Pessimistic Data Integration for Policy Evaluation

Xiangkun Wu,Ting Li,Gholamali Aminian, Armin Behnamnia,
Hamid R.Rabieeand Chengchun Shi.

November 05, 2025

Outline

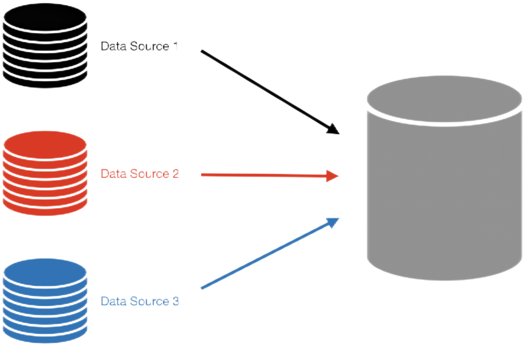
- Motivation
- The Proposed Method
- Theoretical Properties
- Simulation Studies

A/B testing

- A/B testing : A/B testing is a method to compare two methods and measure which performs better .
- Average Treatment Effect (ATE): measures the average difference in outcomes between the treatment group and the control group.

Motivation

- An accurate estimator for ATE is challenging due to the limited size.
- An abundance of data is dispersed across various data centers, presenting a notable opportunity for enhancement.
- This motivates us to capable of integrating data from multiple sources to enhance the accuracy and reliability of ATE.



Data Integration Example: A/B Testing

A/B testing with historical data

Experiment data

Policy **Control**

- limited duration
- weak treatment effect

Historical data

Control

- substantial volume

Data Integration Challenges

- Covariate shift:The covariate distribution of the experiment data is different from that of the historical data.
- Policy shift:The action (treatment allocation policy) differs between the experimental and historical control datasets.
- Posterior shift: The expected reward given the same state and action differs between the experimental and historical datasets.

The Proposed Method

- Experiment Data Only Estimator(EDO):

$$\text{EDO} = \mathbb{E}_n[\psi_1^{(e)}(O^{(e)})] - \mathbb{E}_n[\psi_0^{(e)}(O^{(e)})],$$

- Historical Data Based estimator(HDB)

$$\text{HDB} = \mathbb{E}_n[\psi_1^{(e)}(O^{(e)})] - \mathbb{E}_n[\psi^{(h)}(O^{(h)})].$$

- Adaptive weighted estimator:

$$\widehat{\text{ATE}}(w) = \mathbb{E}_n[\psi_1^{(e)}(O^{(e)})] - \mathbb{E}_n[\psi_w(O^{(e)}, O^{(h)})],$$

Here

$$\psi_w(O^{(e)}, O^{(h)}) = w(S^{(e)})\psi_0^{(e)}(O^{(e)}) + [1 - w(S^{(h)})]\psi^{(h)}(O^{(h)}).$$

$$\psi_a^{(e)}(O^{(e)}) = \mathbb{I}(A = a)R^{(e)}/\pi(a|S^{(e)}). \quad Importance \quad Sampling$$

$$\psi_0^{(h)}(O^{(h)}) = \mu(S^{(h)})R^{(h)}. \quad Density \quad Ratio$$

The Proposed Method

- Adaptively:Context-varing weight
- Treat weight as a policy
- Pessimistic: more robust to deal with heavy-tailed case and shifts.

Toy Example

Illustration of the Bias–Variance Tradeoff

We use a toy example to demonstrate the balance between **bias** and **variance**. The **EDO** method achieves the smallest bias, the **HDB** method shows the lowest variance, while the **Proposed** method strikes a better balance and achieves the smallest overall MSE.

Method	MSE (95% CI)	Bias (95% CI)	Variance (95% CI)
EDO	1.701 (1.598–1.804)	**0.007** (-0.064–0.051)	1.701 (1.598–1.804)
HDB	2.372 (2.289–2.455)	1.400 (1.372–1.428)	**0.413** (0.388–0.438)
Proposed	**1.394** (1.312–1.476)	0.221 (0.170–0.272)	1.345 (1.262–1.428)

All values are averaged over 100 runs with 95% confidence intervals.

Choice of Weight

- We describe the approach we use to determine the weight.

$$\widehat{\mathrm{MSE}_U}(w) = \widehat{\mathrm{bias}_U}^2(w) + \widehat{\mathrm{Var}_U}(w),$$

- $\mathbb{P}(\cap_{w \in \mathcal{W}} \{\widehat{\mathrm{bias}}_U(w) \geq |\mathrm{bias}(w)|\}) \geq 1 - \alpha$ and $\mathbb{P}(\cap_{w \in \mathcal{W}} \{\widehat{\mathrm{Var}}_U(w) \geq \mathrm{Var}(w)\}) \geq 1 - \alpha$
  More robust bias and variance estimator
- We solve for the optimal weight $\widehat{w}$ by minimizing the empirical uncertainty-aware objective:

$$\widehat{w} = \arg\min_{w} \widehat{\mathrm{MSE}}_U(w),$$

and then plug $\widehat{w}$ into the ATE formulation to obtain the final estimator. $\widehat{\mathrm{ATE}}(\widehat{w}) = \mathbb{E}_n[\psi_1^{(e)}(O^{(e)})] - \mathbb{E}_n[\psi_{\widehat{w}}(O^{(e)}, O^{(h)})]$

Theoretical Properties

Theorem 1 (MSE of the proposed estimator)

Under Assumptions 1 and 2, for any $w \in \mathcal{W}$,

$$\mathrm{MSE}\big(\widehat{\mathrm{ATE}}(\widehat{w})\big) - \mathrm{MSE}\big(\widehat{\mathrm{ATE}}(w)\big) \le \mathbb{E}\Big[\widehat{\mathrm{bias}}_U^{\,2}(w) - \mathrm{bias}^2(w)\Big] + \mathbb{E}\Big[\widehat{\mathrm{Var}}_U(w) - \mathrm{Var}(w)\Big] + O(\alpha B^2).$$

Theorem 1 is general in the sense that it holds for any OPE estimator, including direct, importance sampling (IS), and doubly robust (DR) approaches, when estimating the ATE.

The following corollaries demonstrate that our proposed estimator performs comparably to these optimal estimators across all scenarios, maintaining robustness with either heavy-tailed reward residuals or posterior shift.

Corollaries for Many Cases

Scenarios	EDO	HDB	MVE	CWE	NonPessi	Proposed
(i) Heavy-tailed historical rewards	Green	Red	Yellow	Yellow	Yellow	Green
(ii) Heavy-tailed experimental rewards	Red	Green	Yellow	Yellow	Yellow	Green
(iii)Small posterior shifts	Red	Yellow	Green	Yellow	Green	Green
(iv) Moderate posterior shifts	Green	Red	Red	Green	Red	Green
(v) Large posterior shifts	Green	Red	Red	Green	Green	Green

MSEs of different ATE estimators across five scenarios.
Green indicates that the estimator
achieves the oracle property (its MSE is asymptotically equivalent to that of the optimal estimator).
Yellow indicates that the estimator may generally have a high MSE but can attain the oracle property in some special cases.
Red indicates that the estimator exhibits a generally large MSE.

Synthetic-data Simulation

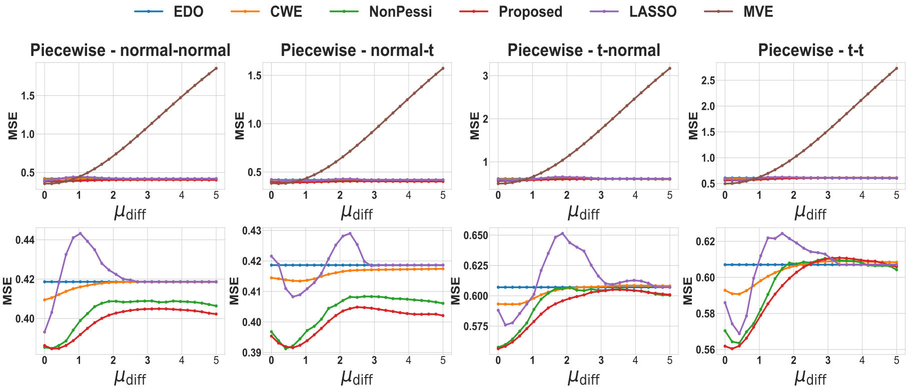


Figure 1: MSEs in Synthetic-data simulation

Top panels show all estimators under Piecewise shifts; bottom panels present results excluding MVE.

Ridesharing-data-based Simulation

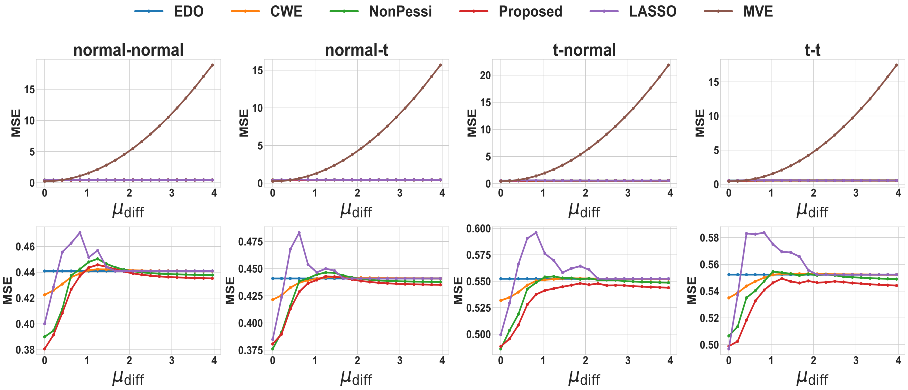


Figure 2: MSEs in Ridesharing-data-based simulation

Top panels show all estimators under different shifts; bottom panels present results excluding MVE.

Clinical-data-based Simulation

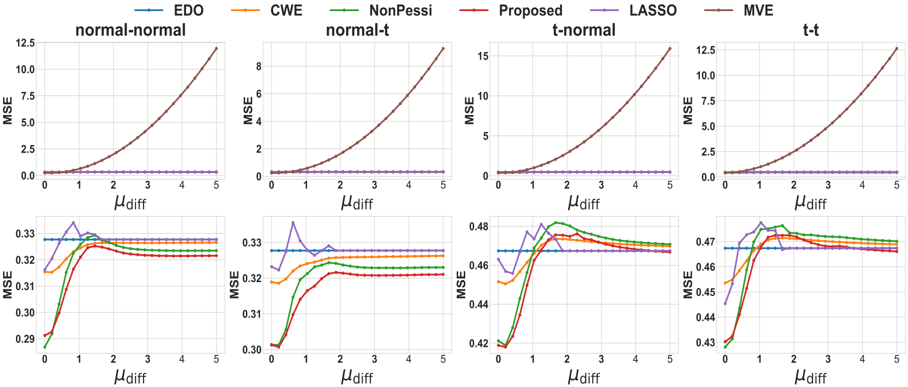


Figure 3: MSEs of ATE estimators in Clinical example. Top: all estimators; Bottom: exclude MVE.

Thank you!