Biometrika (2017), **103**, 1, *pp*. 1–18
Printed in Great Britain Advance Access publication on 31 July 2016

Instrumental variable estimation of marginal structural Cox model for time-varying treatments

BY Y. CUI
Department of Statistics and Data Science, National University of Singapore
cuiy@nus.edu.sg

H. MICHAEL
Department of Mathematics and Statistics, University of Massachusetts
hmichael@math.umass.edu

F. TANSER
Lincoln Institute for Health, University of Lincoln
ftanser@lincoln.ac.uk

AND E. TCHETGEN TCHETGEN
Department of Statistics, Wharton School of the University of Pennsylvania
ett@wharton.upenn.edu

SUMMARY

Robins (1998) introduced marginal structural models, a general class of counterfactual models for the joint effects of time-varying treatments in complex longitudinal studies subject to time-varying confounding. Robins (1998) established the identification of marginal structural model parameters under a sequential randomization assumption, which rules out unmeasured confounding of treatment assignment over time. The marginal structural Cox model is one of the most popular marginal structural models to evaluate the causal effect of time-varying treatments on a censored failure time outcome. In this paper, we establish sufficient conditions for identification of marginal structural Cox model parameters with the aid of a time-varying instrumental variable, when sequential randomization fails to hold due to unmeasured confounding. Our instrumental variable identification condition rules out any interaction between an unmeasured confounder and the instrumental variable in its additive effects on the treatment process, the longitudinal generalization of the identifying condition of Wang & Tchetgen Tchetgen (2018). We describe a large class of weighted estimating equations that give rise to consistent and asymptotically normal estimators of the marginal structural Cox model, thereby extending the standard inverse probability of treatment weighted estimation of marginal structural models to the instrumental variable setting. Our approach is illustrated via extensive simulation studies and an application to estimate the effect of community antiretroviral therapy coverage on HIV incidence.

Some key words: Causal inference; Survival analysis; Marginal structural models; Unmeasured confounding; Instrumental variable; Observational studies

1. INTRODUCTION

Robins (1998, 1999) introduced a new class of counterfactual models known as marginal structural models that encode the joint causal effects of time-varying treatments subject to time-varying confounding. Marginal structural models are particularly powerful as they estimate the causal effects of time-dependent treatments in the presence of time-dependent confounders that are affected by prior treatments. For identification of marginal structural model parameters, Robins (1998, 1999) relied on a sequential randomization assumption, also known as sequential exchangeability, which rules out unmeasured confounding of treatment assignment over time. Applications of marginal structural models abound in health and social sciences, for instance, see Robins et al. (2000); Hernán et al. (2001); Cole & Hernán (2008); Cerdá et al. (2010).

Right-censored data are of common occurrence in epidemiologic studies where the clinical outcome may not always be observed due to censoring. In such settings, marginal structural models extend to time-dependent Cox models. Unlike standard time-dependent Cox models, marginal structural Cox models (Hernán et al., 2000) allow for adjustment of time-varying confounders through the use of inverse probability of treatment weighting, and thus can be utilized to estimate the causal effects of time-varying treatments in the presence of time-varying confounders; see de Keyser et al. (2014); Karim et al. (2014); Ali et al. (2016) for recent applications of marginal structural Cox models in clinical studies.

However, sequential randomization can seldom be guaranteed in observational studies, even if one adjusts for a large number of covariates in an effort to make the assumption credible. The instrumental variable method (Goldberger, 1972; Imbens & Angrist, 1994; Angrist et al., 1996; Wooldridge, 2010) is a well-known approach to estimate causal effects subject to unmeasured confounding in observational studies. An instrumental variable is defined as a pre-treatment variable that is independent of all unmeasured confounders, and does not have a direct causal effect on the outcome other than through the treatment. In a double-blind placebo-controlled randomized trial, random assignment is a common example of an ideal instrumental variable for the causal effect of treatment when some patients fail to comply to assigned treatment, provided that double-blinding is maintained. To our knowledge, Michael et al. (2020) were the first to consider identification and estimation of marginal structural mean models in the context of a time-varying treatment and a time-varying instrumental variable. However, additional challenges arise with censored survival data, which were not addressed in Michael et al. (2020).

The first formal instrumental variable approach for right-censored survival outcome was proposed by Robins & Tsiatis (1991) who parameterized the treatment effect under a structural accelerated failure time model. The approach, which applies to both point and time-varying treatments, can be challenging to implement in practice due to the need for artificial censoring in order to estimate the structural model parameters (Robins & Tsiatis, 1991; Joffe et al., 2012). More recently, Li et al. (2015); Tchetgen Tchetgen et al. (2015) considered estimating the conditional hazard difference under an additive hazard model (Aalen, 1989). Tchetgen Tchetgen et al. (2015) also considered instrumental variable estimation for a Cox structural model under rare disease. MacKenzie et al. (2014) considered instrumental variable estimation of a marginal structural Cox model for point exposure, while their estimator generally fails to be consistent as their proposed estimating equation fails to be unbiased. Martinussen et al. (2017b) developed instrumental variable estimators under a semiparametric structural cumulative survival model, which is closely related to the additive hazard model of Li et al. (2015); Tchetgen Tchetgen et al. (2015). In contrast, Martinussen et al. (2017a); Sørensen et al. (2019) considered estimating the causal hazard ratio among the treated. The literature on instrumental variable methods for complier causal effect is also well developed for survival data, e.g., Loeys et al. (2005); Cuz-

ick et al. (2007); Nie et al. (2011); Yu et al. (2015); Kianian et al. (2019). However, none of these approaches has to date been extended to time-varying settings, although some progress was made by Yende-Zuma et al. (2019) under somewhat restricted conditions including that the unmeasured confounder cannot be an effect of prior treatment.

In fact many instrumental variables that have been used in point exposure studies are often valid time-varying instrumental variables in settings where longitudinal data are available. For instance, in many well-established longitudinal observational studies such as the Multicenter AIDS Cohort Study and the Women's Interagency HIV Study, both HIV treatment assignment as well as treatment adherence information between follow-up visits are routinely collected and therefore treatment assignment is a potential time-varying instrumental variable for the causal effects of time-varying antiretroviral therapy actually taken on HIV related outcomes. Validity of treatment assignment as a valid instrumental variable relies on an assumption that confounding by indication can be fully accounted for, an assumption which is well justified in the literature (Robins et al., 2000; Hernán et al., 2001). Nowadays, even if not measured directly, such adherence information is routinely inferred from pharmacy data in large electronic medical records when studying the causal effect of time-varying treatments on a variety of disease outcomes beyond HIV. Other examples of a common instrumental variable that is also inherently time-varying include physician treatment preference as physician preferences evolve with clinical practice and needs (Brookhart & Schneeweiss, 2007), distance from nearest college as individuals move over time, distance for nearest needle exchange program (Frangakis et al., 2004), calendar period (Cain et al., 2009), as well as differential distance between nearest low level neonatal intensive care units and nearest high level neonatal intensive care units in analyses of causal effects of delivery at a high level vs low level neonatal intensive care units on birth outcomes (Zubizarreta et al., 2013; Yang et al., 2014). We acknowledge that sometimes instrumental variables may be related to initiating a treatment but not adhering to it, so one should exercise caution in selecting a valid time-varying instrumental variable.

In this paper, we consider an instrumental variable approach and establish sufficient conditions for identification of marginal structural Cox model parameters encoding the joint effects of binary time-varying treatments by leveraging a binary time-varying instrumental variable when the sequential randomization assumption fails to hold. Our identifying conditions extend those of Wang et al. (2018) to the longitudinal treatment setting, and require that in each risk set, no unobserved confounder interacts with the instrumental variable in its additive effects on the treatment process. Our proposed semiparametric estimators extend standard inverse probability of treatment weighted estimation, which is the most popular approach for estimating marginal structural Cox models under the sequential randomization assumption, by incorporating time-varying instrumental variables through a modified set of weights. We formally establish identification of our modified weighted estimating equations, and we provide the asymptotic theory which allows, in absence of model misspecification, for valid inference about the marginal structural Cox model parameters.

2. Marginal structural Cox models

2·1. *Notation*

In this section, we first introduce some notation. Continuous time is denoted by t and is measured in weeks/months since the beginning of a subject's follow-up. The index j is often used when we indicate an integer of weeks/months, and J corresponds to the administrative end of follow-up. Let $A(j)$ and $L(j)$ denote a binary treatment taken by a subject and a vector of relevant prognostic factors for survival outcome in time interval $(j, j+1]$, respectively, where

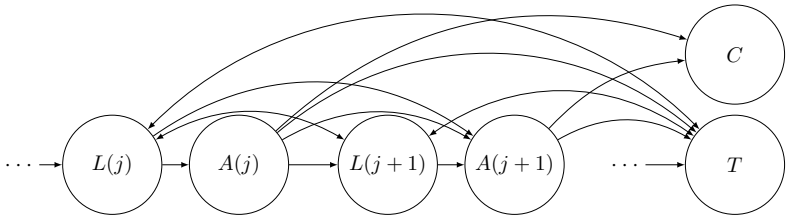

Fig. 1. A causal graph of longitudinal confounding with bi-directed arrows.

$j = 0, \ldots, J-1$. We assume that $L(j)$ temporally precedes $A(j)$. Though time t is considered to be continuous, we assume that recorded data on the treatment and prognostic factors do not change except at integer times. For any time-dependent variable X, we use $\overline{X}(t)$ to denote the history of that variable up to time t. For example, the covariate process through t is $\overline{L}(t) = (L(0), L(1), \ldots, L(\lceil t \rceil - 1))$, where $\lceil t \rceil$ denotes the smallest integer greater than or equal to t. Furthermore, let $\widetilde{C}$ denote censoring time, observed time $Y = \min(T, C)$, and censoring indicator $\delta = I(T \leq C)$, where $C = \min(\widetilde{C}, J)$. Throughout the paper, unless necessary, we suppress the subscript denoting individual because we assume that the observed data of each subject is drawn independently from a distribution common to all subjects.

Figure 1 gives a causal graph representation of longitudinal confounding. The Markov property encoded in this graph implies that a node is independent of non-descendants conditional on its parents. Bi-directed edges into the covariate process $L(j)$ and outcome T indicate the possibility that there may be unmeasured common causes confounding their association. As pointed out by Robins (1997), a standard Cox model of the joint effects of treatment on time to event, which would typically condition on $\overline{L}(j)$, is subject to collider bias and therefore is bound to incorrectly report a causal effect linking treatment to the event time even under the sharp null of no causal effect. This potential flaw of standard regression models motivated Robins (1998) to develop marginal structural model as a principled approach to circumvent this difficulty. Importantly, there are no unmeasured common causes of the treatment process with the outcome T, therefore encoding sequential randomization which we formally define in Section 2. Thus, under a data generating mechanism consistent with Figure 1, the joint causal effect of treatment on T would in fact be point identified under a nonparametric model for the observed data using Robins' g-formula (Robins, 1997).

To conclude this section, we introduce counterfactuals which are key to defining marginal structural models. Neyman (1923) proposed to use counterfactual outcomes to define the causal effect of time-independent treatments in randomized experiments. Rubin (1974) further used potential outcomes in the analysis of causal effects of time-independent treatments in observational studies. Robins (1986, 1987) proposed a formal counterfactual theory of causal inference that extended Neyman's time-independent treatment theory to longitudinal studies with both direct and indirect effects and time-varying treatments and confounders. For a specific fixed treatment history $\overline{a} \equiv \overline{a}(J-1)$, $\overline{L}_{\overline{a}}(T_{\overline{a}})$ is defined to be the random vector representing a subject's covariate process had the subject been treated with the particular treatment regime $\overline{a}$ rather than his or her observed treatment history; $T_{\overline{a}}$ is defined to be the subject's time to death had the subject been treated with the particular treatment regime $\overline{a}$. Throughout the paper, we assume that the future cannot cause the past, for example, $\lambda_{T_{\overline{a}}}(t) = \lambda_{T_{\overline{a}(\lceil t \rceil - 1)}}(t)$ and $L_{\overline{a}}(j) = L_{\overline{a}(j-1)}(j)$, $j = 1, \ldots, J-1$.

2·2. *Identification of marginal structural Cox models under the sequential randomization assumption*

Suppose that we are interested in estimating the parameter ψ_0 indexing a marginal structural Cox model, which encodes the causal effect of all potential treatment histories,

$$\lambda_{T_{\overline{a}}}(t|V) = \lambda_0\left(t\right)\exp\left\{m\left(\overline{a}\left(t\right), t, \psi_0, V\right)\right\}, \tag{1}$$

where $m(\cdot)$ is a known function which satisfies $m\left(\overline{0}(t), t, \psi_0, 0\right) = 0$, $\lambda_0(t)$ is an unspecified baseline hazard function, and $V \in L(0)$ are baseline covariates. Suppose we observe data $O = \left\{\overline{A}\left(Y\right), \overline{L}\left(Y\right), Y, \delta\right\}$. We use $\overline{A} \equiv \overline{A}\left(T\right), \overline{L} \equiv \overline{L}\left(T\right)$ to denote the treatment history and covariate history up to failure, respectively. In this section, we assume an independent censoring mechanism, i.e., $C_{\overline{a}} \perp (\overline{L}_{\overline{a}}, T_{\overline{a}})$. Three important assumptions are sufficient to the identification of ψ_0 from the observed data in marginal structural Cox models.

We make the standard consistency assumption that $T = T_{\overline{A}}$ almost surely, which states that the observed failure time corresponds to the potential failure time under a potential intervention that sets treatment process to the observed treatment history.

The next assumption is the well-known sequential randomization assumption proposed in Robins (1998),

$$T_{\overline{a}} \perp A(j) \mid \overline{A}(j-1) = \overline{a}(j-1), \overline{L}(j), T \geq j, \qquad j = 0, \ldots, J-1, \tag{2}$$

where the term $\overline{A}(-1)$ is defined as an empty set throughout the paper. This assumption generalizes the assumption of ignorable treatment assignment (Rosenbaum & Rubin, 1983) to longitudinal studies with time-varying treatments and confounders. It states that, conditional on treatment history up to time $j-1$ and recorded covariates up to time j, treatment at time j is independent of the counterfactual outcome $T_{\overline{a}}$. The sequential randomization assumption holds if all common causes of T and $A(j)$ are included in $\{\overline{L}(j), \overline{A}(j-1)\}$, therefore ruling out unmeasured confounding of the treatment process as encoded in graph of Figure 1.

Finally, we make the following treatment positivity assumption:

$$f(A(j) = a(j)|\overline{L}(j), \overline{A}(j-1), T \geq j) > 0, \quad \text{if} \quad f(\overline{L}(j), \overline{A}(j-1), T \geq j) > 0,$$

for $a(j) \in \{0, 1\}$, $j = 0, \ldots, J-1$. This assumption states that conditional on observed history, there is a positive probability of receiving either treatment value at any given time. This assumption makes it possible to draw inferences about longitudinal treatment comparisons encoded in the marginal structural models.

Define time-varying weights as

$$\overline{W}(t) = \prod_{j=0}^{\lceil t \rceil - 1} W_j = \prod_{j=0}^{\lceil t \rceil - 1} \frac{f\left(A(j)|\overline{L}(j), \overline{A}(j-1), Y \geq j\right)}{f^*\left(A(j)|V, \overline{A}(j-1), Y \geq j\right)}, \tag{3}$$

where f^* is a user-specified function of the treatment process $A(\cdot)$, Robins (1998) showed that ψ_0 in Equation (1) is the solution to the following weighted estimating equation:

$$\mathbb{E}\int dN(t)\{\overline{W}(t)\}^{-1}\left\{h\left(\overline{A}, t, V\right) - \frac{\mathbb{E}\left[h\left(\overline{A}, t, V\right)\exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\}\{\overline{W}(t)\}^{-1}I(Y \geq t)\right]}{\mathbb{E}\left[\exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\}\{\overline{W}(t)\}^{-1}I(Y \geq t)\right]}\right\} = 0,$$

where $N(t) = I(Y \leq t, \delta = 1)$ is the counting process for the outcome variable, and h is a vector-valued function which has the same dimension as the causal parameter of interest ψ_0. In particular, given the following marginal structural Cox model,

$$\lambda_{T_{\overline{a}}}(t) = \lambda_0(t)\exp\{\psi_0 a(t)\}, \tag{4}$$

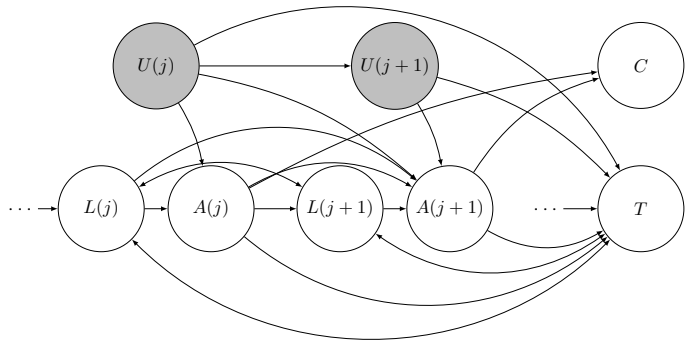

Fig. 2. A causal graph of longitudinal confounding with unobserved confounders.

and taking $h\left(t, \overline{A}, V\right) = A(t)$, ψ_0 is the solution to the following weighted estimating equation:

$$\mathbb{E}\int dN(t)\{\overline{W}(t)\}^{-1}\left\{A(t) - \frac{\mathbb{E}[A(t)\exp\{\psi A(t)\}\{\overline{W}(t)\}^{-1}I(Y \geq t)]}{\mathbb{E}[\exp\{\psi A(t)\}\{\overline{W}(t)\}^{-1}I(Y \geq t)]}\right\} = 0.$$

The weights $\overline{W}(t)$ are generally unknown and need to be estimated from the observed data. Hernán et al. (2001) proposed using pooled logistic regression to estimate a model for the treatment process in both numerator and denominator of $\overline{W}(t)$. The estimator obtained by substituting $\widehat{\overline{W}}(t)$ for $\overline{W}(t)$ and evaluating expectations under empirical distribution, is $n^{1/2}$-consistent and asymptotically linear under standard regularity conditions including $n^{1/2}$-consistency of an estimator for the treatment process $f(A(j)|\overline{L}(j), \overline{A}(j-1), T \geq j)$.

3. Instrumental variable identification and inference of marginal structural Cox model

3·1. *Identification*

It is not always possible to ensure that a sufficiently rich set of variables $\overline{L}$ was collected for sequential randomization to hold. In this vein, suppose that $U(j)$ denotes an unobserved common cause of $A(j), A(j+1), \ldots, A(\lceil T \rceil - 1)$ and T, such that Equation (2) fails. Therefore the treatment process is endogenous, i.e., subject to unmeasured confounding. On the other hand, suppose that one has observed a time-varying instrumental variable $Z(j)$ which satisfies instrumental variable conditions described below. Next, we develop an approach for leveraging such an instrumental variable process to identify and estimate the marginal structural Cox model parameter ψ_0.

A variety of instrumental variable models have been considered in existing literature in point treatment case; see Swanson et al. (2018) for a comprehensive review. In the following, we will adopt the latent counterfactual instrumental variable model described in Swanson et al. (2018) for the proposed method. Specifically, we consider a setting such as the one depicted in Figure 2, where the sequential randomization assumption fails to hold as we allow for unmeasured time-varying covariates which confound the treatment process. Suppose that $\overline{U}(j)$ is a common cause of $\underline{A}(j)$ and T, where $\underline{A}(j) = (A(j), \ldots, A(\lceil T \rceil - 1))$. Thus, we make the following assumption of latent sequential randomization:

Assumption 1. *(Latent Sequential Randomization)*
$T_{\overline{a}} \perp A(j) \mid \overline{A}(j-1) = \overline{a}(j-1), \overline{L}(j), \overline{U}(j), \overline{Z}(j), T \geq j, \qquad j = 0, \ldots, J-1.$

Denote the observed data as $\mathcal{O} = \left\{\overline{A}(Y), \overline{L}(Y), \overline{Z}(Y), Y, \delta\right\}$. A binary time-varying instrumental variable $Z(j)$ is observed just prior to $A(j)$ but after $L(j)$ for $j = 0, \ldots, J-1$, and satisfies the following time-varying instrumental variable conditions for $j = 0, \ldots, J-1$:

Assumption 2. *(Instrumental Variable Relevance)* $Z(j) \not\perp A(j) \mid H(j), T \geq j$, with history process defined as $H(j) \equiv \{\overline{A}(j-1), \overline{L}(j), \overline{Z}(j-1)\}$.

Assumption 3. *(Exclusion Restriction)*

$$\left(\overline{L}_{\overline{a}\overline{z}}, \overline{U}_{\overline{a}\overline{z}}, T_{\overline{a}\overline{z}}\right) = \left(\overline{L}_{\overline{a}}, \overline{U}_{\overline{a}}, T_{\overline{a}}\right).$$

Assumption 4. *(Instrumental Variable Independence)*

$$\left(T_{\overline{a}}, \overline{L}_{\overline{a}}, \overline{U}_{\overline{a}}\right) \perp Z(j) | \overline{A}(j-1) = \overline{a}(j-1), \overline{L}(j), \overline{Z}(j-1), T \geq j.$$

Assumption 5. *(Instrumental Variable Positivity)*

$$0 < \Pr(Z(j) = 1 | H(j), T \geq j) < 1.$$

Assumptions 2-4 are core instrumental variable conditions, while Assumption 5 is needed for nonparametric identification (Greenland, 2000; Hernán & Robins, 2006). Assumption 2 requires that the instrumental variable is associated with the treatment conditional on history process. Note that Assumption 2 does not rule out confounding of the $Z(j)$-$A(j)$ association by an unmeasured factor, however, if present, such factor must be independent of $\overline{U}(j)$. We will refer to $\overline{Z}$ as causal instrumental variables in case no such confounding is present. Assumption 3 states that there can be no direct causal effect of $Z(j)$ on $\underline{L}(j+1)$, $\underline{U}(j+1)$ and T not mediated by $\underline{A}(j)$. Assumption 4 essentially states that the null direct causal effect of $Z(j)$ on $\underline{L}(j+1)$, $\underline{U}(j+1)$, and T would be identified conditional on history process if one could intervene and set $\overline{A} = \overline{a}$.

Finally, we also require the following condition, which states that there is no additive interaction between $\overline{U}(j)$ and $Z(j)$ in a model for $A(j)$ given $\overline{U}(j)$, $H(j)$, and $T \geq j$.

Assumption 6. *(Independent Compliance Type)*

$$\mathbb{E}\left[A(j) | \overline{U}(j), H(j), T \geq j, Z(j) = 1\right] - \mathbb{E}\left[A(j) | \overline{U}(j), H(j), T \geq j, Z(j) = 0\right] = \Delta_j\left(H(j)\right). \tag{5}$$

This assumption is a longitudinal generalization of the assumption made by Wang & Tchetgen Tchetgen (2018); Wang et al. (2018) in the case of point exposure.

To interpret Assumption 6, suppose that the causal effect of $Z(j)$ on $A(j)$ is unconfounded given $\overline{U}(j), H(j)$ and $T \geq j$, i.e.,

$$Z(j) \perp A_{z(j)}(j) | \overline{U}(j), H(j), T \geq j,$$

for $j = 0, \ldots, J-1$, then

$$\mathbb{E}\left[A_{z(j)=1}(j) - A_{z(j)=0}(j) | \overline{U}(j), H(j), T \geq j\right] = \Delta_j\left(H(j)\right).$$

The causal interpretation is that while the unmeasured confounders $\overline{U}(j)$ may confound the causal effects of $\overline{A}(j)$, $\overline{U}(j)$ does not predict compliance type expressed in terms of an

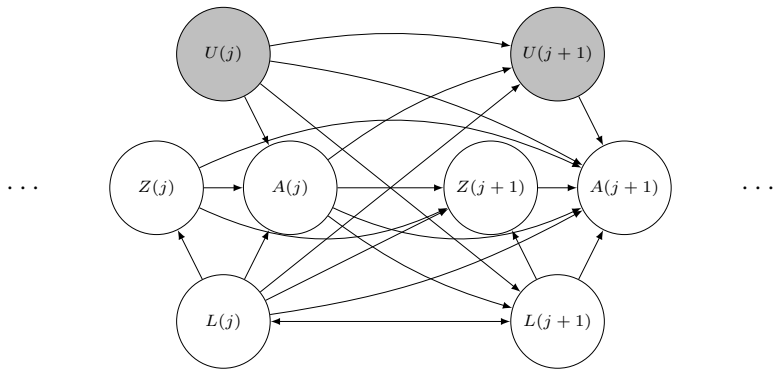

Fig. 3. A causal graph of longitudinal confounding with unobserved confounders and instrumental variables.

individual's potential treatment status under hypothetical instrumental variable interventions $\left\{A_{z(j)=1}(j), A_{z(j)=0}(j)\right\}$ at any j. Figure 3 gives a causal graph representation of longitudinal confounding with unobserved confounders $\overline{U} \equiv \overline{U}(T)$ and causal instrumental variables $\overline{Z} \equiv \overline{Z}(T)$. Note that in principle we only require Equation (5) so that $Z(j)$ may not be a causal instrumental variable, i.e., the association between $Z(j)$ and $A(j)$ may be subject to uncontrolled confounding, provided the confounder is independent of $\overline{U}$.

Finally, we make a positivity assumption regarding censoring, i.e., $S_C(J) > 0$, and the following standard independent censoring assumption.

Assumption 7. $C \perp (T, \overline{A}, \overline{Z}, \overline{U}, \overline{L})$.

It might not always be reasonable to make such independence assumption about loss-to-follow-up censoring; in fact in Section 5 we implement inverse probability of censoring weights which appropriately account for dependent censoring by time-varying covariates under the weaker condition $C \perp (\overline{U}, T) | (\overline{L}(j), \overline{A}(j), \overline{Z}(j), T \geq j)$.

3·2. *Identification of marginal structural Cox model parameter*

Before presenting the main identification result, the following lemma states that under Assumptions 4 and 6, $\Delta_j(h(j))$ is empirically identified from the observed data.

LEMMA 1. *Under Assumptions* 4 *and* 6*, we have that for* $j = 0, \ldots, J-1$,

$$\Delta_j(h(j)) = \mathbb{E}\left[A(j)|H(j) = h(j), Y \geq j, Z(j) = 1\right] - \mathbb{E}\left[A(j)|H(j) = h(j), Y \geq j, Z(j) = 0\right].$$

Next, we define the following novel time-varying weights:

$$\overline{W}^{\star}(t) \equiv \prod_{j=0}^{\lceil t \rceil - 1} W_j^{\star} = \prod_{j=0}^{\lceil t \rceil - 1} W^{\star}{}_{j,1} W^{\star}{}_{j,2}, \tag{6}$$

where

$$W_{j,1}^{\star} = \frac{f\left(Z(j)|H(j), Y \geq j\right) \Delta_j\left(H(j)\right)}{(-1)^{1-Z(j)}},$$

$$W^{\star}_{j,2} = \frac{1}{(-1)^{1-A(j)}\, f^{*}\left(A(j)|V, \overline{A}(j-1), Y \geq j\right)},$$

and f^* is user-specified, possibly empirically determined. Given that these weights can be identified from the observed data, we have the following theorem establishing that ψ_0 in marginal structural Cox model (1) is identified.

THEOREM 1. *Suppose that consistency, positivity, and Assumptions* 1–7 *hold. Then* ψ_0 *solves the population moment equation* $\mathbb{E}\{U_h(\psi)\} = 0$*, where*

$$U_h(\psi) \equiv \int \frac{dN(t)}{\overline{W}^{\star}(t)} \left\{ h\left(\overline{A}, t, V\right) - \frac{\mathbb{E}\left[h\left(\overline{A}, t, V\right) \exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t)\right]}{\mathbb{E}\left[\exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t)\right]} \right\}, \tag{7}$$

$N(t) = I(Y \leq t, \delta = 1)$*, and* $h(\overline{A}, t, V)$ *is a* $dim(\psi)$*-dimensional vector-valued function such that* $\partial\mathbb{E}[U_h(\psi)]/\partial\psi|_{\psi=\psi_0}$ *is invertible.*

3·3. *Instrumental variable based weighted estimator and large sample properties*

Theorem 1 motivates a weighted estimating equation of the parameter ψ_0 for the marginal structural Cox model defined in Equation (1) in the presence of unmeasured confounding, i.e., ψ_0 is the solution of the following population estimating equation

$$\mathbb{E}\int \frac{dN(t)}{\overline{W}^{\star}(t)} \left\{ h\left(\overline{A}, t, V\right) - \frac{\mathbb{E}\left[h\left(\overline{A}, t, V\right) \exp\left\{m\left(\overline{A}\left(t\right), t, \psi_0, V\right)\right\} \{\overline{W}^{\star}(t)\}^{-1} I(Y \geq t)\right]}{\mathbb{E}\left[\exp\left\{m\left(\overline{A}\left(t\right), t, \psi_0, V\right)\right\} \{\overline{W}^{\star}(t)\}^{-1} I(Y \geq t)\right]} \right\} = 0. \tag{8}$$

However, the weights $\overline{W}^{\star}(t)$ are unknown and need to be estimated from the observed data. We propose to use various parametric models to estimate these densities. For example, one may estimate densities $f\left(A(j)|H(j), Z(j), Y \geq j\right)$, $f^{*}(A(j)|\ V, \overline{A}(j-1), Y \geq j)$, and $f\left(Z(j)|H(j), Y \geq j\right)$ with the following logistic regressions,

$$\begin{aligned}
\text{logit}\left\{\Pr\left(A(j) = 1|H(j), Z(j), Y \geq j\right)\right\} &= \widehat{\alpha}_j^T(1, H(j), Z(j)), \\
\text{logit}\left\{\Pr^{*}\left(A(j) = 1|V, \overline{A}(j-1), Y \geq j\right)\right\} &= \widehat{\beta}_j^T(1, V, \overline{A}(j-1)), \\
\text{logit}\left\{\Pr\left(Z(j) = 1|H(j), Y \geq j\right)\right\} &= \widehat{\gamma}_j^T(1, H(j)),
\end{aligned}$$

estimated by standard maximum likelihood. Let $\widehat{\Pr}(A(j) = 1|H(j), Y \geq j, Z(j) = 1)$ and $\widehat{\Pr}(A(j) = 1|H(j), Y \geq j, Z(j) = 0)$ denote the maximum likelihood estimation of $\Pr(A(j) = 1|H(j), Z(j) = 1, Y \geq j; \widehat{\alpha}_j)$ and $\Pr(A(j) = 1|H(j), Z(j) = 0, Y \geq j; \widehat{\alpha}_j)$, respectively, and let $\widehat{f}\left(Z(j)|H(j), Y \geq j\right) \equiv f\left(Z(j)|H(j), Y \geq j; \widehat{\gamma}_j\right)$. The compliance type $\Delta_j\left(H(j)\right)$ can then be estimated by

$$\begin{aligned}
\widehat{\Delta}_j(H(j)) =& \widehat{\Pr}\Big(A(j) = 1|H(j), Y \geq j, Z(j) = 1\Big) \\
&- \widehat{\Pr}\Big(A(j) = 1|H(j), Y \geq j, Z(j) = 0\Big),\ j = 0, \ldots, J-1.
\end{aligned}$$

We denote the estimated weights as

$$\widehat{\overline{W}}^{\star}(t) = \prod_{j=0}^{\lceil t \rceil - 1} \frac{\widehat{f}\left(Z(j)|H(j), Y \geq j\right) \widehat{\Delta}_j\left(H(j)\right)}{(-1)^{2-A(j)-Z(j)}\, \widehat{f}^{*}\left(A(j)|V, \overline{A}(j-1), Y \geq j\right)}.$$

Our final estimator $\widehat{\psi}$ solves an empirical version of Equation (8), i.e., in which $\mathbb{E}$ is replaced by $\mathbb{P}_n$. By standard M-estimation theory, $\widehat{\psi}$ is asymptotically linear with first order expansion given by

$$\begin{aligned} & n^{1/2}(\widehat{\psi} - \psi_0) \\ = & -\mathbb{E}\left\{\frac{\partial U_h(\psi, \eta_0)}{\partial \psi}\Big|_{\psi=\psi_0}\right\}^{-1} \left[U_h^c(\psi_0, \eta_0) + \mathbb{E}\left(\frac{\partial U_h(\psi_0, \eta)}{\partial \eta}\Big|_{\eta=\eta_0}\right) IF_{\widehat{\eta}}\right] + o_p(1), \end{aligned}$$

where

$$U_h(\psi, \eta) \equiv \int \frac{dN(t)}{\overline{W}^{\star}(t;\eta)} \left\{ h\left(\overline{A}, t, V\right) - \frac{\mathbb{E}\left[h\left(\overline{A}, t, V\right) \exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t;\eta)\right]}{\mathbb{E}\left[\exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t;\eta)\right]} \right\},$$

$$U_h^c(\psi, \eta) \equiv \int \frac{dN(t) - I(Y \geq t) d\Lambda_0(t)}{\overline{W}^{\star}(t;\eta)} \left\{ h\left(\overline{A}, t, V\right) - \frac{\mathbb{E}\left[h\left(\overline{A}, t, V\right) \exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t;\eta)\right]}{\mathbb{E}\left[\exp\left\{m\left(\overline{A}\left(t\right), t, \psi, V\right)\right\} I(Y \geq t)/\overline{W}^{\star}(t;\eta)\right]} \right\},$$

$\eta = \{(\alpha_j, \gamma_j), j = 0, \ldots, J-1\}$, $IF_{\widehat{\eta}}$ is the influence function of $\widehat{\eta}$, and Λ_0 is the baseline cumulative hazard function. Confidence intervals can then be constructed either using an estimate of asymptotic variance of $\widehat{\psi}$ based on the influence function representation given above or via the nonparametric bootstrap.

4. Simulation studies

In this section, we report simulation studies in a two occasion setting $J = 2$ to compare the proposed estimator with existing estimators. As a benchmark, we considered an oracle weighted estimator of marginal structural Cox models which used correctly specified weight $f\left(A(j)|\overline{L}(j), \overline{U}(j), \overline{A}(j-1), Y \geq j\right)$ rather than $f\left(A(j)|\overline{L}(j), \overline{A}(j-1), Y \geq j\right)$ in Equation (3). This oracle estimator is clearly not feasible in practice because $\overline{U}$ would not be observed. In addition, we implemented both a marginal structural Cox model estimated via inverse probability of treatment weighting incorrectly assuming the sequential randomization assumption given $\overline{L}$ process, and a time-varying Cox model which directly adjusted for $L(Y)$ in the regression model.

Generating failure time outcomes under a specific marginal structural Cox model is not straightforward. We adopted the approach of Tchetgen Tchetgen (2006) which we outline in the next section.

4·1. *Generating potential outcomes under a marginal structural Cox model*

Let $T_{\mathcal{A}} = \{\overline{T}_{\overline{a}} : \overline{a} \in \mathcal{A}\}$, $\mathcal{A} = \{(0,0), (0,1), (1,0), (1,1)\}$, denote a person's set of potential outcomes. Under our identifying assumptions and further imposing

$$\begin{aligned} & f(L(j), U(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}}) \\ = & f(L(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}}) f(U(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}}), \end{aligned}$$

the full data $\{T_{\mathcal{A}}, \overline{A}, \overline{L}, \overline{U}, \overline{Z}\}$ have a joint likelihood that factorizes as:

$$f_{T_{\mathcal{A}}}(T_{\mathcal{A}}) \prod_{j=0}^{J-1} f(L(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}})$$
$$\prod_{j=0}^{J-1} f(U(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}})$$
$$\prod_{j=0}^{J-1} f(Z(j)|\overline{A}(j-1), \overline{L}(j), \overline{Z}(j-1)) \prod_{j=0}^{J-1} f(A(j)|\overline{A}(j-1), \overline{L}(j), \overline{U}(j), \overline{Z}(j)).$$

Suppose we wish to generate $T_{\mathcal{A}}$ under the following marginal structural Cox model,

$$\lambda_{T_{a_0,a_1}}(t) = \lambda \exp\{\psi_0 a(t)\}, \tag{9}$$

where $a(t) = a_0$ if $t \leq 1$ and $a(t) = a_1$ if $t > 1$. Lemma S1 presented in the Supplementary Material shows that by generating $T_{0,0}$ from an exponential density function with constant hazard λ, then defining $T_{\mathcal{A}}$ under the following accelerated failure time model,

$$T_{0,0} = \int_0^{T_{a_0,a_1}} \exp\{\psi_0 a(t)\} dt,$$

implies the marginal structural Cox model in Equation (9). We further specified

$$f(L(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}}) = f(L(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{0,0}),$$
$$f(U(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{\mathcal{A}}) = f(U(j)|\overline{A}(j-1), \overline{L}(j-1), \overline{U}(j-1), \overline{Z}(j-1), T_{0,0}),$$

where $j = 0, 1$. Upon generating $T_{\mathcal{A}}$, we simulated the processes $\overline{L}, \overline{U}, \overline{Z}, \overline{A}$, respectively. Censoring time was generated independently, and the observed time and censoring indicator were defined as $Y = \min(T, C)$, and $\delta = I(T \leq C)$, respectively.

4·2. *Simulation settings*

We considered two settings, with $\psi_0 = -1$ for the first scenario and $\psi_0 = 0$ for the second scenario. For both scenarios, the random errors ϵ's were generated from $N(0, 0.5^2)$. The baseline hazard $\lambda_0(t) = 1$. For convenience, in a slight abuse of notation, we omit the input argument of $\Delta_j(H(j))$ and write Δ_j hereinafter. The data generating mechanism for treatment and survival time is described as follows:

$L(0) = \mu_{L(0)} + \epsilon_1$, where $\mu_{L(0)} = 1.5T_{0,0}$.

$U(0) = \mu_{U(0)} + \epsilon_2$, where $\mu_{U(0)} = 1.5T_{0,0}$.

$Z(0)|L(0) \sim$ Bernoulli$(\Phi\{-0.5 + 0.4L(0)\})$, where Φ is cumulative distribution function of standard normal distribution.

$A(0)|L(0), U(0), Z(0) \sim$ Bernoulli$(\Phi\{0.2L(0) - 0.8U(0)\}(1 - \Delta_0) + Z(0)\Delta_0)$, where $\Delta_0 = \Phi\{0.2L(0)\}$.

$L(1) = \mu_{L(1)} + \epsilon_3$, where $\mu_{L(1)} = L(0) - U(0) + 1.5A(0) + T_{0,0}$.

$U(1) = \mu_{U(1)} + \epsilon_4$, where $\mu_{U(1)} = 0.5L(0) - U(0) + A(0) + T_{0,0}$.

$Z(1)|L(1), A(0), Z(0) \sim$ Bernoulli$(\Phi\{-1 + 0.2L(1) + 0.2A(0) + 0.2Z(0)\})$.

$A(1)|L(1), U(1), Z(1) \sim$ Bernoulli$(\Phi\{0.2L(1) - 0.8U(1)\}(1 - \Delta_1) + Z(1)\Delta_1)$, where $\Delta_1 = \Phi\{0.2L(1)\}$.

Censoring time was generated by $C = \min(\widetilde{C}, 2)$, where $\widetilde{C}$ followed $U(0, 4)$ yielding censoring rate approximately equal to 40% for the first scenario, and 30% for the second scenario. We also performed a sensitivity analysis for different censoring rates for each scenario. These

additional results are presented in the Supplementary Material. We considered different sample sizes $n = 2000, 4000, 8000$. Each simulation was repeated 500 times. We performed a sensitivity analysis for violation of various instrumental variable assumptions for each scenario. These additional results are presented in the Supplementary Material. In addition, we examined the empirical coverage of 95% confidence intervals of ψ_0 for both scenarios at sample size $n = 1000$. Confidence intervals were obtained by nonparametric bootstrap with 200 replications.

Both standard inverse probability of treatment weighted estimator and our proposed weighted estimator of marginal structural Cox model (4) require estimation of weights. The density functions $f\left(A(j)|\overline{L}(j), \overline{A}(j-1), Y \geq j\right)$ and $f^*\left(A(j)|V, \overline{A}(j-1), Y \geq j\right)$ in Equations (3) and (6) were estimated via maximum likelihood of time-specific logistic regression models,

$$\text{logit}\left\{\Pr\left(A(j) = 1|\overline{L}(j), \overline{A}(j-1), Y \geq j\right)\right\} = \widehat{\xi}_j^T(1, L(j), A(j-1)),$$
$$\text{logit}\left\{\Pr^*\left(A(j) = 1|V, \overline{A}(j-1), Y \geq j\right)\right\} = \widehat{\beta}_j^T(1, A(j-1)),$$

where $j = 0, 1$. The conditional density of Z was estimated via maximum likelihood estimation using the probit model,

$$\text{probit}\left\{\Pr\left(Z(j) = 1|H(j), Y \geq j\right)\right\} = \widehat{\gamma}_j^T(1, L(j), A(j-1), Z(j-1)),$$

where probit link function is the inverse standard normal cumulative distribution function. The data generating mechanism specified a model $\Pr(A(j) = 1|L(j), Z(j)) = \Phi(L(j))(1 - \Delta_j) + Z(j)\Delta_j$. The estimation of compliance type was therefore performed by maximum likelihood of Bernoulli condition on observations still at risk at time j,

$$l(\theta_j, \alpha_j|Y \geq j) = \log\Big[\Phi\left\{\theta_j^T(1, L(j))\right\}\left[1 - \Phi\left\{\alpha_j^T(1, L(j))\right\}\right] + Z(j)\Phi\left\{\alpha_j^T(1, L(j))\right\}\Big]A(j)$$
$$+ \log\Big(1 - \Big[\Phi\left\{\theta_j^T(1, L(j))\right\}\left[1 - \Phi\left\{\alpha_j^T(1, L(j))\right\}\right] + Z(j)\Phi\left\{\alpha_j^T(1, L(j))\right\}\Big]\Big)\Big[1 - A(j)\Big],$$

where $j = 0, 1$. Thus, the estimated compliance type $\widehat{\Delta}_j = \Phi\{\widehat{\alpha}_j^T(1, L(j))\}$, where $(\widehat{\theta}_j, \widehat{\alpha}_j)$ is the maximizer of $l(\theta_j, \alpha_j)$.

4·3. *Simulation results*

We report both squared bias and mean squared error of the estimated causal parameter in Figure 4 for scenario 1, Figure 5 for scenario 2, respectively. In both figures, "cox" denotes a standard time-varying Cox model which adjusted for $L(Y)$ directly in the regression; "sra" denotes standard inverse probability of treatment weighted estimation of marginal structural Cox model which assumes the sequential randomization assumption; "iv" denotes the proposed instrumental variable estimator; "sra.o" denotes the oracle inverse probability of treatment weighted estimation of marginal structural Cox model which includes U_0 and U_1 in the treatment model.

From the left panels of Figures 4 and 5, we see that Cox propositional hazard model and standard weighted estimation of marginal structural Cox model have severe bias. The former because there exist time-dependent confounders that are affected by previous treatments and unmeasured confounding; the latter because of unmeasured confounding. Our proposed estimator outperforms both Cox propositional hazard model and standard weighted estimation of marginal structural Cox model in terms of bias and mean squared error. The oracle weighted estimator of marginal structural Cox model performs as well as the proposed method, and both bias and mean squared error converge to zero as sample size increases. This confirms that our proposed instrumental variable approach performs nearly as well as the infeasible inverse probability of treatment weighted estimator had $\overline{U}$ been observed both in terms of bias and can outperform the latter in terms of efficiency.

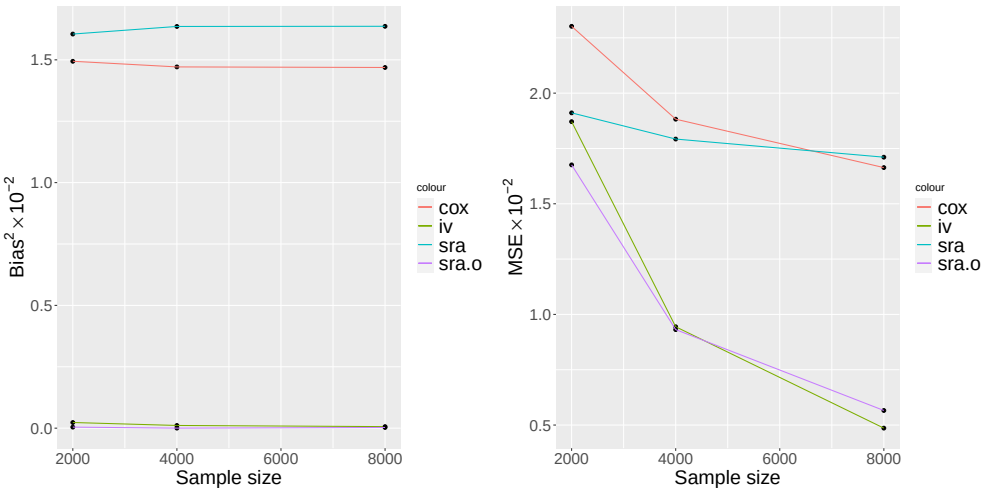

Fig. 4. Scenario 1: Monte Carlo squared bias and mean squared error of $\widehat{\psi}$, respectively

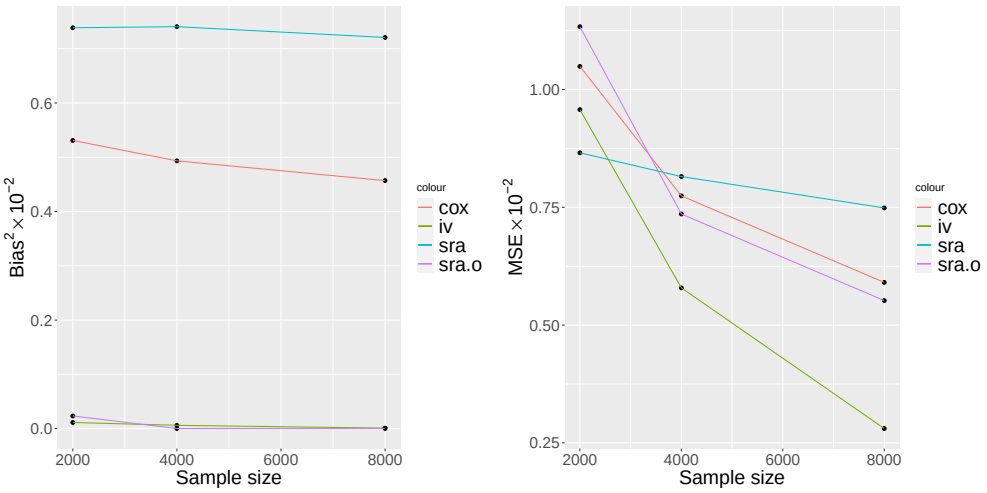

Fig. 5. Scenario 2: Monte Carlo squared bias and mean squared error of $\widehat{\psi}$, respectively

Table 1 presents empirical coverage of 95% nonparametric bootstrap confidence intervals. The proposed instrumental variable method and oracle inverse probability of treatment estimator achieved the nominal coverage. However, the confidence intervals of standard weighted estimation of marginal structural Cox model and standard time-varying Cox proportional hazard model both failed to attain nominal coverage.

Table 1. Coverage of 95% confidence intervals (%)

	iv	sra	sra.o	cox
Scenario 1 ($\psi_0 = -1$)	96.4	87.4	93.6	67.2
Scenario 2 ($\psi_0 = 0$)	97.2	87.2	94.2	66.4

5. DATA APPLICATION ESTIMATING THE EFFECT OF COMMUNITY ANTIRETROVIRAL THERAPY COVERAGE ON HIV ACQUISITION

We applied the proposed method to an HIV study analyzed in Tanser et al. (2013), which found evidence that significant reduction of HIV incidence can be achieved by nurse-led, devolved, public-sector antiretroviral therapy (ART) programs in rural sub-Saharan African settings where complete coverage of therapy under existing treatment guidelines has not yet been attained. Their analysis was based on a standard time-varying Cox proportional model. Our goal was to examine whether unmeasured confounding biased the reported association between high coverage of ART and the decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa.

We reanalyzed the dataset considered in Tanser et al. (2013), one of Africa's largest population-based prospective cohort studies to follow up individuals who were HIV-uninfected at baseline. Our analysis was restricted to 6093 individuals who were enrolled in the study and were known to be HIV negative on $2008/06/05$ with complete covariate and instrumental variable data. The objective of our analysis was to determine the joint effects of living in a high coverage community at two time occasions, $t_0 = 2008/06/05$ and $t_1 = 2011/01/01$ on HIV incidence. The overall cumulative proportion of events for the outcome was 6.3%. We considered the following six time-varying covariates: Number of partners in the past 12 months; Current marital status; Wealth index in quintiles; Age and gender; Location of residence; Community HIV prevalence. ART coverage was defined as the proportion of all HIV-infected individuals receiving ART at every location (Tanser et al., 2013). HIV prevalence and ART coverage of an individual's surrounding community were determined for every year of observation. ART coverage and HIV prevalence around each individual were measured by means of a moving two-dimensional Gaussian kernel of three kilometers search radius for each year of observation (Tanser et al., 2013). Here ART coverage was dichotomized at 30%, $A = 1$ if ART coverage $\geq 30\%$, and 0 otherwise, such that 40% of A equal to 1 over person-time.

Travel distance to the nearest ART facility defined our instrumental variable, $Z = 1$ if travel distance to nearest ART facility was less than 3.8 km and $Z = 0$ otherwise, such that 65% of $Z = 1$ over person-time. Travel distance to the closest ART facility was found to be strongly associated with ART coverage. Adjusted log odds ratios (95% confidence intervals) for the association between travel distance to the closest ART facility and community ART coverage were $1.22\ (1.02, 1.42)$ and $0.79\ (0.52, 1.05)$ at t_0 and t_1, respectively, which justify instrumental variable relevance Assumption 2. Furthermore, it was reasonable to assume that the mechanism by which local density of ART clinic, and therefore travel distance to nearest ART clinic might impact HIV incidence, was primarily through ART coverage, thus the exclusion restriction Assumption 3 holds.

We specified the marginal structural model given by (4), and estimated the various models needed to construct standard inverse probability of treatment weighted weights as well as our proposed instrumental variable weights under the following model specification:

$$\begin{aligned}
\text{logit}\left\{\Pr\left(A(j)=1|\overline{L}(j),\overline{A}(j-1),Y\geq j\right)\right\} &= \widehat{\xi}_j^T(1,L(j),A(j-1)),\\
\text{logit}\left\{\Pr\left(A(j)=1|H(j),Z(j),Y\geq j\right)\right\} &= \widehat{\alpha}_j^T(1,L(j),A(j-1),Z(j)),\\
\text{logit}\left\{\Pr^*\left(A(j)=1|V,\overline{A}(j-1),Y\geq j\right)\right\} &= \widehat{\beta}_j^T(1,V,A(j-1)),\\
\text{logit}\left\{\Pr\left(Z(j)=1|H(j),Y\geq j\right)\right\} &= \widehat{\gamma}_j^T(1,L(j),A(j-1),Z(j-1)).
\end{aligned}$$

We also considered incorporating weights which account for dependent censoring in the Supplementary Material, by adopting the approach of Robins & Rotnitzky (1992) to our approach under assumption $C \perp (\overline{U},T)|(\overline{L}(j),\overline{A}(j),\overline{Z}(j),T\geq j)$. This entails multiplying $\widehat{\overline{W}}^\star(t)$ by an

estimate of

$$\overline{W}_C(t) = \prod_{j=0}^{\lceil t \rceil - 1} \frac{\Pr(C > j | \overline{A}(j-1), \overline{L}(j-1), \overline{Z}(j-1), C > j-1)}{\Pr^*(C > j | V, \overline{A}(j-1), C > j-1)}.$$

Results with or without censoring weights were similar indicating no evidence of dependent censoring. We obtained point estimates and 95% confidence intervals of hazard ratio with the nonparametric bootstrap with 1000 replications for standard weighted estimation and instrumental variable estimation of marginal structural Cox model, respectively. To alleviate the impact of extreme values of weights, we truncated 2.5th and 97.5th percentiles of the weights for the proposed estimator (Cole & Hernán, 2008). The histogram of weights and results of untruncated weights are reported in Section S10 of the Supplementary Material. Untruncated instrumental variable weighted estimator has a similar point estimate as the truncated instrumental variable weighted estimator but wider confidence intervals due to outliers in distribution of estimated weights. As can be seen from Table 2, the instrumental variable point estimate is much smaller, which we interpret, under our instrumental variable assumptions, as appropriately accounting for unmeasured confounding, suggesting that Tanser et al. (2013) might have underestimated the true effect of HIV coverage.

Table 2. The effect of community ART coverage on HIV acquisition

	sra	iv*
Hazard ratio	0.45	0.19
95% confidence intervals	(0.20,0.92)	(0.06,0.83)

* Truncation at 2.5th and 97.5th percentiles

6. Discussion

The proposed method may be improved or extended in multiple ways. It might be of interest to look at truncation of the weights in simulations. Formal justification for truncating weights as a means for stabilization of inverse probability weighting analyses is currently lacking and represents a fruitful avenue of future research. Another potential extension is in the direction of semiparametric efficiency and enhanced robustness to partial model misspecification of nuisance parameters. The efficient influence function for the proposed marginal structural model is significantly more complicated than that in the marginal structural mean model (Tchetgen Tchetgen et al., 2018) and beyond the scope of this paper.

Acknowledgements

We thank the two reviewers, associate editor, and editor for many useful comments which led to an improved manuscript. Yifan Cui is supported by NUS grant R-155-000-229-133 and AcRF grant R-155-000-231-114. Eric Tchetgen Tchetgen is supported by NIH grants R01CA222147, R01AG065276, R01GM139926, and R01AI127271.

Supplementary Material

Supplementary material available at *Biometrika* online includes all proofs and additional simulation scenarios.

REFERENCES

AALEN, O. O. (1989). A linear regression model for the analysis of life times. *Statistics in Medicine* **8**, 907–925.

ALI, M. S., GROENWOLD, R. H. H., BELITSER, S. V., SOUVEREIN, P. C., MARTIN, E., GATTO, N. M., HUERTA, C., GARDARSDOTTIR, H., ROES, K. C. B., HOES, A. W., DE BOER, A. & KLUNGEL, O. H. (2016). Methodological comparison of marginal structural model, time-varying cox regression, and propensity score methods: the example of antidepressant use and the risk of hip fracture. *Pharmacoepidemiology and Drug Safety* **25**, 114–121.

ANGRIST, J. D., IMBENS, G. W. & RUBIN, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* **91**, 444–455.

BROOKHART, M. A. & SCHNEEWEISS, S. (2007). Preference-based instrumental variable methods for the estimation of treatment effects: assessing validity and interpreting results. *The international journal of biostatistics* **3**.

CAIN, L. E., COLE, S. R., GREENLAND, S., BROWN, T. T., CHMIEL, J. S., KINGSLEY, L. & DETELS, R. (2009). Effect of highly active antiretroviral therapy on incident aids using calendar period as an instrumental variable. *American journal of epidemiology* **169**, 1124–1132.

CERDÁ, M., DIEZ-ROUX, A. V., TCHETGEN TCHETGEN, E. J., GORDON-LARSEN, P. & KIEFE, C. I. (2010). The relationship between neighborhood poverty and alcohol use: estimation by marginal structural models. *Epidemiology* **21 4**, 482–9.

COLE, S. R. & HERNÁN, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology* **168 6**, 656–64.

CUZICK, J., SASIENI, P., MYLES, J. & TYRER, J. (2007). Estimating the effect of treatment in a proportional hazards model in the presence of non-compliance and contamination. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**, 565–588.

DE KEYSER, C. E., LEENING, M. J. G., ROMIO, S. A., JUKEMA, J. W., HOFMAN, A., IKRAM, M. A., FRANCO, O. H., STIJNEN, T. & STRICKER, B. H. (2014). Comparing a marginal structural model with a cox proportional hazard model to estimate the effect of time-dependent drug use in observational studies: statin use for primary prevention of cardiovascular disease as an example from the rotterdam study. *European Journal of Epidemiology* **29**, 841–850.

FRANGAKIS, C. E., BROOKMEYER, R. S., VARADHAN, R., SAFAEIAN, M., VLAHOV, D. & STRATHDEE, S. A. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association* **99**, 239–249.

GOLDBERGER, A. S. (1972). Structural equation methods in the social sciences. *Econometrica* **40**, 979–1001.

GREENLAND, S. (2000). An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* **29**, 722–729.

HERNÁN, M., BRUMBACK, B. & ROBINS, J. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology (Cambridge, Mass.)* **11**, 561–70.

HERNÁN, M. & ROBINS, J. (2006). Instruments for causal inference: An epidemiologist's dream? *Epidemiology (Cambridge, Mass.)* **17**, 360–72.

HERNÁN, M. A., BRUMBACK, B. & ROBINS, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association* **96**, 440–448.

IMBENS, G. W. & ANGRIST, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* **62**, 467–475.

JOFFE, M. M., YANG, W. P. & FELDMAN, H. (2012). G-estimation and artificial censoring: Problems, challenges, and applications. *Biometrics* **68**, 275–286.

KARIM, M. E., GUSTAFSON, P., PETKAU, J., ZHAO, Y., SHIRANI, A., KINGWELL, E., EVANS, C., VAN DER KOP, M., OGER, J. & TREMLETT, H. (2014). Marginal structural cox models for estimating the association between β-interferon exposure and disease progression in a multiple sclerosis cohort. *American journal of epidemiology* **180**, 160–171.

KIANIAN, B., KIM, J. I., FINE, J. P. & PENG, L. (2019). Causal proportional hazards estimation with a binary instrumental variable. *arXiv:1901.11050* .

LI, J., FINE, J. & BROOKHART, A. (2015). Instrumental variable additive hazards models. *Biometrics* **71**, 122–130.

LOEYS, T., GOETGHEBEUR, E. & VANDEBOSCH, A. (2005). Causal proportional hazards models and time-constant exposure in randomized clinical trials. *Lifetime Data Analysis* **11**, 435–449.

MACKENZIE, T. A., TOSTESON, T. D., MORDEN, N. E., STUKEL, T. A. & O'MALLEY, A. J. (2014). Using instrumental variables to estimate a cox's proportional hazards regression subject to additive confounding. *Health Services and Outcomes Research Methodology* **14**, 54–68.

MARTINUSSEN, T., NØRBO SØRENSEN, D. & VANSTEELANDT, S. (2017a). Instrumental variables estimation under a structural Cox model. *Biostatistics* **20**, 65–79.

MARTINUSSEN, T., VANSTEELANDT, S., TCHETGEN TCHETGEN, E. J. & ZUCKER, D. M. (2017b). Instrumental variables estimation of exposure effects on a time-to-event endpoint using structural cumulative survival models. *Biometrics* **73**, 1140–1149.

MICHAEL, H., CUI, Y., LORCH, S. & TCHETGEN TCHETGEN, E. (2020). Instrumental variable estimation of marginal structural mean models for time-varying treatment. *arXiv:2004.11769* .

NEYMAN, J. (1923). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science* **5**, 463–480. 1990; transl. by D.M. Dabrowska and T.P. Speed.

NIE, H., CHENG, J. & SMALL, D. S. (2011). Inference for the effect of treatment on survival probability in randomized trials with noncompliance and administrative censoring. *Biometrics* **67**, 1397–1405.

ROBINS, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period - application to control of the healthy worker survivor effect. *Mathematical Modelling* **7**, 1393 – 1512.

ROBINS, J. (1987). "errata to a new approach to causal inference in mortality studies with a sustained exposure period - application to control of the healthy worker survivor effect" mathl modelling 7(9-12), 1393-1512 (1986)". *Computers & Mathematics with Applications* **14**, 917 – 921.

ROBINS, J. (1999). Marginal structural models versus structural nested models as tools for causal inference. *Statistical Models in Epidemiology: The Environment and Clinical Trials* , 95–134.

ROBINS, J. M. (1997). Causal inference from complex longitudinal data. In *Latent Variable Modeling and Applications to Causality*, M. Berkane, ed. New York, NY: Springer New York.

ROBINS, J. M. (1998). Marginal structural models. *1997 Proceedings of the American Statistical Association. Section on Bayesian Statistical Science* , 1–10.

ROBINS, J. M., HERNÁN, M. A. & BRUMBACK, B. A. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* **11 5**, 550–60.

ROBINS, J. M. & ROTNITZKY, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS epidemiology*. Springer, pp. 297–331.

ROBINS, J. M. & TSIATIS, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics - Theory and Methods* **20**, 2609–2631.

ROSENBAUM, P. R. & RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55.

RUBIN, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66**.

SØRENSEN, D. N., MARTINUSSEN, T. & TCHETGEN, E. T. (2019). A causal proportional hazards estimator under homogeneous or heterogeneous selection in an iv setting. *Lifetime data analysis* **25**, 639–659.

SWANSON, S. A., HERNÃ¡N, M. A., MILLER, M., ROBINS, J. M. & RICHARDSON, T. S. (2018). Partial identification of the average treatment effect using instrumental variables: Review of methods for binary instruments, treatments, and outcomes. *Journal of the American Statistical Association* **113**, 933–947. PMID: 31537952.

TANSER, F., BARNIGHAUSEN, T., GRAPSA, E., ZAIDI, J. & NEWELL, M.-L. (2013). High coverage of art associated with decline in risk of hiv acquisition in rural kwazulu-natal, south africa. *Science* **339**.

TCHETGEN TCHETGEN, E. J. (2006). Statistical methods for robust inference in causal and missing data models. *Ph.D. dissertation* .

TCHETGEN TCHETGEN, E. J., MICHAEL, H. & CUI, Y. (2018). Marginal structural models for time-varying endogenous treatments: A time-varying instrumental variable approach. *arXiv:1809.05422* .

TCHETGEN TCHETGEN, E. J., WALTER, S., VANSTEELANDT, S., MARTINUSSEN, T. & GLYMOUR, M. M. (2015). Instrumental variable estimation in a survival context. *Epidemiology* **26 3**, 402–10.

WANG, L. & TCHETGEN TCHETGEN, E. (2018). Bounded, efficient and triply robust estimation of average treatment effects using instrumental variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **80**, 531–550.

WANG, L., TCHETGEN TCHETGEN, E. J., MARTINUSSEN, T. & VANSTEELANDT, S. (2018). Learning causal hazard ratio with endogeneity. *arXiv: 1807.05313* .

WOOLDRIDGE, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

YANG, F., LORCH, S. A., SMALL, D. S. et al. (2014). Estimation of causal effects using instrumental variables with nonignorable missing covariates: application to effect of type of delivery nicu on premature infants. *The Annals of Applied Statistics* **8**, 48–73.

YENDE-ZUMA, N., MWAMBI, H. & VANSTEELANDT, S. (2019). Adjusting the effect of integrating antiretroviral therapy and tuberculosis treatment on mortality for noncompliance: A time-varying instrumental variables analysis. *Epidemiology (Cambridge, Mass.)* **30**, 197–203.

YU, W., CHEN, K., SOBEL, M. E. & YING, Z. (2015). Semiparametric transformation models for causal inference in time-to-event studies with all-or-nothing compliance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77**, 397–415.

ZUBIZARRETA, J. R., SMALL, D. S., GOYAL, N. K., LORCH, S., ROSENBAUM, P. R. et al. (2013). Stronger instruments via integer programming in an observational study of late preterm birth outcomes. *The Annals of Applied Statistics* **7**, 25–50.

[*Received* 2 *January* 2017. *Editorial decision on* 1 *April* 2017]