(u)} \right\} + o_p(n^{-1/2}), \tag{13}
$$

uniformly in $v \in \mathcal{I}$. Note that we have also used that $dS_{V,1}(u) = -f_{V_T}(u)S_{V_C}(u)du$ with f_{V_T} the corresponding density of S_{V_T}. Further, let $N_{i,x,z}(v) = \mathbb{1}(V_i \geq v, X_i = x, Z_i = z)$ and $N_{i,x,z}^1(v) = \mathbb{1}(V_i \geq v, \Delta_i = 1, X_i = x, Z_i = z)$. Adapting Theorem 1 by Lo and Singh (1986) to the conditional Kaplan-Meier, we have that

$$
\widehat{F}_{V_T|X,Z}(u \mid x,z) - F_{V_T|X,Z}(u \mid x,z) = \frac{S_{V_T|X,Z}(u \mid x,z)}{\sum_{i=1}^n \mathbb{1}(X_i = x, Z_i = z)} \sum_{i=1}^n \xi_i(u,x,z) + o_p(n^{-1/2}),
$$

uniformly in $v \in \mathcal{I}$ with

$$
\xi_i(u,x,z) = \int_0^u \frac{N_{i,x,z}(s)dS_{V,1|X,Z}(s \mid x,z)}{S_{V|X,Z}(s \mid x,z)^2} - \int_0^u \frac{dN_{i,x,z}^1(s)}{S_{V|X,Z}(s \mid x,z)}.
$$

Therefore, using again that $V_T \perp\!\!\!\perp X, Z$, we have that (13) is equal to

$$
n^{-1} \sum_{(x,z)\in\mathcal{R}_{X,Z}} \left\{ \int_0^v \frac{f_{V|X,Z}(u \mid x,z) \sum_{i=1}^n \xi_i(u,x,z)}{S_V(u)} du \right.
$$
$$
\left. + \int_0^v \frac{d\left(S_{V|X,Z}(u \mid x,z) \sum_{i=1}^n \xi_i(u,x,z) \right)}{S_V(u)} \right\} + o_p(n^{-1/2}),
$$

which can be simplified to

$$
n^{-1} \sum_{(x,z)\in\mathcal{R}_{X,Z}} \left\{ \int_0^v \frac{S_{V|X,Z}(u \mid x,z) \sum_{i=1}^n d\left(\xi_i(u,x,z) \right)}{S_V(u)} \right\} + o_p(n^{-1/2}).
$$

Clearly, this is equal to

$$
n^{-1} \sum_{i=1}^n \sum_{(x,z)\in\mathcal{R}_{X,Z}} \left\{ \int_0^v \frac{N_{i,x,z}(u)dS_{V,1|X,Z}(u \mid x,z)}{S_{V|X,Z}(u \mid x,z)S_V(u)} - \frac{dN_{i,x,z}^1(u)}{S_V(u)} \right\} + o_p(n^{-1/2}). \tag{14}
$$

Note that

$$dS_{V,1|X,Z}(u \mid x, z) = -f_{V,1|X,Z}(u \mid x, z)du = -f_{V,1}(u)\frac{S_{V_C|X,Z}(u \mid x, z)}{S_{V_C}(u)}du,$$

such that (14) is equal to

$$n^{-1}\sum_{i=1}^{n}\sum_{(x,z)\in\mathcal{R}_{X,Z}}\left\{\int_0^v \frac{N_{i,x,z}(u)dS_{V,1}(u)}{S_V(u)^2} - \frac{dN_{i,x,z}^1(u)}{S_V(u)}\right\} + o_p(n^{-1/2}),$$

which exactly compensates the first two terms of (12), meaning that

$$\log\left(1 - \widehat{F}_{\widehat{V}_T}(v)\right) - \log(1 - v) = o_p(n^{-1/2}),$$

uniformly in $v \in \mathcal{I}$. Using a Taylor expansion, the result follows. We continue with the proof of (ii). Using Lemmas 2 and 3, we can follow the same steps as the proof of (i) to show that

$$\log\left(1 - \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right)$$

$$= -\int_0^v \frac{\widehat{S}_{\widehat{V}|X,W,Z}(u \mid x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)^2}dS_{V,1|X,W,Z}(u \mid x, w, z) + \int_0^v \frac{d\widehat{S}_{\widehat{V},1|X,W,Z}(u \mid x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)} + o_p(n^{-1/2}),$$

uniformly in v, x, w and z. Using Lemma 1, it follows that this can be written as

$$\widehat{p}_{X,W,Z}^{-1}(x, w, z)\left\{-\int_0^v \frac{\widehat{S}_{V,X,W,Z}(u, x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)^2}dS_{V,1|X,W,Z}(u \mid x, w, z)\right.$$

$$+\int_0^v \frac{d\widehat{S}_{V,X,W,Z,1}(u, x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)} - \int_0^v \frac{S_{\widehat{V},X,W,Z}(u, x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)^2}dS_{V,1|X,W,Z}(u \mid x, w, z)$$

$$\left.+\int_0^v \frac{dS_{\widehat{V},X,W,Z,1}(u, x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)}\right\} + o_p(n^{-1/2}),$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. Using a similar Taylor expansion as in the proof of (i) and Theorem 1 by Lo and Singh (1986), it follows that

$$\log\left(1 - \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right)$$

$$= \log\left(1 - \widehat{F}_{V_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right)$$

$$+\left\{\int_0^v \frac{\frac{f_{V|X,W,Z}(u|x,w,z)}{f_{V_T|X,Z}(u|x,z)}[\widehat{F}_{V_T|X,Z}(u \mid x, z) - F_{V_T|X,Z}(u \mid x, z)]}{S_{V|X,W,Z}(u \mid x, w, z)^2}f_{V,1|X,W,Z}(u \mid x, w, z)du\right.$$

$$\left.+\int_0^v \frac{d\left(\frac{f_{V,1|X,W,Z}(u|x,w,z)}{f_{V_T|X,Z}(u|x,z)}[\widehat{F}_{V_T|X,Z}(u \mid x, z) - F_{V_T|X,Z}(u \mid x, z)]\right)}{S_{V|X,W,Z}(u \mid x, w, z)}\right\} \times (1 + o_p(1)) + o_p(n^{-1/2}),$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. Under Assumption (A2), this is equal to

$$\log\left(1 - \widehat{F}_{V_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right)$$

$$+\left\{\frac{f_{V_T|X,W,Z}(v \mid x, w, z)[\widehat{F}_{V_T|X,Z}(v \mid x, z) - F_{V_T|X,Z}(v \mid x, z)]}{f_{V_T}(v)S_{V_T|X,W,Z}(v \mid x, w, z)}\right\} \times (1 + o_p(1)) + o_p(n^{-1/2}),$$

28

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. Using Theorem 1.2 by Gill (1983) and following similar steps to the proof of (i), we now have that

$$
\begin{aligned}
&\log\left(1 - \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right) \\
&= \log\left(1 - \widehat{F}_{V_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right) \\
&\quad + n^{-1} \sum_{i=1}^{n} \frac{f_{V_T|X,W,Z}(v \mid x, w, z) S_{V_T}(v) \xi_i(v, x, z)}{f_{V_T}(v) S_{V_T|X,W,Z}(v \mid x, w, z) p_{X,Z}(x, z)} + o_p(n^{-1/2}),
\end{aligned}
$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$ with $\xi_i(v, x, z)$ defined in the proof of (i). Using Theorem 1 by Lo and Singh (1986), we get that

$$
\begin{aligned}
&\log\left(1 - \widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z)\right) - \log\left(1 - F_{V_T|X,W,Z}(v \mid x, w, z)\right) \\
&= n^{-1} \sum_{i=1}^{n} \left\{ \frac{f_{V_T|X,W,Z}(v \mid x, w, z) S_{V_T}(v) \xi_i(v, x, z)}{f_{V_T}(v) S_{V_T|X,W,Z}(v \mid x, w, z) p_{X,Z}(x, z)} - \frac{\zeta_i(v, x, w, z)}{p_{X,W,Z}(x, w, z)} \right\} \\
&\quad + o_p(n^{-1/2}),
\end{aligned}
$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$, where

$$
\zeta_i(v, x, w, z) = \int_0^v \frac{N_{i,x,w,z}(u) dS_{V,1|X,W,Z}(u \mid x, w, z)}{S_{V|X,W,Z}(u \mid x, w, z)^2} - \int_0^v \frac{dN^1_{i,x,w,z}(u)}{S_{V|X,W,Z}(u \mid x, w, z)},
$$

with $N_{i,x,w,z}(v) = \mathbb{1}(V_i \geq v, X_i = x, W_i = w, Z_i = z)$ and $N^1_{i,x,w,z}(v) = \mathbb{1}(V_i \geq v, \Delta_i = 1, X_i = x, W_i = w, Z_i = z)$. Using a Taylor expansion, we have that

$$
\begin{aligned}
&\widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z) - F_{V_T|X,W,Z}(v \mid x, w, z) \\
&= n^{-1} \sum_{i=1}^{n} \left\{ \frac{S_{V_T|X,W,Z}(v \mid x, w, z) \zeta_i(v, x, w, z)}{p_{X,W,Z}(x, w, z)} - \frac{f_{V_T|X,W,Z}(v \mid x, w, z) S_{V_T}(v) \xi_i(v, x, z)}{f_{V_T}(v) p_{X,Z}(x, z)} \right\} \\
&\quad + o_p(n^{-1/2}),
\end{aligned}
$$

uniformly in $(v, x, w, z) \in \mathcal{I} \times \mathcal{R}_{X,W,Z}$. Note that this can be rewritten as

$$
\begin{aligned}
&\widehat{F}_{V_T|X,W,Z}(v \mid x, w, z) - F_{V_T|X,W,Z}(v \mid x, w, z) \\
&\quad - \frac{f_{V_T|X,W,Z}(v \mid x, w, z)}{f_{V_T}(v)} \left[\widehat{F}_{V_T|X,Z}(v \mid x, z) - F_{V_T|X,Z}(v \mid x, z)\right] + o_p(n^{-1/2}).
\end{aligned}
$$

Using again Theorem 1.2 by Gill (1983), the proof is completed. $\qquad\square$

B   Main proofs

Using the technical results obtained in Appendix A, we are now ready to prove the main theorems stated in Section 3.1.

B.1 Proof of Theorem 1

We start by noticing that

$$
\widehat{p}_{Z|X,W}(z \mid x, w) - p_{Z|X,W}(z \mid x, w)
$$
$$
= n^{-1} \sum_{i=1}^{n} \frac{\mathbb{1}(X_i = x, W_i = w)}{p_{X,W}(x, w)} \left[\mathbb{1}(Z_i = z) - p_{Z|X,W}(z \mid x, w) \right] + o_p(n^{-1/2}), \tag{15}
$$

uniformly in $(x, w, z) \in \mathcal{R}_{X,W,Z}$. Following similar reasoning as the proof of Theorem 1 by Akritas and Van Keilegom (2003), we have that

$$
\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - F_{V_T|X,W}(v \mid x, w)
$$
$$
= \sum_{z \in \mathcal{R}_Z} \left\{ \left[\widehat{F}_{\widehat{V}_T|X,W,Z}(v \mid x, w, z) - F_{V_T|X,W,Z}(v \mid x, w, z) \right] p_{Z|X,W}(z \mid x, w) \right.
$$
$$
\left. + F_{V_T|X,W,Z}(v \mid x, w, z) \left[\widehat{p}_{Z|X,W}(z \mid x, w) - p_{Z|X,W}(z \mid x, w) \right] \right\} + o_p(n^{-1/2}),
$$

uniformly in $(v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W}$. Using Lemma 4 and (15), we get that

$$
\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - F_{V_T|X,W}(v \mid x, w) = n^{-1} \sum_{i=1}^{n} \frac{\omega_i(v, x, w)}{p_{X,W}(x, w)} + o_p(n^{-1/2}),
$$

uniformly in $(v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W}$, where

$$
\omega_i(v, x, w) = S_{V_T|X,W,Z}(v \mid x, w, Z_i)\zeta_i(v, x, w, Z_i) - p_{W|V_T,X,Z}(w \mid v, x, Z_i)S_{V_T}(v)\xi_i(v, x, Z_i)
$$
$$
+ \mathbb{1}(X_i = x, W_i = w) \left[F_{V_T|X,W,Z}(v \mid x, w, Z_i) - F_{V_T|X,W}(v \mid x, w) \right].
$$

Using Lemma 4 again, we now have that

$$
\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - \widehat{F}_{\widehat{V}_T}(v) = n^{-1} \sum_{i=1}^{n} \frac{\omega_i(v, x, w)}{p_{X,W}(x, w)} + F_{V_T|X,W}(v \mid x, w) - v + o_p(n^{-1/2}),
$$

uniformly in $(v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W}$. $\qquad\qquad\square$

B.2 Proof of Theorem 2

To show the weak convergence of the process

$$
\left\{ n^{1/2} \left[\widehat{F}_{\widehat{V}_T|X,W}(v \mid x, w) - \widehat{F}_{\widehat{V}_T}(v) \right] : (v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W} \right\},
$$

under H_0, we will show that the function class

$$
\mathcal{M} = \left\{ (V, \Delta, X, W, Z) \rightarrow \frac{S_{V_T|X,W,Z}(v \mid x, w, Z)\zeta(v, x, w, V, \Delta, X, W, Z)}{p_{X,W}(x, w)} \right.
$$
$$
- \frac{p_{W|V_T,X,Z}(w \mid v, x, Z)(1 - v)\xi(v, x, V, \Delta, X, Z)}{p_{X,W}(x, w)}
$$
$$
\left. + \frac{\mathbb{1}(X = x, W = w)}{p_{X,W}(x, w)} \left[F_{V_T|X,W,Z}(v \mid x, w, Z) - v \right] : (v, x, w) \in \mathcal{I} \times \mathcal{R}_{X,W} \right\},
$$

30

is Donsker, where

$$\zeta(v, x, w, V, \Delta, X, W, Z) = \int_0^v \frac{\mathbb{1}(V \geq u, X = x, W = w)dS_{V,1|X,W,Z}(u \mid x, w, Z)}{S_{V|X,W,Z}(u \mid x, w, Z)^2}$$
$$+ \frac{\mathbb{1}(V \leq v, \Delta = 1, X = x, W = w)}{S_{V|X,W,Z}(V \mid x, w, Z)},$$

and

$$\xi(v, x, V, \Delta, X, Z) = \int_0^v \frac{\mathbb{1}(V \geq u, X = x)dS_{V,1|X,Z}(u \mid x, Z)}{S_{V|X,Z}(u \mid x, Z)^2} + \frac{\mathbb{1}(V \leq v, \Delta = 1, X = x)}{S_{V|X,Z}(V \mid x, Z)}.$$

Note that

$$\mathcal{M} = \bigcup_{(\tilde{x}, \tilde{w}) \in \mathcal{R}_{X,W}} \mathcal{M}^{\tilde{x}, \tilde{w}},$$

where

$$\mathcal{M}^{\tilde{x}, \tilde{w}} = \Big\{ (V, \Delta, X, W, Z) \to \frac{S_{V_T|X,W,Z}(v \mid \tilde{x}, \tilde{w}, Z)\zeta(v, \tilde{x}, \tilde{w}, V, \Delta, X, W, Z)}{p_{X,W}(\tilde{x}, \tilde{w})}$$
$$- \frac{p_{W|V_T,X,Z}(\tilde{w} \mid v, \tilde{x}, Z)(1 - v)\xi(v, \tilde{x}, V, \Delta, X, Z)}{p_{X,W}(\tilde{x}, \tilde{w})}$$
$$+ \frac{\mathbb{1}(X = \tilde{x}, W = \tilde{w})}{p_{X,W}(\tilde{x}, \tilde{w})} \big[F_{V_T|X,W,Z}(v \mid \tilde{x}, \tilde{w}, Z) - v \big] : v \in \mathcal{I} \Big\}.$$

Since the finite union of Donsker classes is again Donsker, it is sufficient to show that $\mathcal{M}^{\tilde{x}, \tilde{w}}$ is Donsker for all $(\tilde{x}, \tilde{w}) \in \mathcal{R}_{X,W}$. Using the same reasoning as in the proof of Lemma 1, $\mathcal{M}^{\tilde{x}, \tilde{w}}$ is Donsker if

$$\mathcal{M}^{\tilde{x}, \tilde{w}}_{\tilde{X}, \tilde{W}, \tilde{Z}} = \Big\{ (V, \Delta) \to \frac{S_{V_T|X,W,Z}(v \mid \tilde{x}, \tilde{w}, \tilde{Z})\zeta(v, \tilde{x}, \tilde{w}, V, \Delta, \tilde{X}, \tilde{W}, \tilde{Z})}{p_{X,W}(\tilde{x}, \tilde{w})}$$
$$- \frac{p_{W|V_T,X,Z}(\tilde{w} \mid v, \tilde{x}, \tilde{Z})(1 - v)\xi(v, \tilde{x}, V, \Delta, \tilde{X}, \tilde{Z})}{p_{X,W}(\tilde{x}, \tilde{w})}$$
$$+ \frac{\mathbb{1}(\tilde{X} = \tilde{x}, \tilde{W} = \tilde{w})}{p_{X,W}(\tilde{x}, \tilde{w})} \big[F_{V_T|X,W,Z}(v \mid \tilde{x}, \tilde{w}, \tilde{Z}) - v \big] : v \in \mathcal{I} \Big\},$$

is Donsker for all $(\tilde{X}, \tilde{W}, \tilde{Z}) \in \mathcal{R}_{X,W,Z}$. Noticing that for all $(\tilde{x}, \tilde{w}, \tilde{X}, \tilde{W}, \tilde{Z}) \in \mathcal{R}_{X,W} \times \mathcal{R}_{X,W,Z}$ we have that $\mathcal{M}^{\tilde{x}, \tilde{w}}_{\tilde{X}, \tilde{W}, \tilde{Z}}$ is a linear combination of functions that are monotone and uniformly bounded in $v \in \mathcal{I}$, the weak convergence follows from Theorem 2.7.5 by Van der Vaart and Wellner (1996) and Lemma 9.25 by Kosorok (2008). The convergence of T_n^{KS} follows from the continuous mapping theorem. Moreover, the convergence of T_n^{CM} follows from Assumption (A4) and using similar arguments as in the proof of Corollary 3.1 by Fève et al. (2018). $\qquad \square$