BIOMETRICS xx, 1–22 DOI: 000
December xxxx

Instrumental Variable Based Estimation Under the Semiparametric Accelerated Failure Time Model

Jared Huling[1], Menggang Yu[2,*], and A. James O'Malley[3]

[1]Department of Statistics, The Ohio State University, Columbus, Ohio, U.S.A.

[2]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.

[3]Department of Biomedical Data Science and The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine, Dartmouth, Hanover, New Hampshire, U.S.A.

**email:* meyu@biostat.wisc.edu

SUMMARY: Randomized controlled trials are the gold standard for estimating causal effects of treatments or interventions, but in many cases are too costly, too difficult, or even unethical to conduct. Hence, many pressing medical questions can only be investigated using observational studies. However, direct statistical modeling of observational data can result in biased estimates of treatment effects due to unmeasured confounding. In certain cases, instrumental variable based techniques can be used to remove such biases. These techniques are indeed widely studied and used in econometrics under parametric outcome models, however limited works have focused on the utilization of instrumental variables in survival analysis, where semiparametric models are often necessary. The additional challenge in analyzing survival data is the presence of censoring. In this paper, we introduce an instrumental variable method that relaxes the strong assumptions of previous works and provides consistent estimation of the causal effect of a treatment on a survival outcome. We demonstrate the efficacy of our method in various simulated settings and an analysis of Medicare enrollment data comparing two prevalent surgical procedures for abdominal aortic aneurysm from an observational study.

KEY WORDS: Causal inference; Instrumental Variables; Observational data; Survival analysis.

This paper has been submitted for consideration for publication in *Biometrics*

1. Introduction

In medical research, randomized controlled trials (RCTs) are the standard for evaluating the causal effect of a treatment or intervention on a relevant outcome. While RCTs have resulted in invaluable medical and scientific knowledge, they are not always feasible to conduct due to cost, implementation difficulty, or ethical concerns. Furthermore, the scope of RCTs can often be narrow. As a result, many pressing medical questions must be addressed using observational studies. However, the estimation of causal effects from observational data is materially hindered by bias due to confounding. Techniques such as propensity score matching (Rosenbaum and Rubin, 1983; D'Agostino, 1998) can mitigate bias due to observed confounders but cannot account for unmeasured confounding, which is often the norm rather than the exception.

Instrumental variable (IV) based methods are a common approach to handling bias due to unmeasured confounding. They have found widespread use and impact in economics and have recently been increasingly utilized in medical applications. A simple example of an IV is the assignment of a treatment in a clinical trial. The assignment of a treatment is randomized and strongly related to the actual treatment received and hence can be viewed as an imperfect measure of the treatment status. Furthermore it is not contaminated by unmeasured confounding. Another example of an IV comes from the motivating application for this paper, in which we seek to compare the efficacy of two different surgical repair procedures, open vs. endovascular, for abdominal aortic aneurysm. Clinical trials comparing the two procedures (Prinssen et al., 2004; Lederle et al., 2012) are underpowered and lack the long-term follow-up necessary for a thorough comparison. To overcome these concerns, we utilize a large Medicare enrollment dataset with long-term follow-up. Hospital-level preferences for one type of surgery over another, measured by the proportion of endovascular repairs over the year prior to surgery, are highly predictive of the actual procedure received and thus we seek to use it in this work as an IV. This proportion has also been used as an IV in O'Malley et al. (2011).

It is well-known that IV estimators cannot identify causal effects without additional assumptions (Pearl, 1995). Therefore most existing methods make two types of modeling assumptions (Rubin, 2008): those on the scientific or outcome model and those on the IV model. In the classical IV methods, linear regression models are assumed for the (continuous) outcome whereas the IV assumptions, known as independence and monotonicity conditions, are very general. In particular, noncompliance in the clinical trial setting has been frequently dealt with under both these two types of conditions (Angrist, Imbens, and Rubin, 1996; Clarke and Windmeijer, 2012; Robins and Tsiatis, 1991). Moving beyond linear outcome models, recent work has extended IV methods to nonlinear parametric outcome models (Terza, Basu, and Rathouz, 2008). However, a fairly restrictive assumption was imposed on the IV model. In particular, the relationship between IV and treatment was assumed to be deterministic to facilitate a class of two-stage estimators. Tan (2006) and Clarke and Windmeijer (2012) considered less restrictive IV assumptions to deal with nonlinear and semiparametric outcome models.

When the outcome of interest is a duration measure such as mortality, the development of IV methods has been hindered by substantial challenges. Besides the need to deal with censoring, the hazard function, as a common modeling objective of many survival models, has significant limitation for causal interpretation (Hernan, 2010). Work to incorporate IVs in time-to-event analysis has relied on restrictive assumptions either on the outcome model or on the IV model. The majority of such methods impose strong outcome modeling assumptions by using parametric structural equation models (Tang and Lee, 1998; Muthen and Masyn, 2005; Chen, Hsiao, and Wang, 2011). Recent work has moved away from parametric survival models (MacKenzie et al., 2014; Tchetgen Tchetgen et al., 2015; Lin et al., 2014; Li et al., 2015). However, parametric assumptions on the IV models were still required. Robins and Tsiatis (1991) considered rank preserving structural failure time models in the clinical trial setting with noncompliance, assuming censoring time is always observed. The work of Martinussen et al. (2017) allows for flexible

semiparametric estimation with no distributional assumptions on the exposure and can incorporate covariate adjustment as long as the model for the conditional mean of the IV given covariates is correctly specified.

The focus of this paper is to develop an IV method which imposes only the core IV conditions, i.e. the independence conditions, while retaining the flexibility of semiparametric models for the time-to-event outcome. Similar to Martinussen et al. (2017), our approach can handle certain forms of dependent censoring, however we only require consistent estimation of the censoring distribution. In particular, we consider the semiparametric accelerated failure time (AFT) model which postulates a direct linear relationship between the log transformed failure time and covariates (Tsiatis, 1990; Ying, 1993). The remainder of the paper is organized as follows. In Section 2, we introduce the rank-based estimation framework for the semiparametric AFT model. Section 3 introduces a rank-based estimation framework of the semiparametric AFT model which incorporates IV estimation and derive the theoretical properties of the resulting estimator. Simulation studies are given in Section 4 to evaluate finite sample properties of our method. Finally, in Section 5 we compare the two surgical repair procedures for AAA using the Medicare enrollment data.

2. Preliminaries

Let $T_i^\dagger$ be a log-transformed duration outcome. The semiparametric AFT model postulates a direct relationship between failure time and covariates,

$$T_i^\dagger = \beta_0 X_i + e_i, \quad i = 1, \ldots, n, \tag{1}$$

where X_i is a covariate of interest and the e_i's are independent and identically distributed with an unknown distribution F_0, which does not depend on X_i. Here X_i is one-dimensional for the sake of simplicity; extensions to d-dimensional settings are straightforward. Now suppose $T_i^\dagger$ is subject to censoring by C_i. We observe $T_i \equiv T_i^\dagger \wedge C_i$ and the failure indicator $\Delta_i \equiv I\{T_i^\dagger \leqslant C_i\}$. The censoring variable C_i is commonly assumed to be independent of e_i conditioning on X_i.

A popular estimation method for the semiparametric AFT model is the rank-based method introduced by Tsiatis (1990). Let $\epsilon_i^\beta = T_i - \beta X_i$ be the observed residual for subject i, $N_i(t,\beta) = I(\epsilon_i^\beta \leqslant t, \Delta_i = 1)$ the residual counting process and $Y_i(t,\beta) = I(\epsilon_i^\beta \geqslant t)$ the residual at-risk process. Tsiatis (1990) introduced the following rank based estimating function for β_0,

$$S_n(\beta) \;=\; \frac{1}{n}\sum_{i=1}^{n}\int \rho_n(t,\beta)\left\{X_i - \eta_X(t,\beta)\right\}\, dN_i(t,\beta) \tag{2}$$

where $\rho_n(t,\beta)$ is a weight function and $\eta_X(t,\beta) \equiv D_n^{(1)}(t,\beta)\Big/ D_n^{(0)}(t,\beta)$ with

$$D_n^{(0)}(t,\beta) = \frac{1}{n}\sum_{j=1}^{n} Y_j(t,\beta) \text{ and } D_n^{(1)}(t,\beta) = \frac{1}{n}\sum_{j=1}^{n} X_j Y_j(t,\beta). \tag{3}$$

Hence $\eta_X(t,\beta)$ estimates the mean covariate value for patients in the residual risk set at time t. We have included β in our notations to emphasize dependence on β. In particular, the dependence of $N_i(t,\beta)$ and $Y_i(t,\beta)$ on β makes the estimating function $S_n(\beta)$ non-smooth, which poses computational challenges (Lin and Geyer, 1992; Tian et al., 2004; Yu and Nan, 2006).

Tsiatis (1990) showed that $S_n(\beta)$ is asymptotically linear in a $\sqrt{n}$-neighborhood of the true value β_0 under some regularity assumptions. Ying (1993) relaxed the assumptions and proved a stronger version of the same result. As $S_n(\beta)$ is not continuous, the estimator of β_0 is defined as a zero-crossing of $S_n(\beta)$. The asymptotic normality of a zero-crossing of $S_n(\beta)$ was also proved. By taking $\rho_n(t,\beta) = \omega_n(t,\beta) D_n^{(0)}(t,\beta)$, we can write

$$S_n(\beta) = \frac{1}{n}\sum_{i=1}^{n}\int \omega_n(t,\beta)\left\{D_n^{(0)}(t,\beta)X_i - D_n^{(1)}(t,\beta)\right\}\, dN_i(t,\beta).$$

When $\omega_n(t,\beta) \equiv 1$, $S_n(\beta)$ is the Gehan weighted estimating equation. We aim to extend this modeling and estimation framework to allow for IV estimation in the presence of unmeasured confounding.

3. Estimation with instrumental variables

3.1 *Instrumental Variable Assumptions and Potential Estimators*

In the presence of possible unmeasured confounding, we assume that the true model is

$$T_i^\dagger = \beta_0 X_i + U_i, \quad i = 1, \ldots, n, \tag{4}$$

The term U here is general and can represent an error term which is partly comprised of confounding factors. In contrast to (1), X and U can be correlated. When X is univariate and represents treatment, then the correlation can be due to 'differential selection' into the treatment. In the more general case of multivariate X, we assume that one component of X represents a treatment, and the effects of the other components of X on $T_i^\dagger$ are also of interest to us.

Obviously, the rank-based method for the usual AFT model does not work any more due to this correlation. Now assume that there is an IV Z that satisfies the following well-known core conditions (Clarke and Windmeijer, 2012; Didelez and Sheehan, 2007):

(1) Independence: $Z \perp\!\!\!\perp U$

(2) Exclusion Restriction: $T^\dagger \perp\!\!\!\perp Z \mid X, U$

(3) Relatedness: $X \not\perp\!\!\!\perp Z$, for all components of X

The relationship among the variables can also be depicted in Figure 1. These three assumptions are critical for consistent estimation with IVs. The first of these assumptions assures that the instrument is uncontaminated by unmeasured confounding. The second assumption requires that the instrument does not affect the survival time except through its effect on X. If this is violated, then there is no way to determine how a change in the instrument affects the outcome through its change on the treatment. Clearly it is necessary for there to be a relationship between the instrument and the treatment for any hope of consistent IV estimation.

[Figure 1 about here.]

Instead of the observable quantity based formulation, counterfactuals can be used to provide an

equivalent description of our model. Such equivalent description has appeared in recent literature. In particular, Wang and Tchetgen Tchetgen (2016) gave such a counterfactual formulation in their Section 2 for causal inference with IVs. Specifically, for the case when X is univariate and represents a dichotomous treatment, let $T^{\dagger}(x, z)$ be the counterfactual log-transformed survival time corresponding to treatment level $X = x$ and IV level $Z = z$. Then the independence assumption becomes $Z \perp\!\!\!\perp (X(z), T^{\dagger}(x, z))$ and the exclusion restriction assumption becomes $T^{\dagger}(x, z) = T^{\dagger}(x)$ for any z. The underlying structural model takes a similar form as Equation (4). However the underlying effect of the unmeasured confounder U on $T^{\dagger}(x)$ needs to be the same for different treatment levels. Otherwise only a local average treatment effect (LATE) is identifiable without further assumptions (Wang and Tchetgen Tchetgen, 2016; Tan, 2006). Whereas the counterfactual framework makes the underlying assumptions more transparent, we will use our observable quantity based formulation for the remainder of the article to make it accessible to a wider audience.

An IV can be viewed as an experimental handle. When the IV is changed, the treatment status may be changed, and only through this change the outcome is impacted. This viewpoint is a key motivation of our proposed estimator and provides insight into its consistency. Another view of IVs is that the instrument is an imperfect measurement of the treatment status, but unlike the treatment status, it is not affected by unmeasured confounding. This view of IVs leads to an imputation-based estimator such as a two-stage least squares type estimator.

One possible IV estimator under the AFT model is a two-stage least squares approach. Such an approach would first posit a linear relationship between X and Z as $X = \gamma Z + \nu$, where $E(\nu) = 0$ and var$(\nu) < \infty$. Using this model, imputed values $\hat{X}$ of the treatment X would be computed as $\hat{X} = Z(Z'Z)^{-1}Z'X$ and would replace X everywhere in (2). This may rely on the validity of the posited relationship between X and Z and hence if this relationship does not hold, the resulting estimator is potentially inconsistent. This assertion is demonstrated in our simulation studies.

Let $\epsilon_i^\beta = T_i - \beta X_i$ be the observed residual for subject i, $N_i(t,\beta) = I(\epsilon_i^\beta \leqslant t, \Delta_i = 1)$ the residual counting process and $Y_i(t,\beta) = I(\epsilon_i^\beta \geqslant t)$ the residual at-risk process. Here, Δ is as defined in Section 2. In light of the view of IVs as experimental handles, the development of an estimator for β_0 should capture how changes in Z affect X and how this change affects $T^\dagger$. To estimate β_0, it would then seem intuitive to use

$$\tilde{\Psi}_n(\beta) = \frac{1}{n}\sum_{i=1}^{n} \int \rho_n(t,\beta)\left\{Z_i - \eta_Z(t,\beta)\right\}\, \mathrm{d}N_i(t,\beta) \tag{5}$$

where $\rho_n(t,\beta)$ is a weight function and $\eta_Z(t,\beta) \equiv \tilde{D}_n^{(1)}(t,\beta)\Big/ D_n^{(0)}(t,\beta)$ with

$$\tilde{D}_n^{(1)}(t,\beta) = \frac{1}{n}\sum_{j=1}^{n} Z_j Y_j(t,\beta) = \frac{1}{n}\sum_{j=1}^{n} Z_j I(\epsilon_j^\beta \geqslant t) = \frac{1}{n}\sum_{j=1}^{n} Z_j I(T_j - \beta X_j \geqslant t).$$

and $D_n^{(0)}(t,\beta)$ defined in (3).

Hence $\eta_Z(t,\beta)$ estimates the mean covariate value of the IV Z for patients in the residual (defined using X) risk set at time t. In other words, the time scale of (5) is constructed using X, but the covariate comparison is done using Z. However, even under the strong assumption that C is independent of $(Z, X, T^\dagger)$, this formulation induces non-ignorability of the censoring random variable C_i on the residual time scale and hence is not consistent for β_0. To remedy this issue, we next consider unbiased estimation of β_0 in (4) using the IV.

Note that the estimating equation of Robins and Tsiatis (1991), with extensions thereof proposed by Bijwaard (2009), is closely related to (5). Considering covariates without time-dependence and a dichotomous X, Robins and Tsiatis (1991) replace $Y_i(t,\beta)$ and $N_i(t,\beta)$ with $Y_i^*(t,\beta) = I(v_i^\beta \geqslant t)$ and $N_i^*(t,\beta) = I(\epsilon_i^\beta \leqslant t, \Delta_i^* = 1)$, respectively, where $v_i^\beta = \min(T_i \exp(\beta X_i), C_i^*)$, the artificial censoring variable $C_i^* = C_i I(\beta \geqslant 0) + C_i \exp(\beta) I(\beta < 0)$, and $\Delta_i^* = I(T_i \exp(\beta X_i) < C_i^*)$. The population parameter of interest $\widetilde{\beta}_0$ under the model of Robins and Tsiatis (1991) represents $-\beta_0$ of (1). The induced non-ignorability of censoring is handled via the artificial censoring variable C_i^*, which is no longer non-ignorable. Note that this approach requires C_i observed for all subjects.

3.2 *Proposed Estimator*

To mitigate the induced non-ignorability of censoring, we propose an estimator $\hat{\beta}_n$ of β_0 as the zero crossing of the following estimating equation

$$\Psi_n(\beta) = \frac{1}{n}\sum_{i=1}^{n}\int \rho_n(t,\beta)\left\{Z_i - \hat{\eta}_{\hat{G},n}(t,\beta)\right\}\frac{\mathrm{d}N_i(t,\beta)}{\hat{G}(t+\beta X_i|X_i,Z_i)} \tag{6}$$

where $\rho_n(t,\beta)$ is a weight function, $\hat{G}$ is any uniformly consistent estimator of the survival function G_0 for the censoring random variable C, which may depend on the covariate X and instrument Z but not U, and

$$\hat{\eta}_{\hat{G},n}(t,\beta) \equiv \frac{D^{(1)}_{\hat{G},n}(t,\beta)}{D^{(0)}_{\hat{G},n}(t,\beta)} \tag{7}$$

with

$$D^{(1)}_{\hat{G},n}(t,\beta) = \frac{1}{n}\sum_{j=1}^{n}\frac{Z_jY_j(t,\beta)}{\hat{G}(t+\beta X_j|X_j,Z_j)} \quad \text{and} \quad D^{(0)}_{\hat{G},n}(t,\beta) = \frac{1}{n}\sum_{j=1}^{n}\frac{Y_j(t,\beta)}{\hat{G}(t+\beta X_j|X_j,Z_j)}. \tag{8}$$

Our proposed estimating equation is based on quantities weighted according to the inverse probability of censoring, which accounts for noninformative censoring by creating a pseudopopulation that would have been observed had there been no censoring. Here we explicitly put $\hat{G}$ in the notation as it is generally unknown and needs to be estimated. We assume that $C \perp\!\!\!\perp T^{\dagger}|(X,Z)$. A common model for G when censoring depends on covariates is Cox's proportional hazards model, $\lambda(c|X,Z) = \lambda_0(c)\exp(X\alpha_X + Z\alpha_Z)$, under which $\hat{G}(t|X,Z) = [\exp(-\int_0^t \hat{\lambda}_0(u)\mathrm{d}\,u)]^{\exp(X\hat{\alpha}_X + Z\hat{\alpha}_Z)}$. If C is independent of $(Z,X,T^{\dagger})$, then we can take $\hat{G}$ as the estimated survival function of C based on the Kaplan-Meier estimator. Although we define our estimator as the zero crossing of (6), an approximate root $\hat{\beta}_n$ satisfying $\Psi_n(\hat{\beta}_n) = o_{p*}(n^{-1/2})$ satisfies the conditions required for our asymptotic derivations in Section 3.3.

Dividing by zero will never occur in (8) or in (6). The denominator, $\hat{G}(t+\beta X_j|X_j,Z_j)$, can only be zero if $t+\beta X_j > \tau^*$, where τ^* represents the longest follow-up time. However, if $t+\beta X_j > \tau^*$, then the term in the numerator $Y_j(t) = I(T_j \geqslant t+\beta X_j) = 0$. Then the terms in (8) can be defined

as 0 without issue. Similarly, dividing by zero does not occur in (6) as $t+\beta X_i$ by definition cannot be greater than τ^*.

Before we study asymptotic properties of the solution $\hat{\beta}_n$ of $\Psi_n(\beta) = 0$, we justify the unbiasedness of the population level version of (6) for 0 when β is set to β_0. Thus this should imply Fisher consistency of $\hat{\beta}_n$ for estimating β_0 under some regularity conditions such as uniqueness of the solution.

The term $D^{(1)}_{\hat{G},n}(t,\beta_0)$ is an estimator of its population level counterpart

$$\begin{aligned} d^{(1)}_{G_0}(t,\beta_0) &= E\big\{ZI(U \geqslant t)I(C \geqslant t+\beta_0 X)G_0(t+\beta_0 X|X,Z)^{-1}\big\} \text{ where } U = T^\dagger - \beta_0 X \\ &= E\big\{ZI(U \geqslant t)G_0(t+\beta_0 X|X,Z)^{-1}E(I(C-\beta_0 X \geqslant t) \mid X,Z,U)\big\} \\ &= E\big\{ZI(U \geqslant t)G_0(t+\beta_0 X|X,Z)^{-1}G_0(t+\beta_0 X|X,Z)\big\} \\ &= E\big\{ZI(U \geqslant t)\big\} = E(Z)P(U \geqslant t) \end{aligned} \tag{9}$$

Similarly, $D^{(0)}_{\hat{G},n}(t,\beta_0)$ is an estimator of its population level counterpart

$$d^{(0)}_{G_0}(t,\beta_0) = E\big\{I(U \geqslant t)I(C \geqslant t+\beta_0 X)G_0(t+\beta_0 X|X,Z)^{-1}\big\} = P(U \geqslant t) \tag{10}$$

Therefore $D^{(1)}_{\hat{G},n}(t,\beta_0)/D^{(0)}_{\hat{G},n}(t,\beta_0)$ converges to $E(Z)$ for any t. Assuming that $\rho_n(t,\beta_0)$ converges to a function of t, say $\rho_0(t)$, we then show that $\Psi_n(\beta_0)$ basically acts as an estimator of 0.

$$\begin{aligned} \Psi_n(\beta_0) &\to E\left[\rho_0(U)\left\{Z - \frac{d^{(1)}_{G_0}(U)}{d^{(0)}_{G_0}(U)}\right\}\Delta G_0(T^\dagger|X,Z)^{-1}\right] \\ &= E\left[E\left(\rho_0(U)\left\{Z-E(Z)\right\}G_0(T^\dagger|X,Z)^{-1}\Delta\middle| T^\dagger,X,Z,U\right)\right] \\ &= E\left[\rho_0(U)\left\{Z-E(Z)\right\}G_0(T^\dagger|X,Z)^{-1}E\left(I(C>T^\dagger)\middle| T^\dagger,X,Z,U\right)\right] \\ &= E\left[\rho_0(U)\left\{Z-E(Z)\right\}G_0(T^\dagger|X,Z)^{-1}G_0(T^\dagger|X,Z)\right] \\ &= E\left[\rho_0(U)\{Z-E(Z)\}\right] \\ &= 0, \end{aligned}$$

where the third equality holds by assumption (C.5) below and the last equality holds due to the independence of U and Z.

3.3 *Asymptotic properties*

We utilize the notation of van der Vaart and Wellner (1996); Kosorok (2008) for empirical processes. For a given function f of a random variable X, we denote $Pf = \int f(x)\ \mathrm{d}P(x), \mathbb{P}_n f = n^{-1}\sum_{i=1}^{n} f(X_i)$, and $\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P)f$. Additionally, for a sequence of random variables $\{X_n\}$ and a sequence of constants a_n, $n = 1, 2, \ldots$, we say $X_n = o_{p^*}(a_n)$ if X_n/a_n converges to 0 in outer probability P^* corresponding to P (van der Vaart and Wellner, 1996; Kosorok, 2008). Similarly, we say $X_n = O_{p*}(a_n)$ if X_n/a_n is stochastically bounded in outer probability, i.e. for any $\epsilon > 0$, there exists $M > 0$ such that $P^*(|X_n/a_n| > M) < \epsilon,\ \forall n$.

In this notation, for any G, we can write $D^{(1)}_{G,n}(t,\beta) = \mathbb{P}_n\{ZI(T-\beta X \geqslant t)G^{-1}(t+\beta X|X,Z)\}$ and $D^{(0)}_{G,n}(t,\beta) = \mathbb{P}_n\{I(T-\beta X \geqslant t)G^{-1}(t+\beta X|X,Z)\}$. Their corresponding limits are $d^{(1)}_G(t,\beta) \equiv P\{ZI(T-\beta X \geqslant t)G^{-1}(t+\beta X|X,Z)\}$ and $d^{(0)}_G(t,\beta) \equiv P\{I(T-\beta X \geqslant t)G^{-1}(t+\beta X|X,Z)\}$. Then $\eta_0(t,\beta) \equiv d^{(1)}_{G_0}(t,\beta)/d^{(0)}_{G_0}(t,\beta)$ is the limit of $\hat{\eta}_{\hat{G},n}(t,\beta)$. Let τ^* represent the longest follow-up time and τ_* the initial time of the study. The supernorm $||\cdot||$ is therefore defined over the interval $[\tau_*, \tau^*]$. We assume the following conditions for Theorem 1:

(C.1) $\mathcal{B}$, the parameter space of β, is compact;

(C.2) $\beta_0 \in \mathcal{B}$ is the unique solution of $\Psi_0(\beta) = 0$ where $\Psi_0(\beta)$ is the limit of $\Psi_n(\beta)$;

(C.3) $G_0(t|X,Z)$ and $\hat{G}(t|X,Z)$ are bounded away from zero when $\tau_* \leqslant t \leqslant \tau^*$;

(C.4) X has bounded support;

(C.5) $C \perp\!\!\!\perp T^\dagger|(X,Z)$.

THEOREM 1: *Assume conditions (C.1)-(C.5) are true, that both $\rho_n(t,\beta)$ in $\Psi_n(\beta)$ defined in (6) and $\hat{G}(t)$ belong to Glivenko-Cantelli classes, and that $\|\hat{G} - G_0\| = o_{p^*}(1)$. Then an approximate root $\hat{\beta}_n$ satisfying $\Psi_n(\hat{\beta}_n) = o_{p*}(1)$ is consistent. In particular when $\hat{\rho}_{\hat{G},n}(t,\beta) = D^{(0)}_{\hat{G},n}(t,\beta)$ for the Gehan type of weight, the resulting estimator is consistent.*

EXAMPLE 1: Assume C follows the Cox's proportional hazards model $\lambda(c|X,Z) = \lambda_0(c)\exp(X\alpha_X +$

$Z\alpha_Z)$. Then take $\hat{G}(t|X,Z) = [\exp(-\int_0^t \hat{\lambda}_0(u)\mathrm{d}\ u)]^{\exp(X\hat{\alpha}_X + Z\hat{\alpha}_Z)}$. From (Kosorok, 2008, pages 54-59), $\{\hat{G}(t|X,Z)\}$ admits a Donsker class (and thus Glivenko-Cantelli class) and hence the conditions for Theorem 1 hold.

EXAMPLE 2: Make the further assumption that C is independent of $(Z, X, T^\dagger)$ and let $\hat{G}$ be the Kaplan-Meier estimator of G_0. From (Kosorok, 2008, pages 24-28), $\{\hat{G}(t|X,Z)\} = \{\hat{G}(t)\}$ admits a Donsker class and hence the conditions for Theorem 1 hold.

Now we consider the asymptotic distribution of $\hat{\beta}_n$. We chose to work with $\gamma(\cdot) \equiv G(\cdot|X,Z)^{-1}$ instead of $G(\cdot)$ for asymptotic investigation due to both the simplicity of notation and theoretical development. For clarity of presentation, in the following theorem, for generic η, ρ, G, we define

$$\Psi(\beta, \eta, \rho, \gamma) \equiv P[\gamma(T)\rho(\epsilon_\beta, \beta)\{Z - \eta(\epsilon_\beta, \beta)\}\Delta].$$

The corresponding empirical version is $\Psi_n(\beta, \eta, \rho, \gamma) = \mathbb{P}_n[\gamma(T)\rho(\epsilon_\beta, \beta)\{Z - \eta(\epsilon_\beta, \beta)\}\Delta]$.

For simplicity of presentation, we focus on $\hat{\rho}_{\hat{G},n}(t,\beta) = D^{(0)}_{\hat{G},n}(t,\beta)$ which converges to $\rho_0(t,\beta) \equiv P(T - \beta X \geqslant t)$. Note that $\eta_0(t,\beta) = P(Z \mid T - \beta X \geqslant t)$. Denote $\gamma_0 = G_0^{-1}$ where G_0 is the survival function for C. We introduce the following conditions for Theorem 2:

(C.6) $\Psi(\beta, \eta_0(\cdot,\beta), \rho_0(\cdot,\beta), \gamma_0(\cdot))$ is differentiable with respect to $\beta \in \mathcal{B}_0$ where $\mathcal{B}_0 \subset \mathcal{B}$ is a neighborhood of β_0;

(C.7) The derivative $\dot{\Psi}_\beta(\beta) \equiv \dot{\Psi}_\beta(\beta, \eta_0(\cdot,\beta), \rho_0(\cdot,\beta), \gamma_0(\cdot))$ is bounded and continuous. Furthermore, its evaluation at β_0, $\dot{\Psi}_\beta(\beta_0) \equiv \dot{\Psi}_\beta(\beta_0, \eta_0(\cdot,\beta_0), \rho_0(\cdot,\beta_0), \gamma_0(\cdot))$, is nonsingular.

Let $\mathcal{T}$, $\mathcal{Z}$, and $\mathcal{X}$ denote the sample spaces of the random variables T, Z, and X respectively. Then condition (C.6) is also implied by the following condition:

(C.6*) $\rho_0(\epsilon_\beta, \beta)$ and $\eta_0(\epsilon_\beta, \beta)$ are differentiable in β with continuous derivatives $\dot{\rho}_{0\beta}$ and $\dot{\eta}_{0\beta}$ respectively, which are uniformly bounded and continuous in $\mathcal{B}_0 \times \mathcal{T} \times \mathcal{Z} \times \mathcal{X}$.

Condition (C.6*) implies that $\Psi(\beta, \eta_0(\cdot,\beta), \rho_0(\cdot,\beta), \gamma_0(\cdot))$ is differentiable in β with bounded continuous derivative $\dot{\Psi}_\beta(\beta, \eta_0(\cdot,\beta), \rho_0(\cdot,\beta), \gamma_0(\cdot))$ in $\mathcal{B}_0$.

We must further define the derivative map of $\Psi(\beta_0, \eta_0, \rho_0, \gamma)$ with respect to the function $\gamma(\cdot)$ as it is estimated by $\hat{\gamma}_n = 1/\hat{G}$. For this, we use the more general differentiability concept for normed spaces (van der Vaart, 1998, Chapter 20). In particular, in the Supplementary Material, we prove that $\Psi(\beta, \eta_0(\cdot, \beta), \rho_0(\cdot, \beta), \gamma(\cdot))$ is Hadamard-differentiable with respect to $\gamma(\cdot)$. Both η_0 and ρ_0 are also functions of γ, however for simplicity this is not indicated in the notation. Denote the corresponding derivative map as

$$\dot{\Psi}_\gamma(\gamma) \equiv \dot{\Psi}_\gamma(\beta_0, \eta_0(\cdot, \beta_0), \rho_0(\cdot, \beta_0), \gamma(\cdot)).$$

Furthermore, let $\widetilde{\beta}_n$ be an approximate root satisfying

$$\Psi_n(\widetilde{\beta}_n, \hat{\eta}_{G_0,n}(\cdot, \widetilde{\beta}_{G_0,n}), \hat{\rho}_{G_0,n}(\cdot, \widetilde{\beta}_n), \gamma_0(\cdot)) = o_{p^*}(n^{-1/2}).$$

Thus, $\widetilde{\beta}_n$ is our proposed estimator when the true censoring distribution is known.

THEOREM 2: *Let $\hat{\beta}_n$ be an approximate root satisfying*

$$\Psi_n(\hat{\beta}_n, \hat{\eta}_{\hat{G},n}(\cdot, \hat{\beta}_{\hat{G},n}), \hat{\rho}_{\hat{G},n}(\cdot, \hat{\beta}_n), \hat{\gamma}_n(\cdot)) = o_{p^*}(n^{-1/2}).$$

Then under the conditions (C.1)-(C.7), and the assumptions of Theorem 1 but with $\{\hat{G}(t)\}$ belonging to Donsker classes, and that $\sqrt{n}\|\hat{G} - G_0\|$ converges to a tight, mean zero Gaussian process, we have $\sqrt{n}(\hat{\beta}_n - \beta_0)$ is asymptotically normal with the asymptotic representation

$$\sqrt{n}(\hat{\beta}_n - \beta_0) = \sqrt{n}(\widetilde{\beta}_n - \beta_0) - \{\dot{\Psi}_\beta(\beta_0)\}^{-1}\dot{\Psi}_\gamma(\sqrt{n}(\hat{\gamma}_n - \gamma_0)) + o_{p^*}(1)$$

Both Theorems 1 and 2 are proven in the Supplementary Material. The asymptotic representation of $\sqrt{n}(\widetilde{\beta}_n - \beta_0)$ is derived in Lemma 2 of the Supplementary Material. Note that both Examples 1 and 2 are covered under Theorem 2. For variance estimation, we use a modification of the computationally efficient approach of Zeng and Lin (2008), outlined in the Supplementary Material.

4. Simulations

We investigate the performance of our method under a variety of simulation settings intended to replicate the presence of unmeasured confounding in observational studies. The aim of the simula-

tions is to evaluate finite sample estimation properties under varying degrees of unmeasured confounding and strengths of the IV. We first generate independent standard normal random variables $\epsilon_{T^\dagger}, \epsilon_X, \epsilon_C, U$, and Z. Then we generate log-normal survival times $T^\dagger = \exp\{\beta_0 X + \beta_U U + \epsilon_{T^\dagger}\}$. The covariate is generated as $X = \alpha_Z \exp\{Z\} + \alpha_U U + \epsilon_X$. The censoring times are log-normally-distributed and either independent of covariates $C = \exp\{\alpha + \epsilon_C\}$ or dependent on covariates $C = \exp\{0.25X - 0.25Z + \epsilon_C\}$. Here α is chosen to be the mean of $\log T^\dagger$ to keep the censoring times on a similar scale as the survival times. The simulation parameters β_U and α_U were varied from 0.25 to 0.75 and α_Z was varied from 0.25 to 1. The true effect of interest, β_0 was set to 1. When the censoring distribution is independent, the Kaplan-Meier estimator is used to estimate $\hat{G}$ and when the censoring distribution is dependent, a Cox proportional hazard model was used for $\hat{G}$. The variable U in this simulation acts as unmeasured confounding. The strengths of the effect of the unmeasured confounding on the survival time $T^\dagger$ and the treatment X are specified by β_U and α_U respectively. The strength of the IV is characterized by α_Z. The effect of the IV Z on X is nonlinear to demonstrate that our estimator does not require correct specification of the structural relationship between Z and X. The censoring rate under the independent censoring simulation settings is approximately 0.48. The censoring rate under the dependent censoring rate is approximately 0.62, which is similar to the censoring rate in the motivating Medicare enrollment dataset.

We evaluate our proposed estimator as well as other estimators. All estimators are based on rank-based estimation utilizing Gehan-type weights. To evaluate the impact of the unmeasured confounding we use as a baseline the standard rank-based estimator for the semiparametric AFT model in Eq. (2). In addition, we evaluate the impact of the induced dependence of the censoring random variable on the residual scale by using the estimator in Eq. (5). To assess the impact of incorrectly specified models in a structural equation modeling based approach, we evaluate a two-stage estimator. In the two stage estimator, X is regressed on Z in a linear model with no nonlinear

terms of Z. Predictions $\hat{X}$ of X using this incorrect model are then generated and X is replaced with $\hat{X}$ in the standard rank-based estimator given in Eq. (2) and the solution of this equation gives a structural equation modeling type estimator for β_0. The two-stage estimator is left unspecified to demonstrate its reliance on the truth of the posited structural equation. Finally, we evaluate the performance of our proposed estimator in Eq. (6). For this method we use the Kaplan-Meier estimator of the censoring distribution when the censoring random variable is independent and we use Cox's proportional hazards model when it is dependent. The results for estimation bias are in Figures 2 and 3 and coverage results are displayed in Figures 4 and 5. Results under Cox's proportional hazards model are presented in the Supplementary Material. Simulation results for scenarios with depending censoring are similar to those with independent censoring and are thus presented in the Supplementary Material. The method denoted "AFT" is the standard semiparametric AFT model with no adjustment for unmeasured confounding, the method denoted "AFT-2SLS" is the semiparametric AFT method with X replaced with $\hat{X}$ from a first stage linear regression model, "AFT-IV" is the semiparametric AFT model with IV adjustment of Eq. (5) that does not account for induced censoring dependence, and finally "AFT-IV-IPCW" is our proposed estimator. All approaches use the Gehan weighting function. The proposed estimator has consistently less empirical bias than other estimators under all settings. All estimators tend to have more bias for larger values of β_U and α_U. The proposed estimator has reduced bias and variance for the larger sample size. The AFT and AFT-2SLS estimators have differing biases as α_Z is changed, however this is an artifact of how the data were simulated. Unsurprisingly, the empirical coverage of all methods other than the proposed method is significantly below the nominal level. The empirical coverage of the proposed method is approximately at the nominal level in most scenarios, however in some settings (in particular when $\alpha_U = 0.75$) its empirical coverage is slightly above the nominal level. Further simulations investigating the robustness of AFT-IV-IPCW to censoring distribution

misspecification and results regarding the root mean squared error of estimates are presented in the Supplementary Material.

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

[Figure 5 about here.]

5. Analysis of Medicare enrollment data

Abdominal aortic aneurysm (AAA) is a life-threatening condition characterized by a pronounced enlargement of the abdominal aorta, increasing the risk for aortal rupture. Aortal rupture is associated with a severely high rate of mortality, resulting in death in approximately 90% of cases (Assar and Zarins, 2009). One treatment option to mitigate the risk of rupture is open surgical repair, however it is associated with higher risk of complication. Endovascular repair (EVAR) is a less invasive repair procedure and is associated with lower short-term mortality. EVAR has found widespread use and now accounts for a large portion of elective repairs, yet there are concerns that EVAR is less effective in the long-term, leading to more frequent reinterventions. On the other hand, evidence from RCTs comparing EVAR with open repair suggest that while there may be a short term mortality reduction for EVAR (Prinssen et al., 2004), there may be little difference in the long-term (Lederle et al., 2012). However these trials are underpowered, one having only 10 AAA-related deaths, and lack long-term follow-up. With the excessive costs involved in conducting RCTs, it is unlikely that it will be possible to conduct a trial large enough and for long enough to definitively answer questions over the comparative advantage or not of EVAR versus open surgical repair over long-term follow-up. However, the concern that the mortality benefits of EVAR may not be durable in the long-term is an important population health concern.

As sufficient data from RCTs are not available, comparisons between the two repair procedures

must be carried out using observational data. One such dataset is taken from Medicare enrollment. The current analysis of this data (Schermerhorn et al., 2008; Edwards et al., 2014) hinges on propensity score matching and thus does not control for bias due to unmeasured confounding. Thus an analysis which accounts for unmeasured confounding could result in a better understanding of AAA repair and specifically whether the recommended surgical procedure varies with the length of follow-up being considered. Such comparative effectiveness predicaments occur in a plethora of domains and in each case clinical researchers, physicians, and surgeons are anxious to learn the best approach.

To illustrate our method, we apply it to analyze a data set of Medicare beneficiaries who received surgical AAA repair. The surgeries were performed from 2001-2008 and patients were followed up until 2009. Patients aged 65 and older were included and those enrolled in Health Maintenance Organizations were excluded from the study. The primary outcome is all-cause survival. The IV used is the proportion of AAA repairs that were EVAR for each hospital over the previous year. There is no conclusive method to demonstrate that the IV independence assumption holds using observable data; we can only justify its use using external arguments. A crucial point is that we include total AAA volume per hospital as a predictor, hence the IV only fails if institutional AAA experience is procedure specific in the sense that the proportion of past EVAR cases is either correlated with an unmeasured confounder or is directly related to the survival time to death. In our analysis we adjust for demographic as well as medical information. Medical information includes instances of previous AAA, myocardial infarction over the past 6 months, valvular heart disease, congestive heart failure, hypertension, and diabetes mellitus among others. By adjusting for these observed predictors in our survival time and EVAR selection models, we hope to block remaining sources of unmeasured confounding. We focus on the subset of patients who experienced a ruptured AAA. Patients with ruptured AAA experience high mortality and thus it is of interest to understand the effects of the two surgical repair procedures in such cases. The data consists of 2853 patients

with ruptured AAA with a censoring rate of approximately 66%. Approximately 77% of surgeries were EVAR. The median follow-up time is approximately 1598 days and the longest follow-up time is 3283 days. The median failure and censoring times are 1253 and 1759 days, respectively.

Without adjusting for unmeasured confounding, analysis using the semiparametric AFT model suggests that there is little difference in long-term survival between those who received EVAR and those who received open repair. However analysis using our method results in a considerably different estimate of the effect of EVAR. While the difference in survival time between the two procedures is not significant, the sign of the effect of EVAR changed from that under the standard AFT model to suggest slightly higher risk under EVAR (see Table 1 for the estimates of the effect of EVAR on $\log T^{\dagger}$ under the various approaches).

[Table 1 about here.]

An obvious counter to the utility of the new AFT-IV-IPCW method is that the same conclusion is reached under all analyses and its estimate is indistinguishable from that of the other IV procedures. However, the lack of significance of the IV procedures is in-part due to the wider confidence intervals accompanying them suggesting that these effects may have been significant if the data set had been larger, and the similarity among the IV estimates reflects a tendency for estimates of different procedures to converge the closer the true effect is to 0 or in situations when the true amount of unmeasured confounding is relatively small.

An issue raised by the empirical analysis is whether the treatment effect of EVAR over open surgical repair is homogeneous. An extension of the current methodology is to allow for treatment effect heterogeneity over follow-up time, initially based on a specified change-point and ultimately allowing for unspecified change-points. This remains an exciting avenue of further work.

6. Discussion

The focus of this paper is to develop an estimator which incorporates IVs in the semiparametric AFT model to mitigate bias from unmeasured confounding for studies with time-to-event outcomes. The resulting estimator relies only on the core IV conditions and provides consistent estimation of the effect of a treatment or intervention in the presence of unmeasured confounding. Furthermore, the asymptotic properties of the estimator are established. The proposed estimator requires estimation of the distribution of the censoring random variable, which seems necessary as the core IV conditions are known to be insufficient to identify the causal parameters under general outcome models. However, the proposed method is robust to misspecification of the censoring distribution in particular when a strong IV is used. Considering more flexible models than (4), our approach can be extended to a more general class of rank preserving structural failure time models (Robins and Tsiatis, 1991) as outlined in Section 2 of the Supplementary Material. A complication of our estimator is the non-monotonicity of the estimating equation even when using Gehan-weights. It would be worthwhile to develop a consistent estimator based on monotone estimating equations in order to ease the computational burden and potentially improve efficiency.

7. Supplementary Materials

Web Appendices and Figures referenced in Sections 3, 4, and 6 in addition to an `R` implementation of the proposed methodology are available with this paper at the Biometrics website on Wiley Online Library.

Acknowledgments

The authors would like to thank the anonymous referees for their helpful comments. Dr. O'Malley's time on this work was in part supported by the Patient-Centered Outcomes Research Institute (PCORI) Award ME-1503-28261. The time of Drs. Huling and Yu on this work was in part supported by the Patient-Centered Outcomes Research Institute (PCORI) Award ME-1409-21219.

All statements in this paper, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee.

REFERENCES

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

Assar, A. and Zarins, C. (2009). Ruptured abdominal aortic aneurysm: a surgical emergency with many clinical presentations. *Postgraduate Medical Journal* **85,** 268–273.

Bijwaard, G. E. (2009). Instrumental variable estimation for duration data. In Henriette, E., Hans-Peter, K., and Alexia, F., editors, *Causal Analysis in Population Studies*, The Springer Series on Demographic Methods and Population Analysis, chapter 6, pages 111–148. Springer.

Chen, S., Hsiao, C., and Wang, L. (2011). Measurement errors and censored structural latent variables models. *Econometric Theory* **28,** 696–710.

Clarke, P. S. and Windmeijer, F. (2012). Instrumental variable estimators for binary outcomes. *Journal of the American Statistical Association* **107,** 1638–1652.

D'Agostino, R. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine* **17,** 2265–2281.

Didelez, V. and Sheehan, N. (2007). Mendelian randomization as an instrumental variable approach to causal inference. *Statistical Methods in Medical Research* **16,** 309–330. PMID: 17715159.

Edwards, S., Schermerhorn, M., O'Malley, A., Bensley, R., Hurks, R., Cotterill, P., and Landon, B. (2014). Comparative effectiveness of endovascular versus open repair of ruptured abdominal aortic aneurysm in the medicare population. *Journal of Vascular Surgery* **59,** 575–582.

Hernan, M. (2010). The hazards of hazard ratios. *Epidemiology* **21,** 13–15.

Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer Series in Statistics. Springer New York.

Lederle, F., Freischlag, J., Kyriakides, T., Matsumura, J., Padberg, F., Kohler, T., Kougias, P., Jean-Claude, J., Cikrit, D., and Swanson, K. (2012). Long-term comparison of endovascular and open repair of abdominal aortic aneurysm. *New England Journal of Medicine* **367,** 1988–1997.

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

Lin, D. Y. and Geyer, C. J. (1992). Computational methods for semiparametric linear regression with censored data. *Journal of Computational and Graphical Statistics* **1,** 77–90.

Lin, H., Li, Y., Jiang, L., and Li, G. (2014). A semiparametric linear transformation model to estimate causal effects for survival data. *Canadian Journal of Statistics* **42,** 18–35.

MacKenzie, T., Tosteson, T., Morden, N., Stukel, T., and O'Malley, A. (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., Vansteelandt, S., Tchetgen Tchetgen, E. J., and Zucker, D. M. (2017). Instrumental variables estimation of exposure effects on a time-to-event endpoint using structural cumulative survival models. *Biometrics* **73,** 1140–1149.

Muthen, B. and Masyn, L. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics* **30,** 27–58.

O'Malley, A., Cotterill, P., Schermerhorn, M., and Landon, B. (2011). Improving observational study estimates of treatment effects using joint modeling of selection effects and outcomes: the case of aaa repair. *Medical care* **49,** 1126.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika* **82,** 669–688.

Prinssen, M., Verhoevenet, E., Buth, J., Cuypers, P., van Sambeek, M., Balm, R., Buskens, E., Grobbee, D., and Blankensteijn, J. (2004). A randomized trial comparing conventional and

endovascular repair of abdominal aortic aneurysms. *New England Journal of Medicine* **351,** 1607–1618.

Robins, J. M. and 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. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70,** 41–55.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *Annals of Applied Statistics* **2,** 808–840.

Schermerhorn, M., O'Malley, A., Jhaveri, A., Cotterill, P., Pomposelli, F., and Landon, B. (2008). Endovascular vs. open repair of abdominal aortic aneurysms in the medicare population. *New England Journal of Medicine* **358,** 464–474.

Tan, Z. (2006). Regression and weighting methods for causal inference using instrumental variables. *Journal of the American Statistical Association* **101,** 1607–1618.

Tang, M. and Lee, S. (1998). Analysis of structural equation models with censored or truncated data via em algorithm. *Computational Statistics and Data Analysis* **27,** 33–46.

Tchetgen Tchetgen, E., Walter, S., Vansteelandt, S., Martinussen, T., and Glymour, M. (2015). Instrumental variable estim ation in a survival context. *Epidemiology* **26,** 402–10.

Terza, J., Basu, A., and Rathouz, P. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of Health Economics* **27,** 531–543.

Tian, L., Liu, J., Zhao, Y., and Wei, L. J. (2004). Statistical inference based on non-smooth estimating functions. *Biometrika* **91,** 943–954.

Tsiatis, A. (1990). Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics* **18,** pp. 354–372.

van der Vaart, A. (1998). *Asymptotic statistics*. Cambridge series in statistical and probabilistic

mathematics. Cambridge University Press, Cambridge (UK), New York (N.Y.). Autre tirage : 2000 (èdition brochèe), 2005, 2006, 2007.

van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics (Springer Series in Statistics)*. Springer.

Wang, L. and Tchetgen Tchetgen, E. (2016). Bounded, efficient and multiply robust estimation of average treatment effects using instrumental variables. *ArXiv e-prints* .

Ying, Z. (1993). A large sample study of rank estimation for censored regression data. *Annals of Statistics* **21,** 76–99.

Yu, M. and Nan, B. (2006). A hybrid newton-type method for censored survival data using double weights in linear models. *Lifetime Data Analysis* **12,** 345–364.

Zeng, D. and Lin, D. (2008). Efficient resampling methods for nonsmooth estimating functions. *Biostatistics* **9.2,** 355–363.

Received October 2100. *Revised February* 2101. *Accepted xxxx* xxxx.

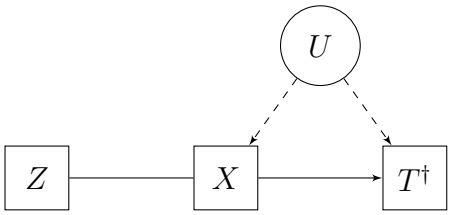

Figure 1. Diagram of core IV assumptions. A line with no arrow indicates dependence and a line with an arrow indicates a causal relationship in a specific direction.

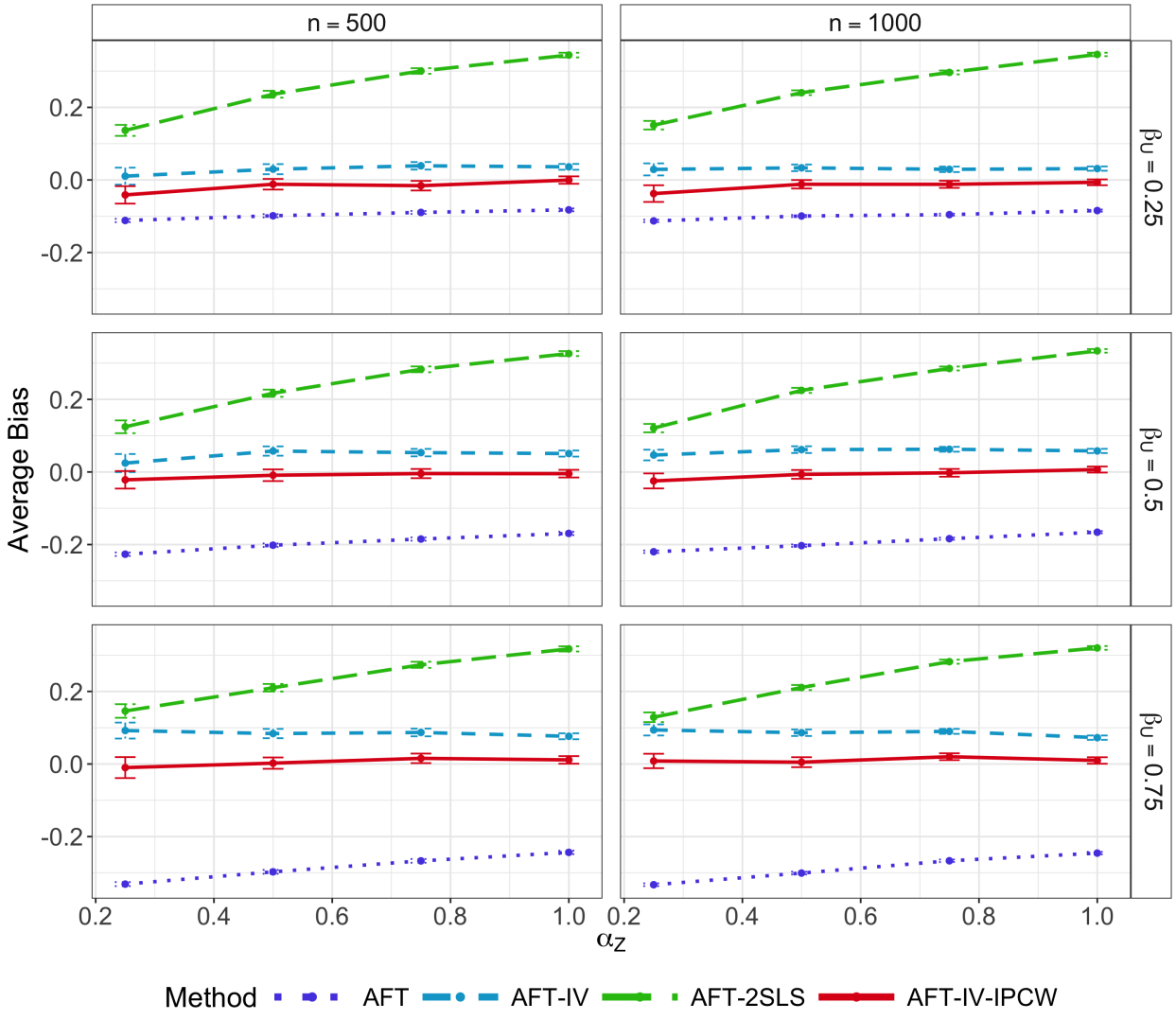

Figure 2. Simulation bias results holding $\alpha_U = 0.75$ and varying β_U, α_Z, and the sample size. The censoring distribution is independent of covariates. This figure appears in color in the electronic version of this article.

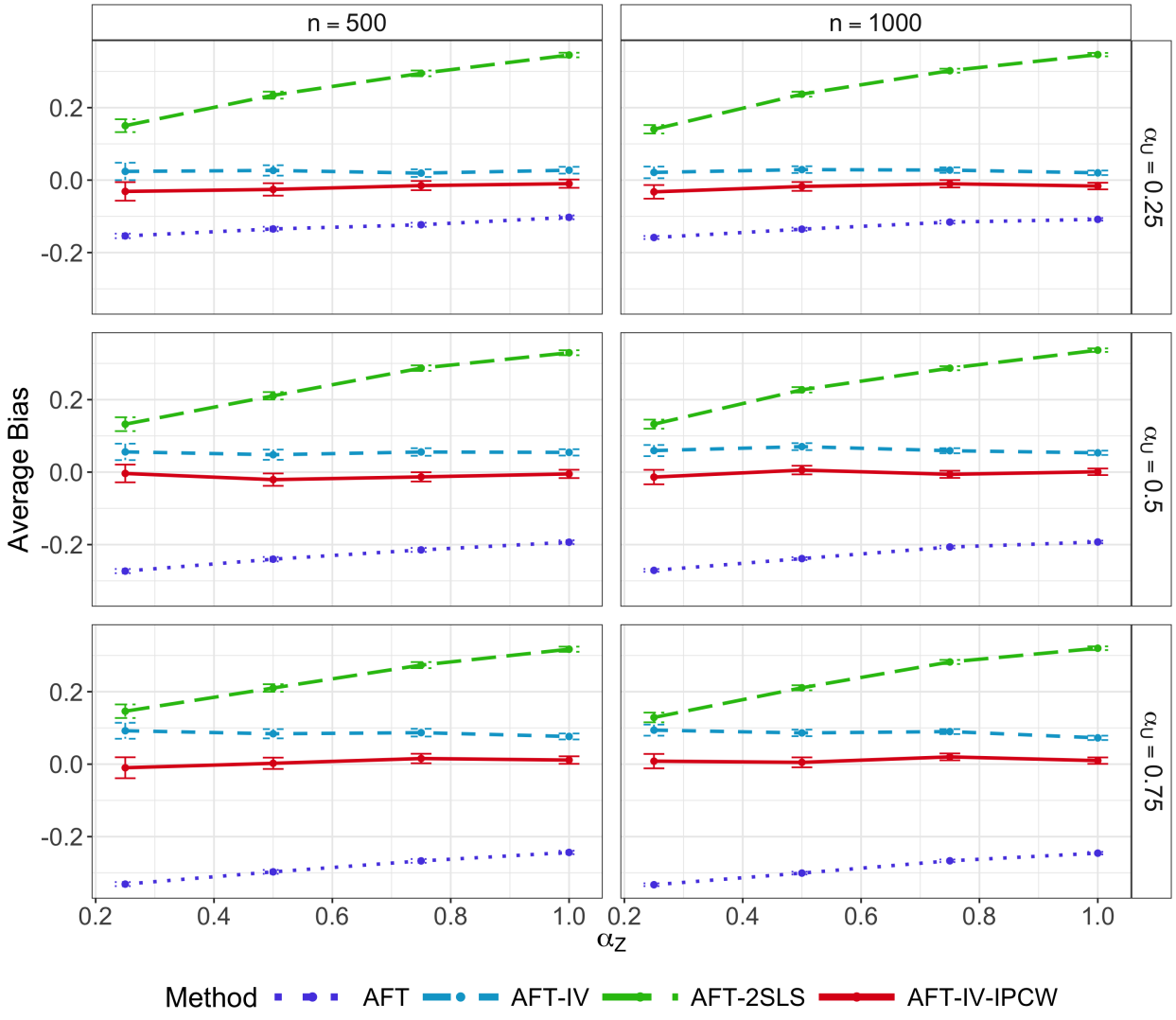

Figure 3. Simulation bias results holding $\beta_U = 0.75$ and varying α_U, α_Z, and the sample size. The censoring distribution is independent of covariates. This figure appears in color in the electronic version of this article.

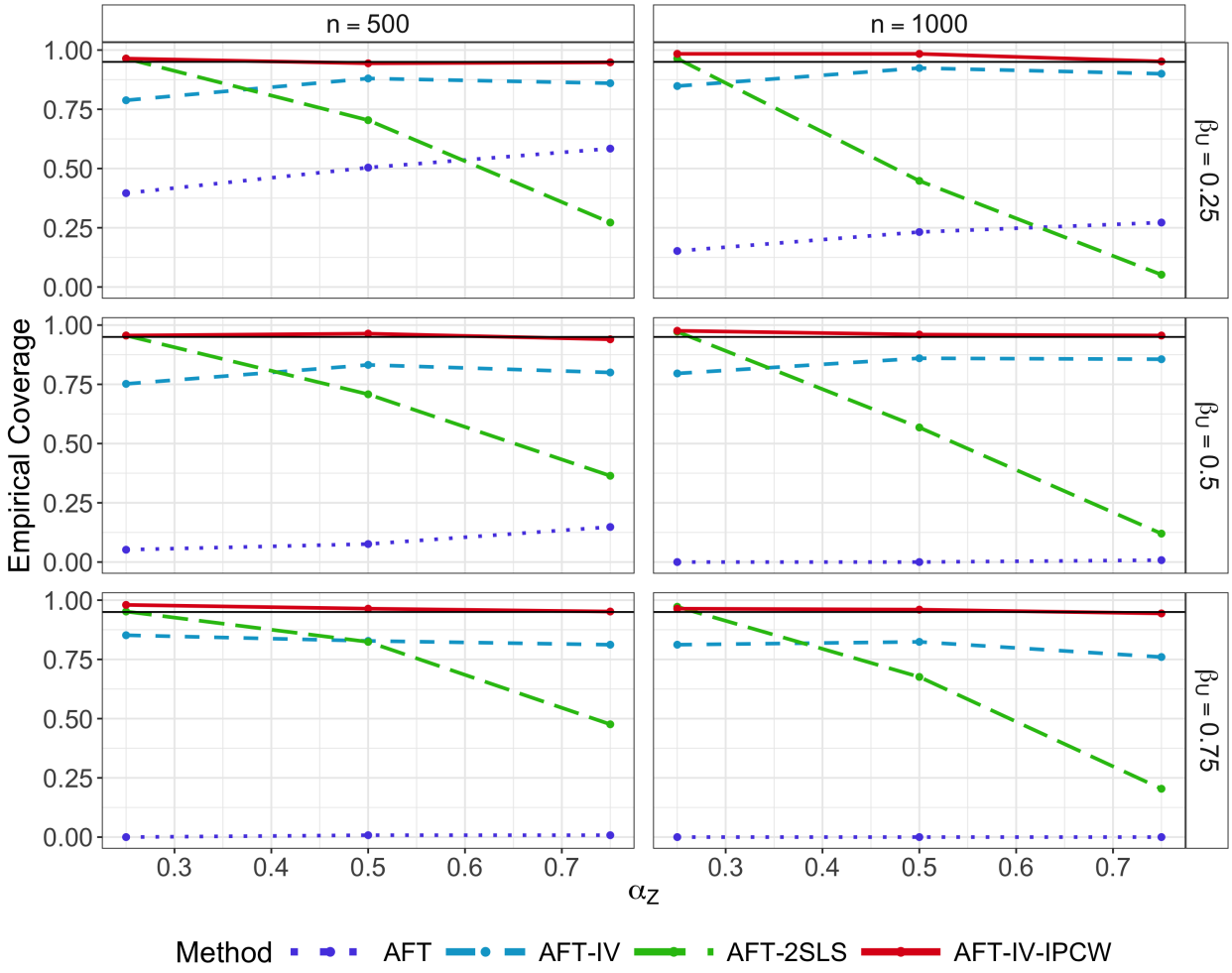

Figure 4. Empirical coverage for 95% confidence intervals holding $\alpha_U = 0.75$ and varying β_U, α_Z, and the sample size. The censoring distribution is independent of covariates. This figure appears in color in the electronic version of this article.

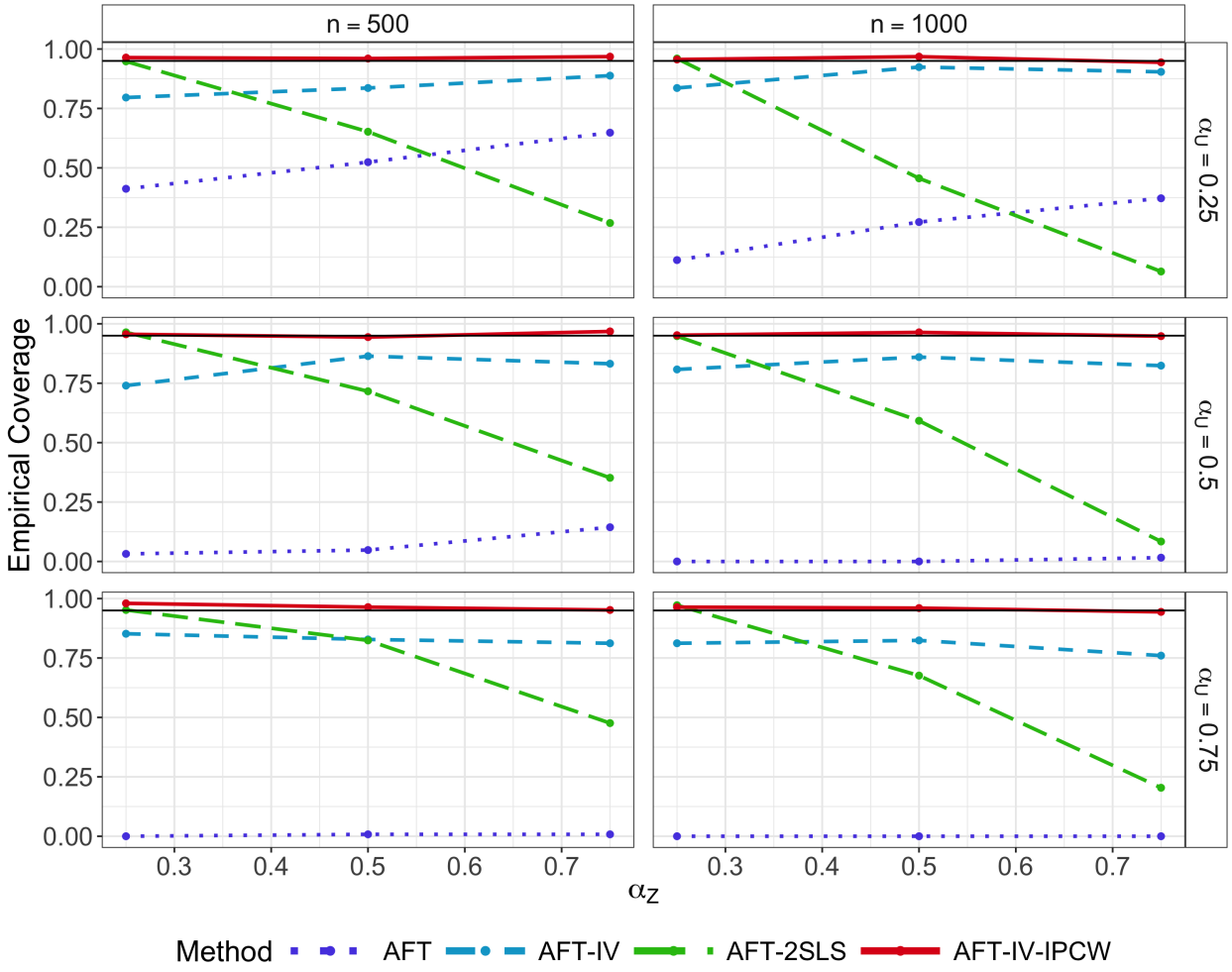

Figure 5. Empirical coverage for 95% confidence intervals holding $\beta_U = 0.75$ and varying α_U, α_Z, and the sample size. The censoring distribution is independent of covariates. This figure appears in color in the electronic version of this article.

Table 1
Estimates and corresponding confidence intervals of the effect of endovascular vs. open repair for rupture cases based on the different estimators.

Estimator	$\hat{\beta}$	(95% Conf.	Interval)
AFT	0.047	-0.063	0.144
AFT-IV	-0.169	-0.420	0.080
AFT-2SLS	-0.175	-0.432	0.074
AFT-IV-IPCW	-0.156	-0.364	0.052