Debiased Front-Door Learners for Heterogeneous Effects

Yonghan Jung

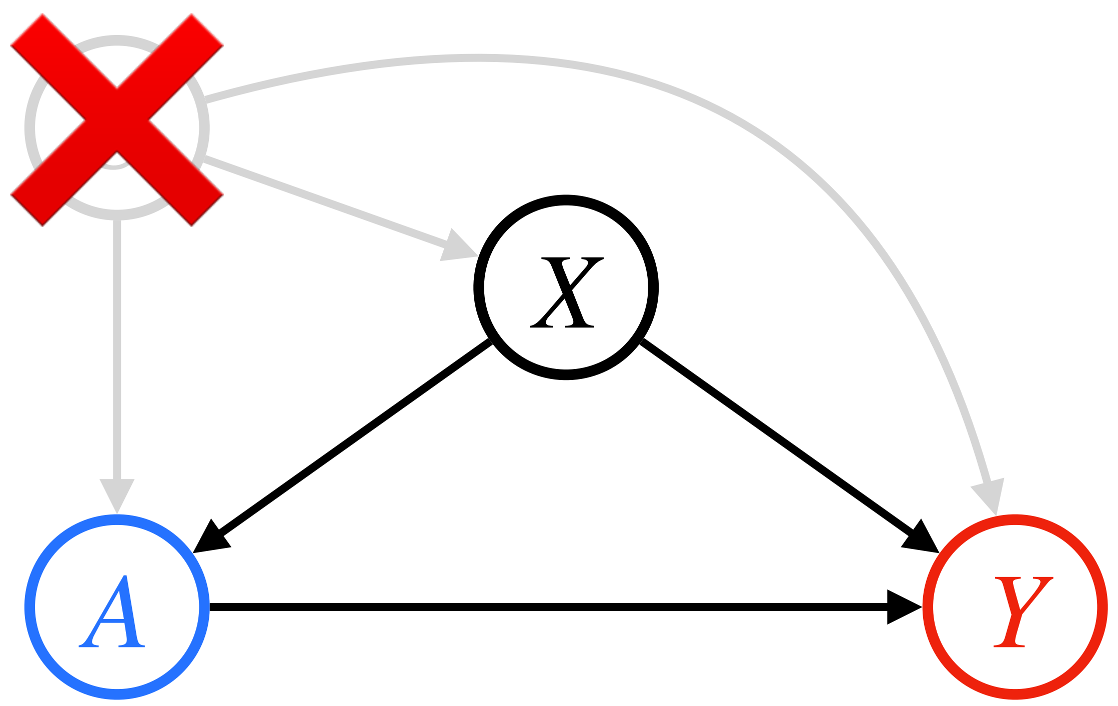
School of Information Sciences (iSchool)

University of Illinois Urbana-Champaign (UIUC)

www.yonghanjung.me

2025 CDSM

No Unmeasured Confounder Assumptions



Between treatments A and outcomes Y,
all confounder are measured (as X)

Unmeasured confounders U doesn't exist.

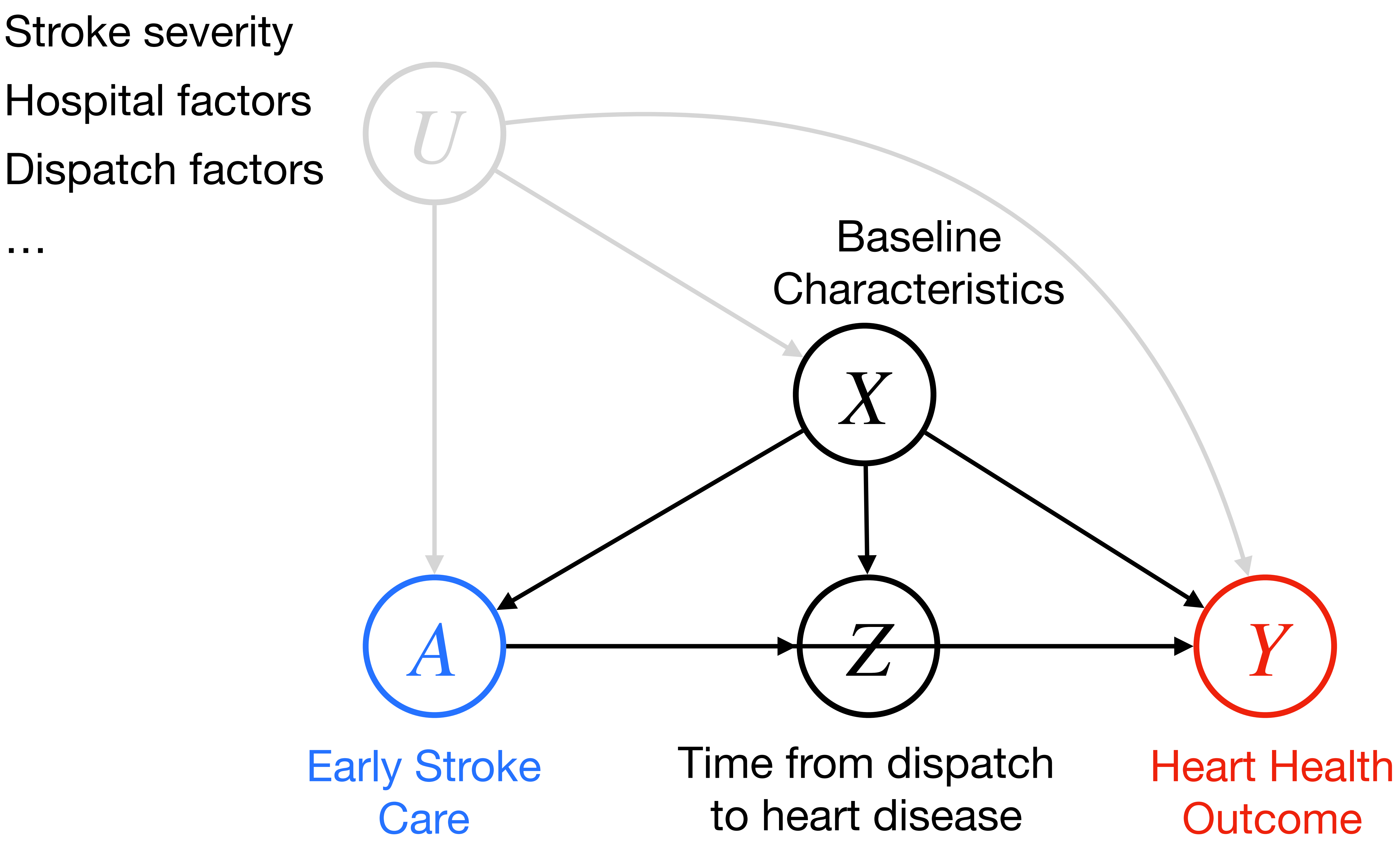
⚠️ Hardly satisfied in practice

Non-identifiable causal effects

A	U	Y
Smoking	Genetic Factor	Cancer
Attendance	Motivation	GPA
Treatment	Disease history	Recovery
Job training	Economic trend	Hiring status

🏆 Award-winning real-world example

Leveraging Mediators: Public Health

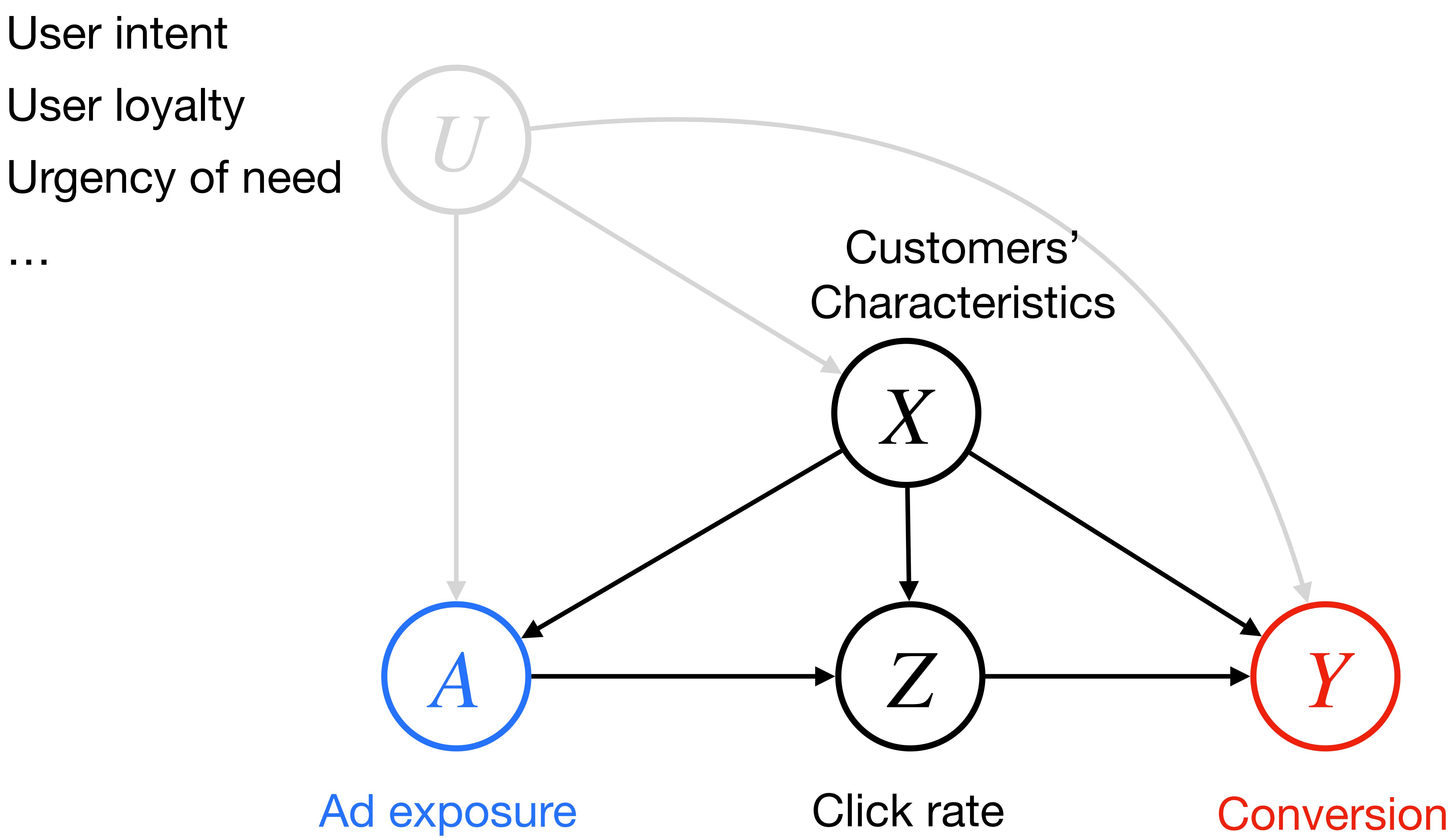


$$\mathbb{E}[Y \mid \mathrm{do}(a)] = \sum_{z,a',x} \mathbb{E}[Y \mid z, a', x] P(z \mid a, x) P(a', x)$$

$\mathbb{E}[Y \mid \mathrm{do}(a)]$ is not identifiable.

“Front-door adjustment”

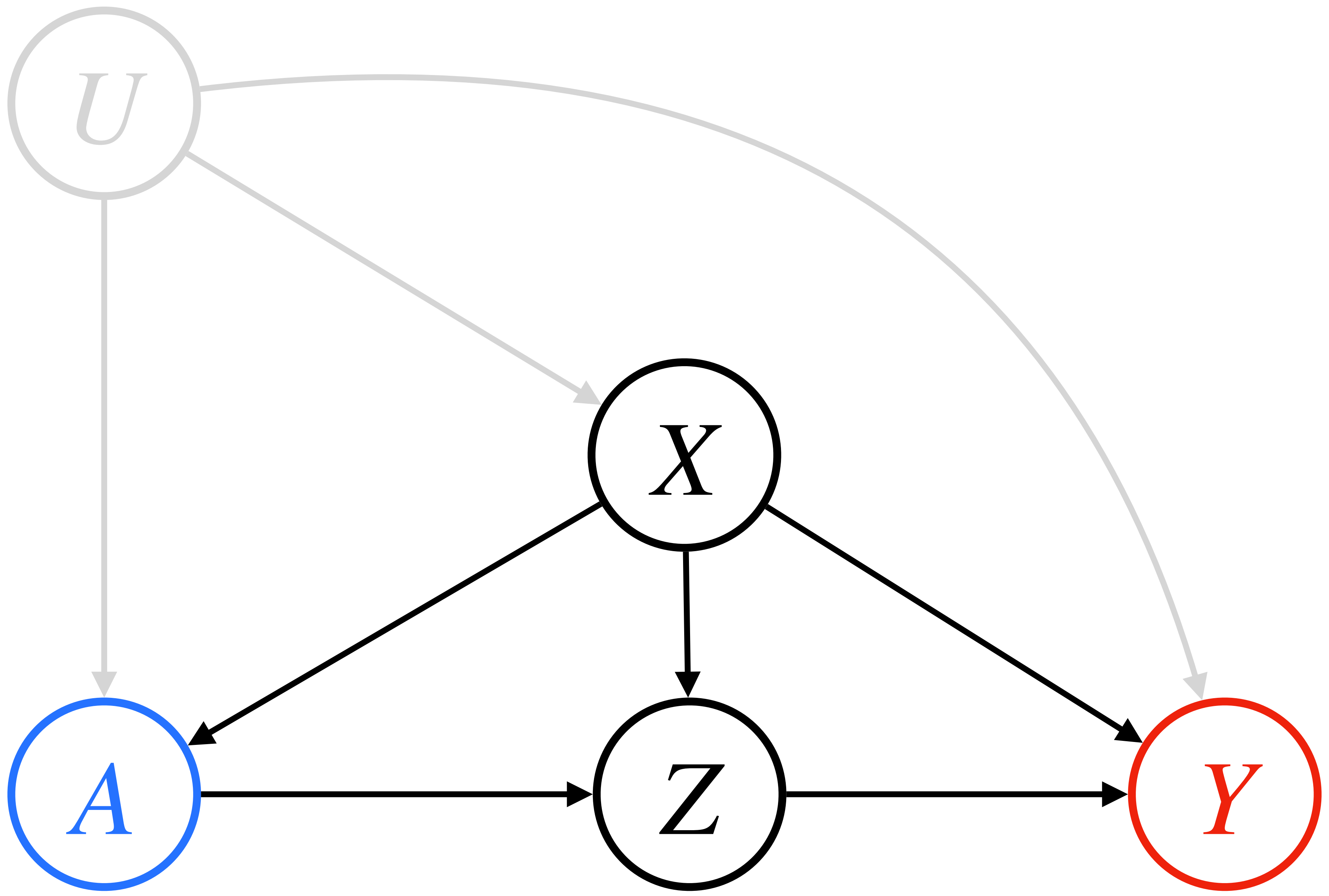
Leveraging Mediators: Digital Experiments



$$\mathbb{E}[Y \mid do(a)] = \sum_{z,a',x} \mathbb{E}[Y \mid z,a',x] P(z \mid a,x) P(a',x)$$

"Front-door adjustment"

Front-Door Model



Front-Door Criterion

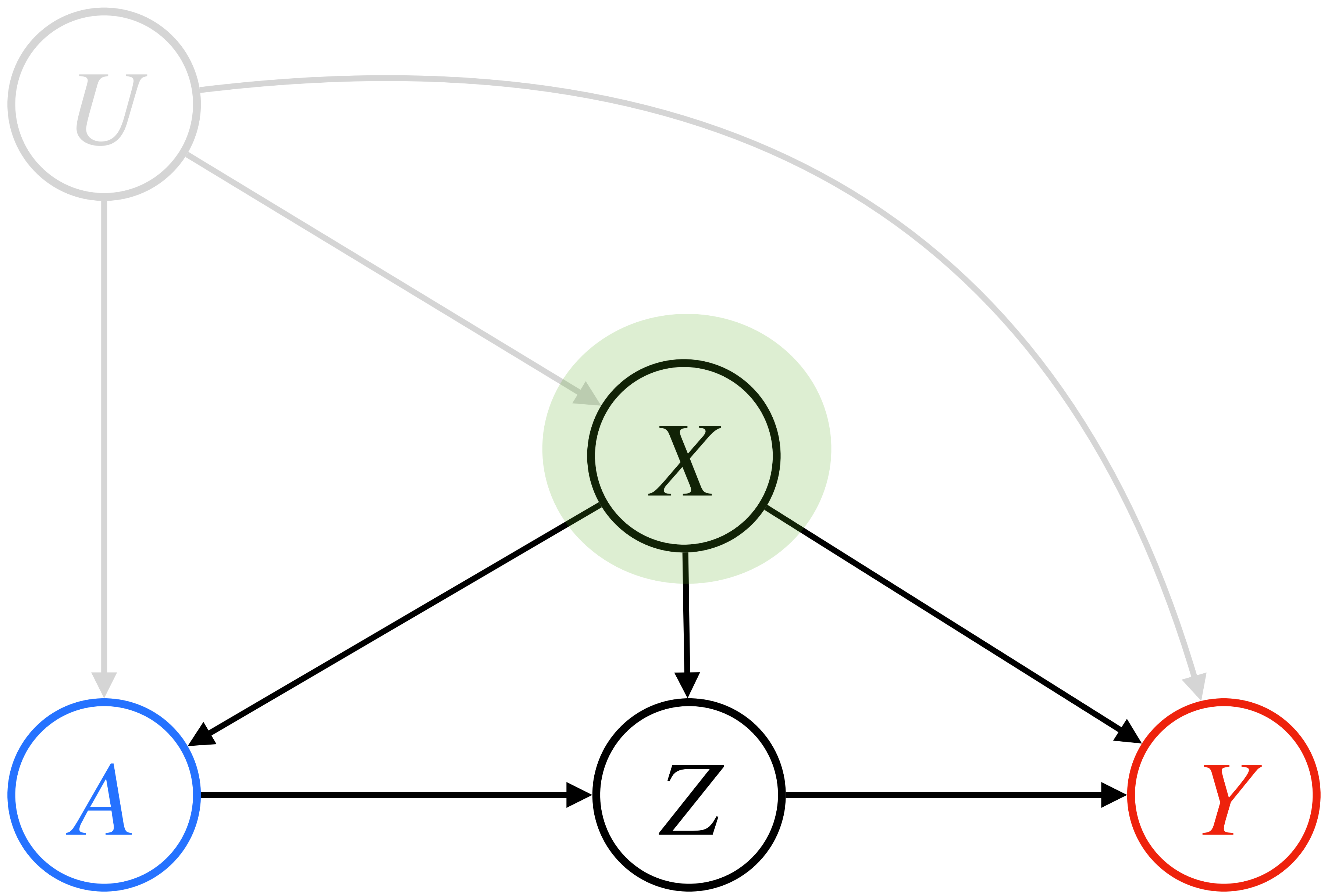
"Good Mediator Z"

1. Z blocks all directed path from A to Y
2. No unmeasured confounders of $A \rightarrow Z$
3. No unmeasured confounders of $Z \rightarrow Y$

$$\mathbb{E}[Y \mid \mathrm{do}(a)] = \sum_{z,a',x} \mathbb{E}[Y \mid z,a',x] P(z \mid a,x) P(a',x)$$

FD Heterogeneous Treatment Effect

NEW



Front-door effect at characteristics $X = x$

$$\mathbb{E}[Y \mid do(a), x]$$

$$= \sum_{z,a',x} \mathbb{E}[Y \mid z, a', x] P(z \mid a, x) P(a' \mid x)$$

FD-CATE: $\mathbb{E}[Y \mid do(A = 1), x] - \mathbb{E}[Y \mid do(A = 0), x]$

Front-door Heterogeneous Treatment Effect

$$\mathbb{E}[Y \mid \mathrm{do}(a), x] = \sum_{z,a',x} \mathbb{E}[Y \mid z,a',x] P(z \mid a,x) P(a' \mid x) = \sum_{z,a',x} m(z,a',x) q(z \mid a,x) e_{a'}(x)$$

FD-CATE $\tau(X)$**:** $\underbrace{\mathbb{E}[Y \mid \mathrm{do}(A=1), X]}_{\tau_1(X)} - \underbrace{\mathbb{E}[Y \mid \mathrm{do}(A=0), X]}_{\tau_0(X)}$

$$= \sum_{z,a'} \{q(z \mid 1,x) - q(z \mid 0,x)\} e_{a'}(x) m(z,a',x)$$

FD-DR-Learner: Doubly Robust Learner for FD

$$\xi_a(ZAX) = \frac{q(Z \mid aX)}{q(Z \mid AX)}$$

$$\pi_a(AX) = \frac{\mathbb{I}(A = a)}{e_A(X)}$$

$$r_{me}(ZX) = \sum_{a'} m(Z, a', X) e_{a'}(X)$$

$$\nu_{meq}(AX) = \sum_{z} r_{me}(zX) q(z \mid A, X)$$

$$s_{mq_a}(AX) = \sum_{z} m(z, A, X) q(z \mid a, X)$$

FD effect $\tau_a(X) = \mathbb{E}[Y \mid \mathrm{do}(a)]$

$$\mathbb{E}[\xi_a(ZAX) \times \{Y - m(ZAX)\}]$$

$$+ \mathbb{E}[\pi_a(AX) \times \{r_{me}(ZX) - \nu_{meq}(AX)\}]$$

$$+ \mathbb{E}[s_{mq_a}(AX)]$$

FD Pseudo-Outcome

$$\varphi_{mqe,a}(YZAX)$$

FD-DR-Learner: Doubly Robust Learner for FD

Learn nuisances from D_{train}		Evaluate functions		Evaluate PO from D_{eval}		Regress $\hat{\varphi}$ onto X
$\hat{m}(ZAX)$, $\hat{q}(Z\|AX)$, $\hat{e}_A(X)$	$\longrightarrow$	$\hat{\xi}_a(ZAX)$, $\hat{\pi}_a(AX)$, $r_{\hat{m}\hat{e}}(ZX)$, $\nu_{\hat{m}\hat{e}\hat{q}}(AX)$, $s_{\hat{m}\hat{q}}(AX)$	$\longrightarrow$	FD Pseudo-Outcome $\hat{\varphi}_{\hat{m}\hat{q}\hat{e},a}(YZAX)$	$\longrightarrow$	$\hat{\mathbb{E}}[\hat{\varphi}_{\hat{m}\hat{q}\hat{e},a}(YZAX)\|X]$ $\hat{\tau}_a(X)$

FD-DR-Learner: $\hat{\tau}_{\text{DR}}(X) = \hat{\tau}_1(X)$ - $\hat{\tau}_0(X)$

FD-DR-Learner: Doubly Robust Learner for FD

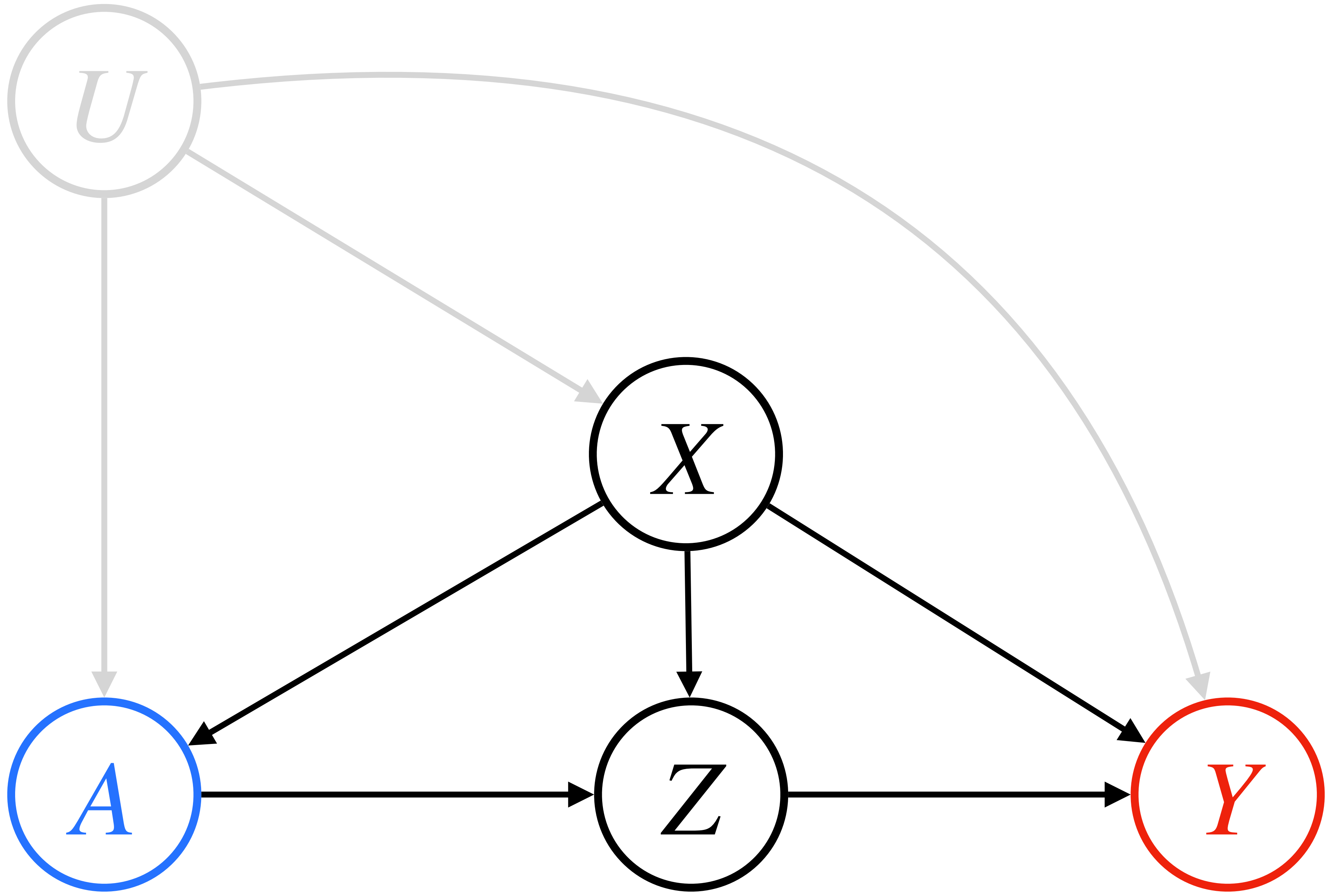
$$\text{error}(\hat{\tau}_{\text{DR}}) \lesssim \text{oracle-rate} + \{\text{error}(\hat{m}) \times \text{error}(\hat{q})\} + \{\text{error}(\hat{q}) \times \text{error}(\hat{e})\}$$

oracle-rate $\quad n^{-1/4} \quad n^{-1/4} \quad n^{-1/4} \quad n^{-1/4}$

- **Double Robustness:** error $\to 0$ *fast* if either error($\hat{m}$)=0 & error($\hat{e}$)=0; or error($\hat{q}$) =0

- **Fast Convergence:** error $\to 0$ *fast* even when error($\hat{m}$), error($\hat{e}$), error($\hat{q}$) goes to zero slowly

FD-R-Learner



FD Parametrization

$$A = e_A(X) + \epsilon_A$$

$$Z = \alpha(X) + A \cdot \beta(X) + \epsilon_Z$$

$$Y = f(AX) + Z \cdot g(AX) + \epsilon_Y$$

$$\mathbb{E}[\epsilon_A \mid X] = 0$$

$$\mathbb{E}[\epsilon_Z \mid AX] = 0$$

$$\mathbb{E}[\epsilon_Y \mid ZAX] = 0$$

FD-R-Learner

FD-CATE $\tau(X)$**:** $\mathbb{E}[Y \mid \mathrm{do}(A=1), X] - \mathbb{E}[Y \mid \mathrm{do}(A=0), X]$

FD Parametrization

$$A = e_A(X) + \epsilon_A$$

$$Z = \alpha(X) + A \cdot \beta(X) + \epsilon_Z$$

$$Y = f(AX) + Z \cdot g(AX) + \epsilon_Y$$

$$\mathbb{E}[\epsilon_A \mid X] = 0$$

$$\mathbb{E}[\epsilon_Z \mid AX] = 0$$

$$\mathbb{E}[\epsilon_Y \mid ZAX] = 0$$

$$\beta(X) = \mathbb{E}[Z \mid \mathrm{do}(A=1), X] - \mathbb{E}[Z \mid \mathrm{do}(A=0), X]$$

$$g(AX) = \mathbb{E}[Y \mid \mathrm{do}(Z=1), AX] - \mathbb{E}[Y \mid \mathrm{do}(Z=0), AX]$$

FD-R-Learner Representation

$$\tau(X) = \beta(X)\mathbb{E}[g(AX) \mid X]$$

FD-R-Learner

Learn nuisances from D_{train}

Evaluate with D_{eval}

FD-R-Learner Representation

$$\tau(X) = \beta(X)\mathbb{E}[g(AX) \mid X]$$

$\hat{m}_Z(X) = \mathbb{E}[Z \mid X]$

$\hat{e}_A(X) = \mathbb{E}[A \mid X]$

BD-R-Learner (Nie and Wager, 2021) → $\widehat{\beta}(X)$

$\hat{m}_Y(AX) = \mathbb{E}[Y \mid AX]$

$\hat{e}_Z(AX) = \mathbb{E}[Z \mid AX]$

BD-R-Learner (Nie and Wager, 2021) → $\widehat{g}(AX)$

$\widehat{\zeta}(AX)$

$$\sum_a \hat{e}_a(X)\hat{g}(aX) + \{A\text{-}\hat{e}_A(X)\}\{\hat{g}(1X) - \hat{g}(0X)\}$$

$\widehat{\zeta}(AX) \xrightarrow{\text{Regress } \widehat{\zeta} \text{ onto } X} \widehat{\gamma}(X)$

$$\widehat{\tau}_{\mathrm{R}}(X) = \widehat{\beta}(X)\widehat{\gamma}(X)$$

FD-R-Learner: Doubly Robust Learner for FD

$n^{-1/4}$ $n^{-1/4}$ $n^{-1/4}$

$$\text{error}(\hat{\tau}_{\mathrm{R}}) \lesssim \text{oracle-rate} + \text{error}(\hat{e}_A)^2 + \{\text{error}(\hat{e}_A) \times \text{error}(\hat{m}_Z)\}$$

oracle-rate

$$+ \text{error}(\hat{e}_Z)^2 + \{\text{error}(\hat{e}_Z) \times \text{error}(\hat{m}_Y)\}$$

$n^{-1/4}$ $n^{-1/4}$ $n^{-1/4}$

- **Fast Convergence:** error $\rightarrow 0$ fast even when

  error($\hat{e}_A$), error($\hat{m}_Z$), error($\hat{e}_Z$), error($\hat{m}_Y$)
  goes to zero slowly

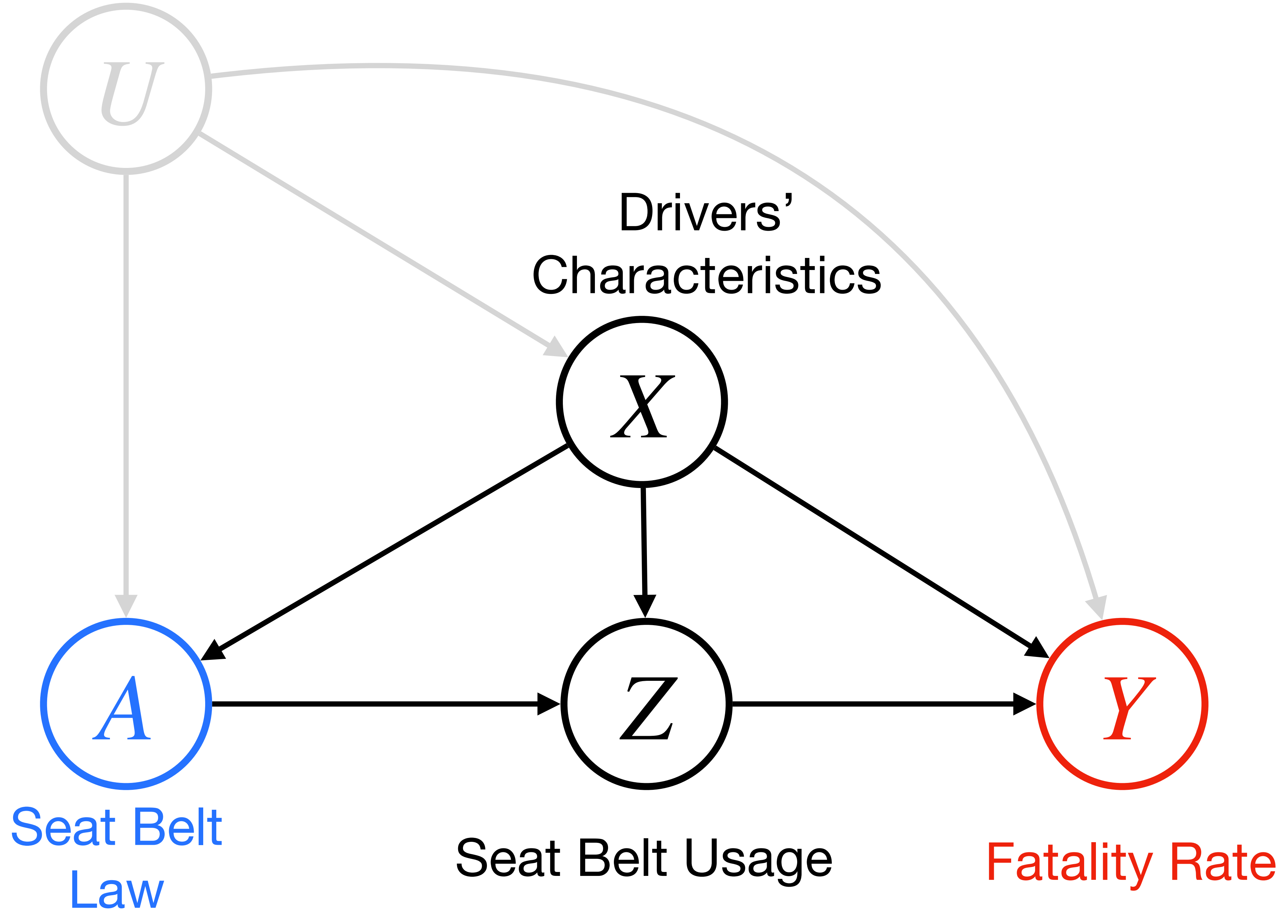
Simulation: Synthetic Data

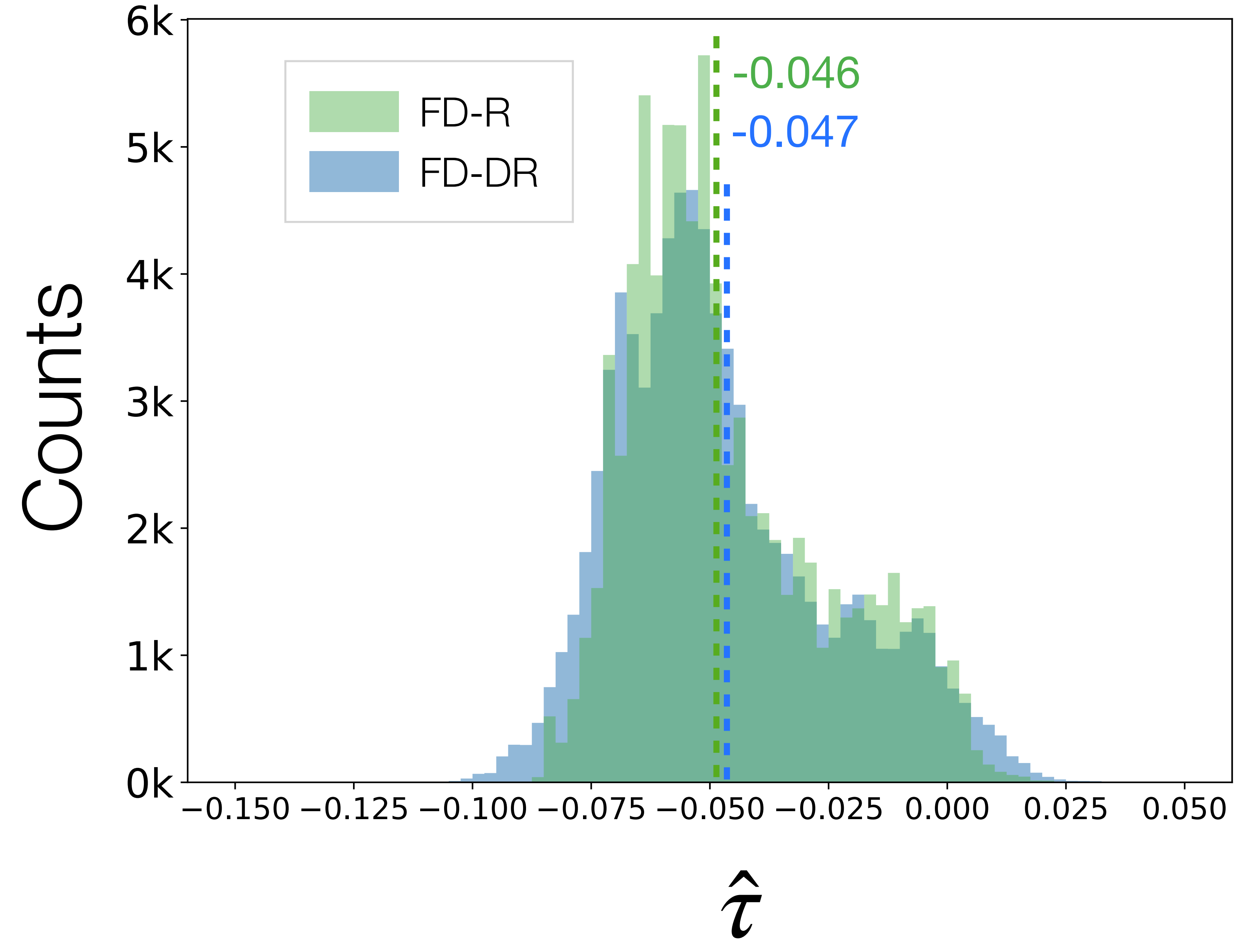
FD-DR & FD-R learners outperforms the plug-in estimator.

FD-DR & FD-R learners are more robust to slow convergences of nuisances

FD-DR & FD-R learners are more robust to weak overlap.

Simulation: NHTSA Dataset





Summary

- Front-door models are increasingly applied in real-world causal analyses.
- ATE estimators have been well-developed.
- However, CATE/HTE estimators have been largely unexplored.

We introduce two CATE estimators for FD:
(FD-DR-Learner, FD-R-Learner)