



# IDENTIFYING DIRECT CAUSAL EFFECTS UNDER UNMEASURED CONFOUNDING

Philippe Boileau<sup>\*1</sup>, Nima S. Hejazi<sup>\*2</sup>, Ivana Malenica<sup>\*1</sup>, Sandrine Dudoit<sup>1</sup>, Mark J. van der Laan<sup>1</sup>

<sup>1</sup>University of California, Berkeley; <sup>2</sup>Weill Cornell Medicine



## Introduction

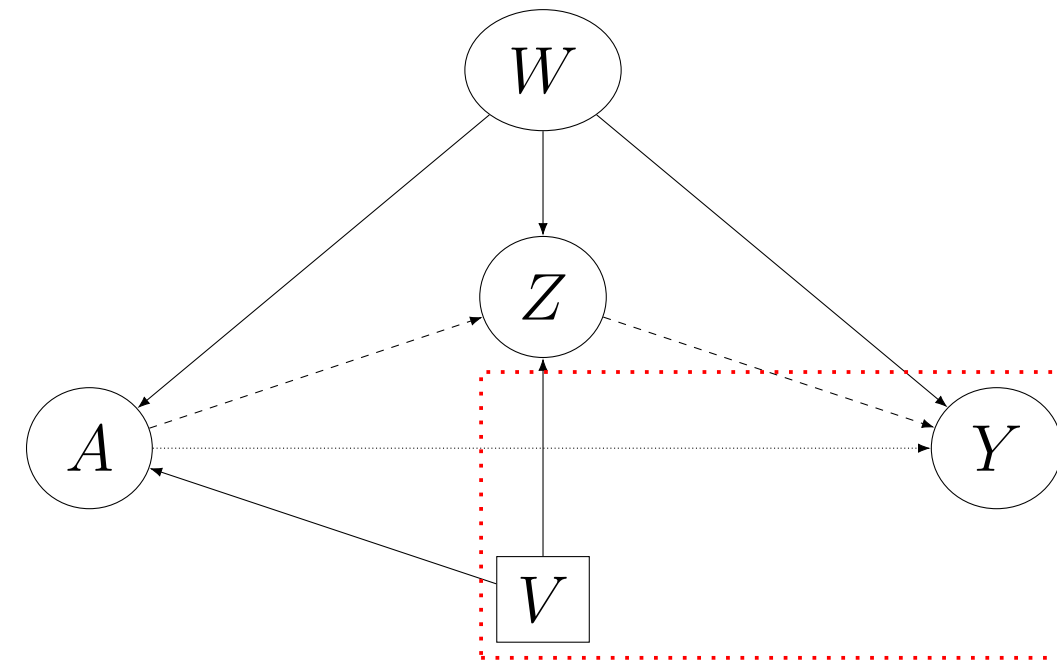
This is the background.

## Statistical Problem

State the causal and statistical models, and estimand.  
The causal target parameter is

$$\Psi^F(P_{U,X,0}) = \int_{w,z} \mathbb{E}[Y(1,z) - Y(0,z) \mid W = w] p_Z(z \mid A = 0, w) p_W(w) dz dw .$$

## Identification



- (A1) No unmeasured endogenous pathways:  
 $f_Y(Z, A, W, V, U_Y) \equiv f_Y(Z, A, W, U_Y)$ .
- (A2) Conditional expectation equivalence:  
 $\mathbb{E}(Y \mid Z, A = 1, W, V) \equiv \mathbb{E}(Y \mid Z, A = 1, W)$

### Theorem

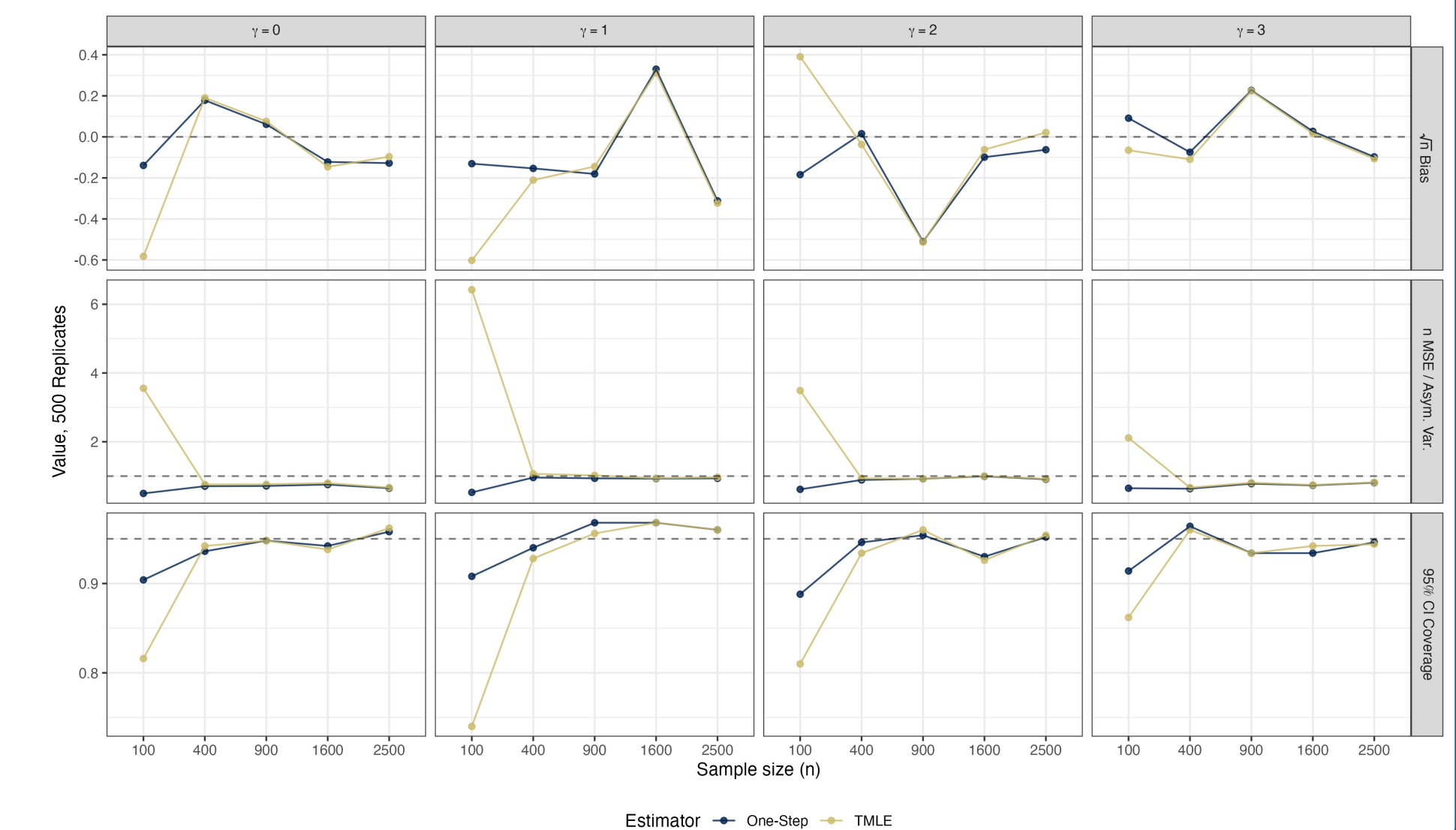
Under assumptions A1 and A2,  $\Psi^F(P_{U,X,0})$  is identified by

$$\Psi(P_0) = \mathbb{E}_{P_0} \mathbb{E}_{P_0} \{ \mathbb{E}_{P_0}(Y \mid W, A = 1, Z) - \mathbb{E}_{P_0}(Y \mid W, A = 0, Z) \mid A = 0, W \} .$$

## Simulation Study

We consider the following data-generating distribution:

$$\begin{aligned} W_1 &\sim \text{Unif}(-1, 1), \quad W_2, V \sim N(0, 1) \\ A \mid W, V &\sim \text{Bern} \left( (1 + \exp\{-W_1 - W_2 - V\})^{-1} \right) \\ Z \mid A, W, V &\sim \text{Bern} \left( (1 + \exp\{-W_1 - W_2 - \gamma V - 3A\})^{-1} \right) \\ Y \mid Z, A, W, V &\sim N(3A + W_1 + W_2 + Z, 1) . \end{aligned}$$



## Inference

Statistical inference is possible using standard methods.

## Conclusions

Here are the important takeaways.

## Funding

Thank you for paying my bills.

## References

List of references.

<sup>\*</sup> indicates shared first-authorship