Understanding the Impact of Competing Events on Heterogeneous Treatment Effect Estimation from Time-to-Event Data

Alicia Curth Mihaela van der Schaar

August 30th 2024

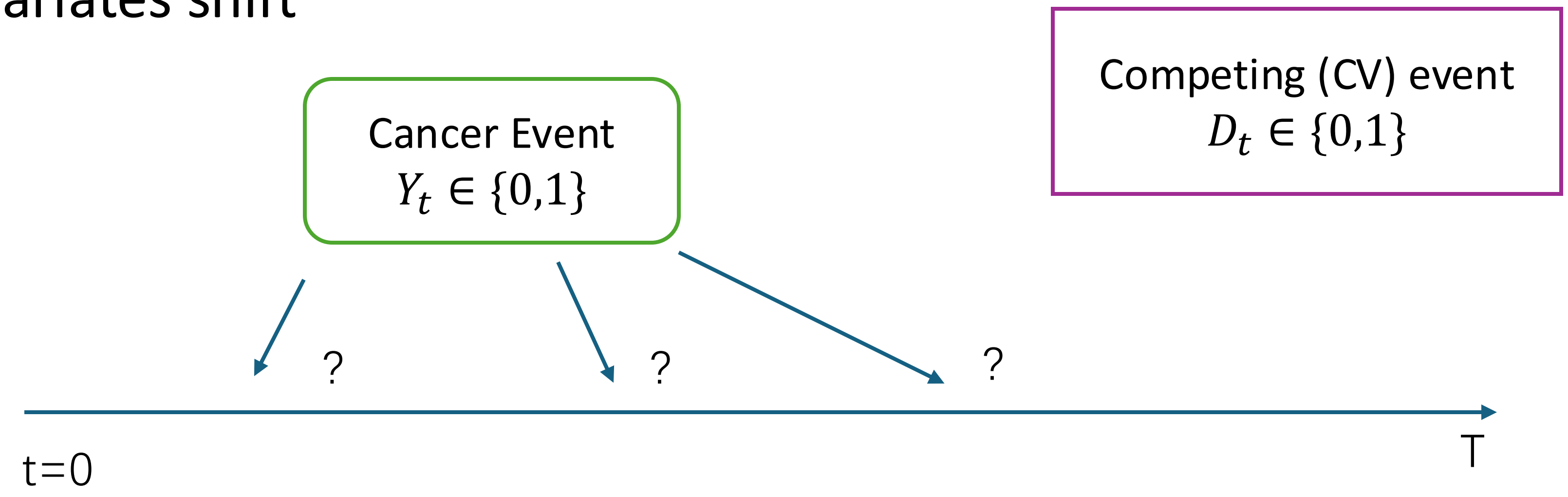
Presented by Yuqi Li

Keywords of this problem

- Time-to-event Setting
- Heterogeneous Treatment Effect
- Competing Events
- Treatment effect of interest
- Covariates shift

X, pre-treatment covariates

A, treatment assignment



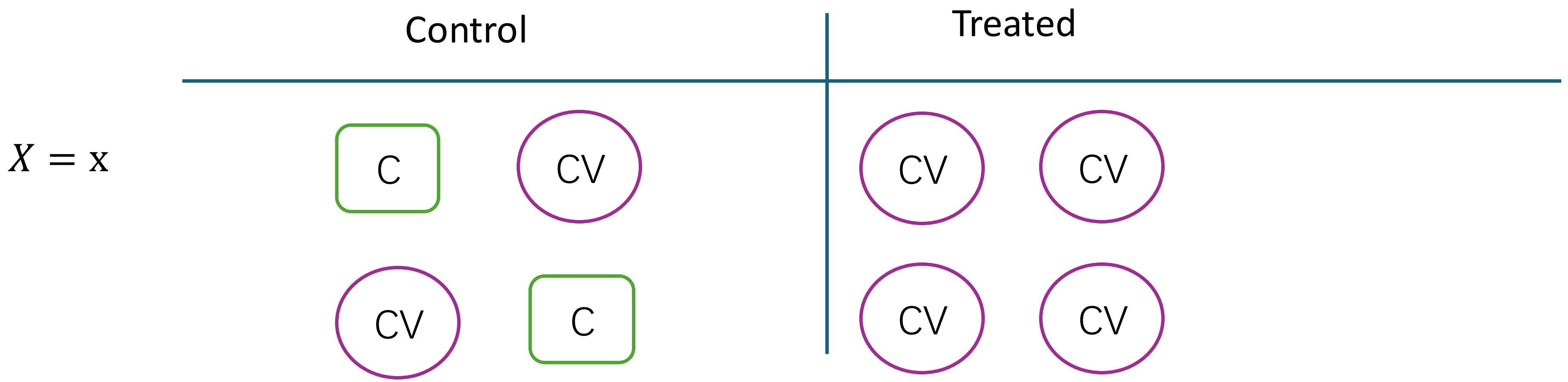
Section 1: Introduction

What if we use standard treatment effect toolbox?

The most basic approach: Estimating heterogeneous treatment effects by comparing potential outcome predictions

$$\hat{P}(Y_T = 1 | A = 0, X = x) \quad \text{vs.} \quad \hat{P}(Y_T = 1 | A = 1, X = x)$$

$$= 1/2 \qquad\qquad = 0$$



Competing event: Conceptual Challenge

Competing events act as mediators of the effect of treatment on the outcome of interest. (Young et al, 2020; Stensrud et al, 2021)

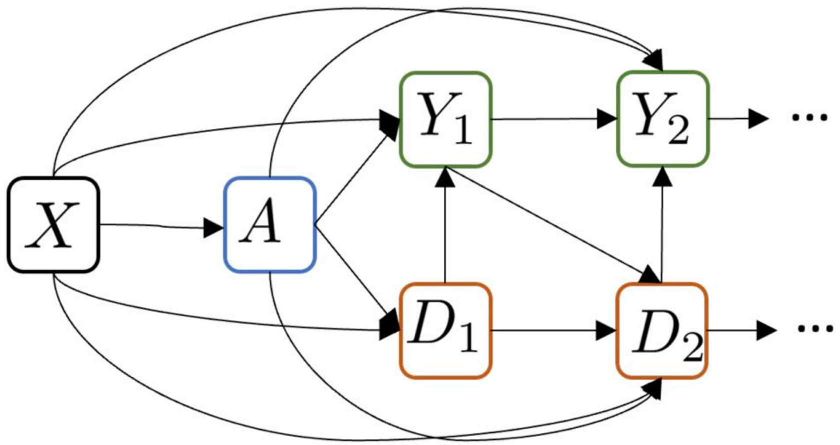


Figure 1. Assumed causal graph

Thus, there is more than one heterogeneous effect of possible interest here.
Assume no loss to follow-up due to censoring for simplicity.

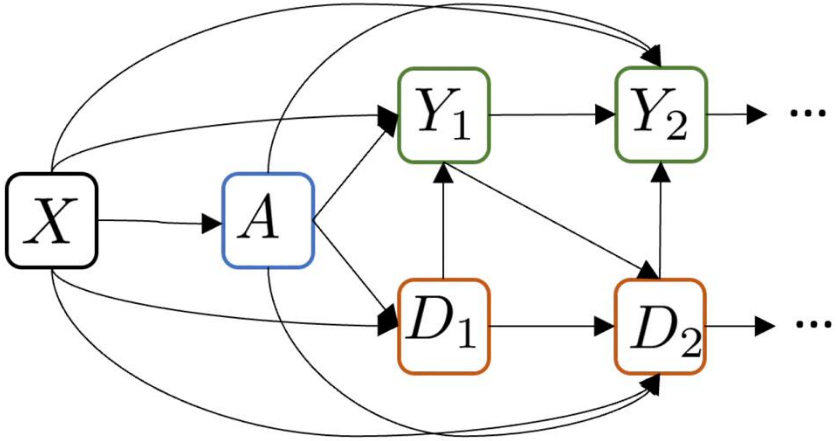
Section 2: Problem Setup

Data format

The data can be presented in a short format (X, A, T, E)

where T indicates the time at which the event occurred (i.e. $T = \min k\colon\ Y_k = 1 \cup D_k = 1$)

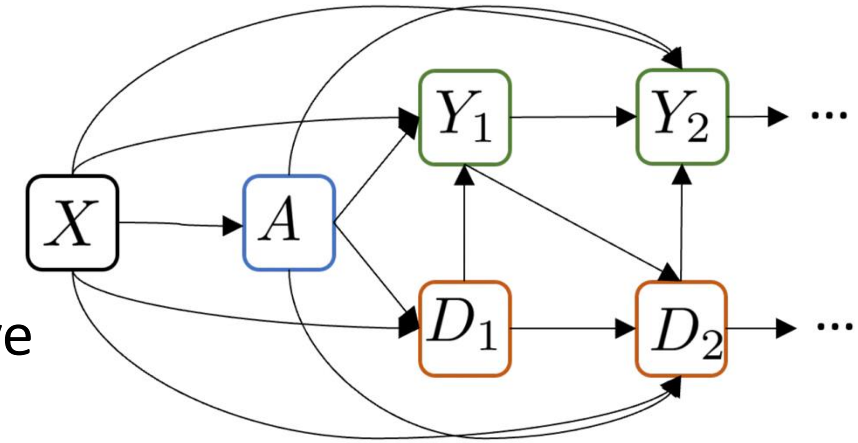
and $E \in \{Y, D\}$ indicates its type (i.e. $E = Y$ if $Y_T = 1$ else D).



Modeling Risk Function

$$h_Y(k,x,a) = \mathbb{P}(Y_k{=}1 \mid \bar{D}_k{=}\bar{Y}_{k-1}{=}0, X{=}x, A{=}a)$$

$$h_D(k,x,a) = \mathbb{P}(D_k{=}1 \mid \bar{D}_{k-1}{=}\bar{Y}_{k-1}{=}0, X{=}x, A{=}a)$$



These can be used to model the cause-specific cumulative incidence function, or risk, of an event occurring by time k (rely on the conditional hazard functions):

$$\mathbb{P}(T \le k, E{=}Y \mid X{=}x, A{=}a) = \mathbb{P}(Y_k{=}1 \mid X{=}x, A{=}a) = \sum_{l=1}^{k} h_Y(l,x,a) \times \prod_{q=1}^{l-1} (1-h_Y(q,x,a))(1-h_D(q,x,a))$$

5/16

Section 3: Defining and Estimating HTE given Competing Events

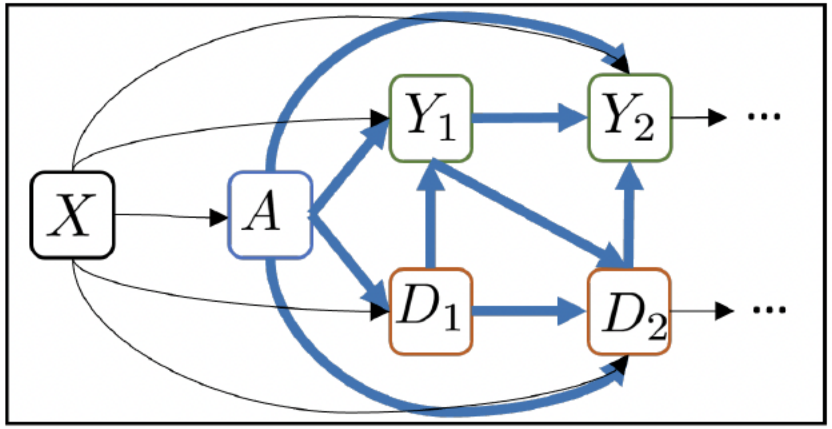
3.1 Total Effects

$$\hat{\tau}(x)=\hat{\mathbb{P}}(Y_K=1|X=x, A=1)-\hat{\mathbb{P}}(Y_K=1|X=x, A=0)$$

Requires ignorability assumption w.r.t treatment A

Outcome modeling:

1) appending A to X like a standard covariate (S-learner)
2) fitting 2 separate prediction models (T-learner)



(A) Total Effect

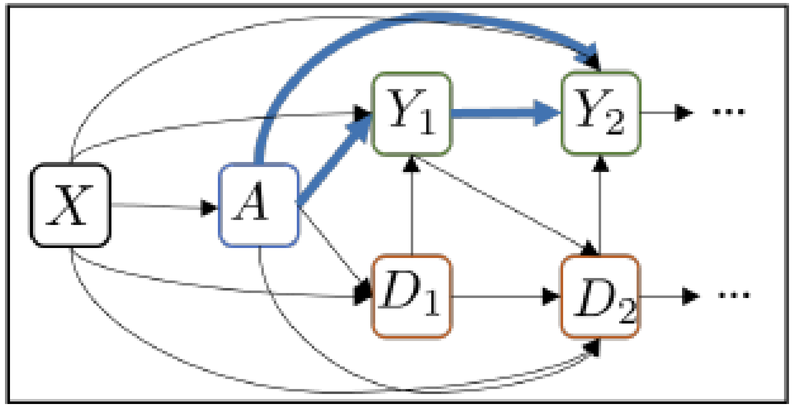
However, with competing events, even when treatment doesn't affect the primary event at all, it is possible that $P(Y_k^a = 1|X = x) \neq P(Y_k^a = 0|X = x)$ if the treatment affects the competing event.

In the presence of competing events, the total effects, which focus on ***all-cause survival***, may not always be the effect of most natural interest to an investigator.

3.2 Direct Effects

One could be interested in estimating $Y_K^{a,\bar{d}=0}$, where an additional intervention is made, and the competing event is eliminated:

$$\tau(x) = \mathbb{P}(Y_K^{1,\bar{d}=0} = 1 | X = x) - \mathbb{P}(Y_K^{0,\bar{d}=0} = 1 | X = x)$$



(B) Direct Effect

This requires ignorability assumption w.r.t A and competing event D.

In this setting, direct effects could be estimated from observational data, as

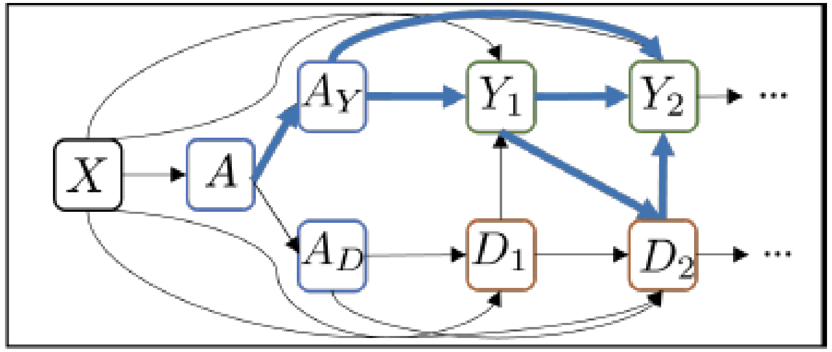
$$\mathbb{P}(Y_K^{a,\bar{d}=0}=1|X = x) = \mathbb{P}(Y_K=1|\bar{D}_K=0, A=a, X=x)$$

$$\mathbb{P}(Y_K^{a,\bar{d}=0}=1|X=x) = \sum_{l=1}^{K} h_Y(l,x,a) \prod_{q=1}^{l-1}(1-h_Y(q,x,a))$$

-- The dependence on h_D has been removed.
-- The direct risk treats competing events like a source of independent censoring.

3.3 Separable Direct Effects

Separable direct and in-direct effects a re path-specific effect that assume A conceptually consists of components A_Y and A_D.



(C) Separable Direct Effect

This requires an ignorability assumption w.r.t A and

Dismissible components

$$\mathbb{P}(Y_k^{a_Y,a_D=1}=1|\bar{Y}_{k-1}^{a_Y,a_D=1}=0,\bar{D}_k^{a_Y,a_D=1}=0,X=x)=\mathbb{P}(Y_k^{a_Y,a_D=0}=1|\bar{Y}_{k-1}^{a_Y,a_D=0}=0,\bar{D}_k^{a_Y,a_D=0}=0,X=x)$$

Risk

$$\mathbb{P}(Y_K^{a_Y,a_D}=1|X=x)=\sum_{l=1}^{K}h_Y(l,x,a_Y)\prod_{q=1}^{l-1}(1-h_Y(q,x,a_Y))(1-h_D(q,x,a_D))$$

Type of treatment effect of interest: motivation

Section 4: Covariates due to competing events

Data available for training follows an observational distribution, while the target quantities are defined with respect to interventions --- **covariate shift arise.**

- Treatment selection bias (confounding)
- Censoring act as an additional source of covariate shift
- Competing events can act as an additional source, depending on the chosen treatment effect of interest

Learning treatment-specific hazard functions

This can be framed as a standard ML classification problem:

"Of those at risk today, who will experience their event today?"

Using ERM
Hazard estimator

$$\hat{h}^{Y}(k, x, a) \in \arg\min_{h \in \mathcal{H}} \hat{R}^{obs}_{a,k}(h)$$

$$\hat{R}^{obs}_{a,k}(h) = \sum_{i \in \mathcal{I}_Y(a,k)} \ell(Y_{k,i}, h(X_i)) = \hat{\mathbb{E}}_{X \sim p^{obs}_{a,k}, Y_k \sim h_Y(k,x,a)}[\ell(Y_k, h(X))]$$

Section 4: Covariates due to competing events

Data distribution mismatch $\arg\min_{h\in\mathcal{H}} R^{int}_{a,k}(h) \neq \arg\min_{h\in\mathcal{H}} R^{obs}_{a,k}(h)$

Importance weights $w^*(x) \propto \frac{\mathbb{P}^{int}_{a,k}(X=x)}{\mathbb{P}^{obs}_{a,k}(X=x)}$

$$\hat{R}^{w,obs}_{a,k}(h) = \sum_{i\in\mathcal{I}_Y(a,k)} w^*(X_i)\ell(Y_{k,i}, h(X_i))$$

When all effects are identified, the main learning challenge lies in covariate shift.

--- Different effects correspond to different conceptual interventions

Section 4: Covariates due to competing events

- Total effect: $\mathbb{P}_{a,k}^{do(A=a)}(X=x)$ $\quad w_{a,k}^{*,do(A=a)}(x) \propto P(A=a|X=x)^{-1}$
  Only confounding-induced shift (if treatment propensities vary)

- Separable effect: $\mathbb{P}_{a,k}^{do(A_Y=a_Y,A_D=a_D)}(X=x)$ $\quad w_{a,k}^{*,do(A=a,\bar{D}_K=0)}(x) \propto \left[P(A=a|X=x)\prod_{l=1}^{k}(1-h_D(l,x,a))\right]^{-1}$
  Additional shift if treatment effect on D is heterogeneous

- Direct effect: $\mathbb{P}_{a,k}^{do(A=a,\bar{D}_K=0)}(X=x)$
  Additional shift if risk of competing event depends on covariates
  $$w_{a,k}^{*,do(A_Y=a_Y,A_D=a_D)}(x) \propto \left[P(A=a|X=x)\frac{\prod_{l=1}^{k}h_D(l,x,a_Y)}{\prod_{l=1}^{k}h_D(l,x,a_D)}\right]^{-1}$$

Section 5: Experiments

Setting 1: There is confounding when $|\xi| > 0$, but treatment has no effect on either event.

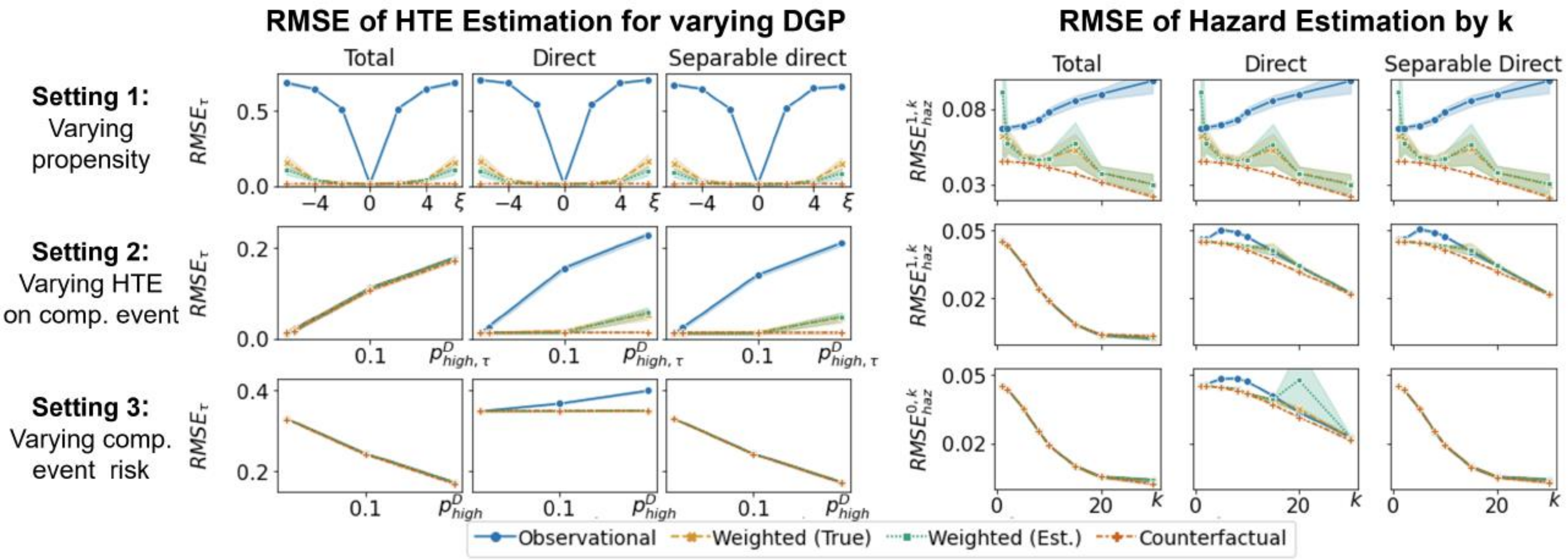
Setting 2: No Confounding but Treatment Affects Competing Events
There is no confounding ($\xi = 0$), treatment has no effect on the main event but affects the competing event

Setting 3: No Confounding but HTE on Main Event
There is no confounding ($\xi = 0$), treatment has a heterogeneous effect on the main event, but has no effect on the competing event. When the high-risk group is also at higher (baseline) risk of the competing event, this may mask the protective effect of treatment on the main event

Estimators

1. Observational Empirical Risk Minimization (ERM)
2. Weighted ERM with True Importance Weights
3. Weighted ERM with Estimated Importance Weights
4. Counterfactual Estimator

Section 5: Experiments



14/16

Conclusions

- **Complexity of HTE Estimation with Competing Events**

The inclusion of competing events complicates heterogeneous treatment effect (HTE) estimation by introducing multiple definitions of effects and sources of covariate shifts. This complexity affects how and when different biases influence the estimation of each effect.