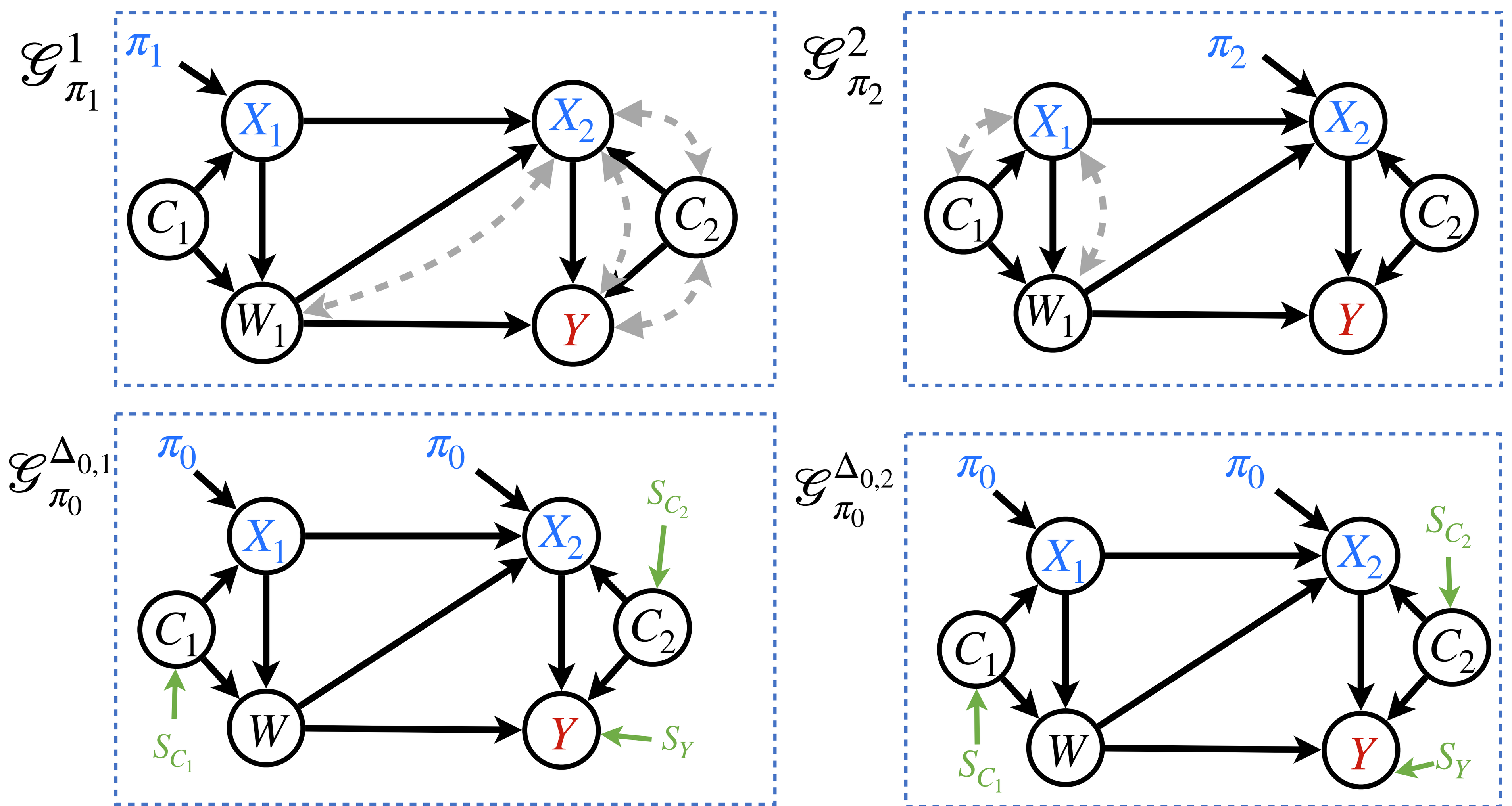
Efficient Policy Evaluation Across Multiple Different Experimental Datasets

Yonghan Jung*[1] and **Alexis Bellot***[2]   [1] Purdue University, [2] Google DeepMind

Yonghan is on the academic job market! Visit www.yonghanjung.me

Motivation and Problems



C_1, C_2: Baselines Year 1,2

X_1: Class size Year 1

π_1: Policy X_1 at source 1

W_1: Grade Year 1

X_2: Class size Year 2

π_2: Policy X_2 at source 2

Y: Grade Year 2

Input

- Graphs $\mathcal{G}^i_{\pi_i}$ for experiments to X_i with π_i on sources i.
- Data $(D^0, D^1_{\pi_1}, D^2_{\pi_2}) \sim, P^0(C_1, C_2), P^1_{\pi_1}(\mathbf{V}), P^2_{\pi_2}(\mathbf{V})$
- A target policy π_0 intervened on (X_1, X_2)
- $\mathcal{G}^{\Delta_{0,i}}_{\pi_0}$ domain discrepancies, indicated S, between source i and target $P^0_{\pi_0}$.

Output

$\mathbb{E}_{P^0_{\pi_0}}[Y]$**:** Effect of π_0 on the target domain P^0.

Contribution

1. We develop the identification criterion for evaluating the target policy on the target domain from multiple *experimental datasets* from *different sources*.
2. We develop *doubly robust* estimator (DML-estimator) for estimating the policy intervention, and provide a finite sample learning guarantee.

Identification Criterion

Invariance of Y over domains
$(Y \perp\!\!\!\perp S \mid C_1, C_2, X_1, X_2, W_1)$ in $\mathcal{G}^{\Delta_{0,2}}_{\pi_0}$

Invariance of Y over policies
$(Y \perp\!\!\!\perp \pi_2 \mid C_1, C_2, X_1, X_2, W_1)$ in $\mathcal{G}_{\pi_2}$

Invariance of W_1 over domains
$(W_1 \perp\!\!\!\perp S \mid C_1, X_1)$ in $\mathcal{G}^{\Delta_{0,1}}_{\pi_0}$

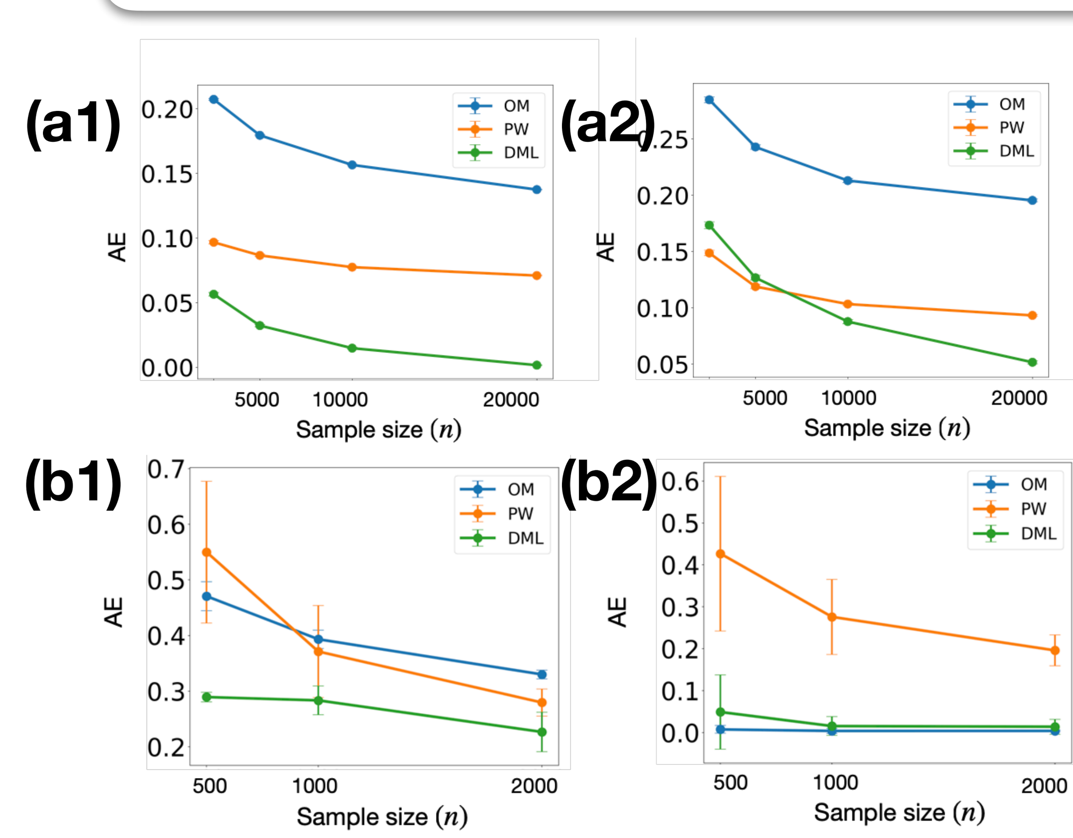
Invariance of W_1 over policies
$(W_1 \perp\!\!\!\perp \pi_1 \mid C_1, X_1)$ in $\mathcal{G}_{\pi_1}$

$$\mathbb{E}_{P^0_{\pi_0}}[Y] = \sum_{w,c_1,c_2,x_1,x_2} \mathbb{E}_{P^2_{\pi_2}}[Y \mid c_1, c_2, w_1, x_1, x_2]\pi_0(x_1 \mid c_1)P^1_{\pi_1}(w_1 \mid c_1, x_1)\pi_0(x_2 \mid c_1, c_2, x_1, w_1)P^0_{\pi_0}(c_1, c_2)$$

A target policy on a target domain is evaluable from multiple experiments from different source domains under these conditions.

DML-based Estimator

DML-estimator$:= \mathbb{E}_{D^2_{\pi_2}}[\hat{\omega}^2\{Y - \hat{\mu}^2\}] + \mathbb{E}_{D^1_{\pi_2}}[\check{\omega}^1\{\check{\mu}^2 - \hat{\mu}^1\}] + \mathbb{E}_{D^0}[\check{\mu}^1]$ ($\{\hat{\mu}^i, \hat{\omega}^i\}$ are nuisance estimates and $D^i_\pi \sim P^i_{\pi_i}$ are samples) satisfies *doubly robustness* (If $\hat{\mu}^i$, $\hat{\omega}^i$ converges to μ^i_0, ω^i_0 at $n^{-1/4}$ rate, the estimator converges at $n^{-1/2}$ rate).



- **(a)** Synthetic and **(b)** real-world (ACTG-175), which assessed therapies for reducing CD4 cell counts in HIV patients.
- **(1)** Noise-free (no noises are added to the estimated nuisance) and **(2)** Noisy environment — noises converging at $n^{-1/4}$ rate is added to witness the doubly robustness behavior.
- For all experiments, the proposed DML-estimator exhibits doubly robustness.