Comparative Effectiveness of Machine Learning Methods for Causal Inference

ML methods for causal inference may find significant estimates that were not found before due to confoundedness or nonlinearity

Syed Badruddoza, Syed Fuad, Modhurima Dey Amin

Department of Agricultural and Applied Economics, Texas Tech University

Correspondence: Syed.M.Fuad@ttu.edu

RESEARCH QUESTIONS

- How do machine learning-based causal models compare in terms of bias in the estimation of the average treatment effect (ATE)?
- Can machine learning-based methods provide further insights into the treatment effects in presence of confoundedness or nonlinearity?

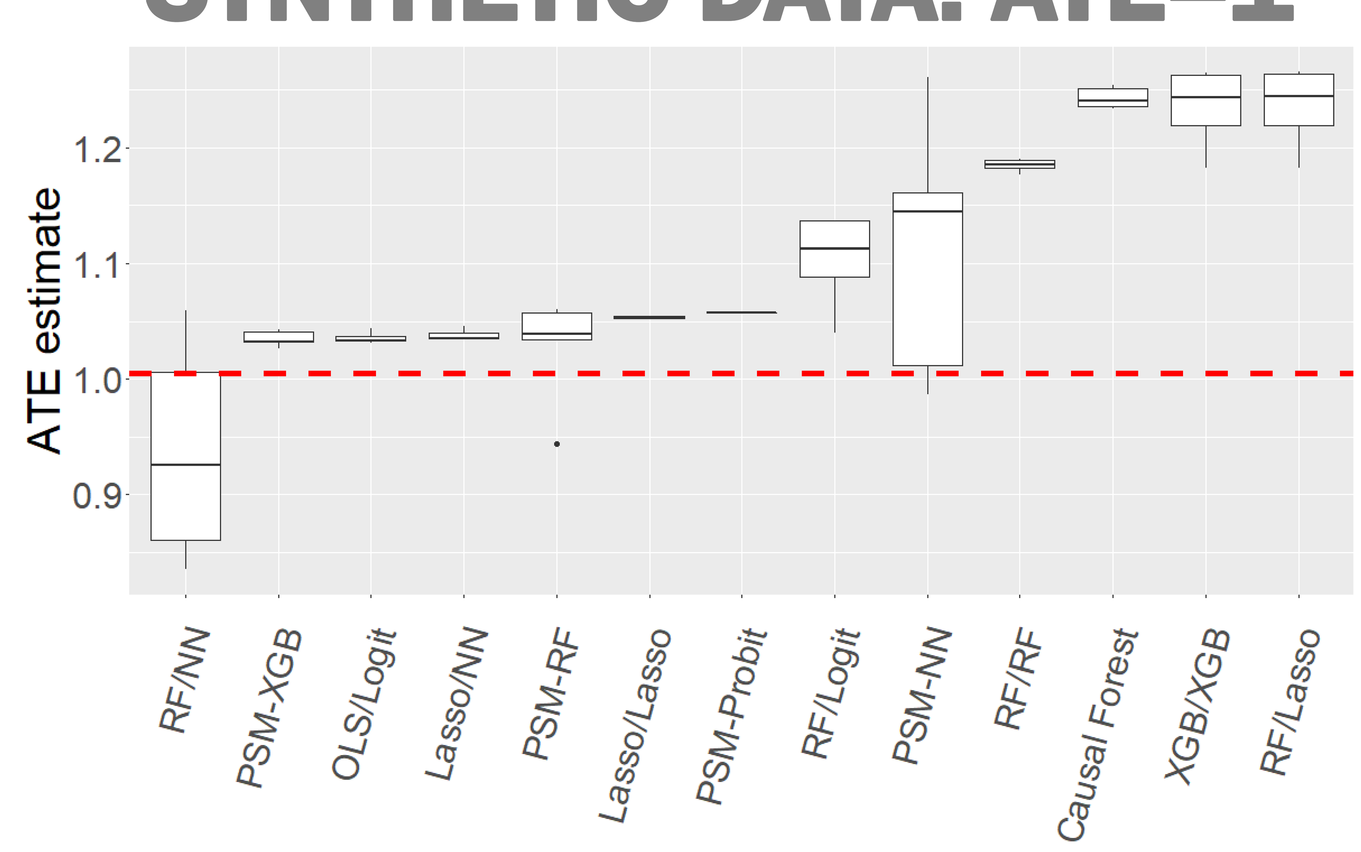
MOTIVATION

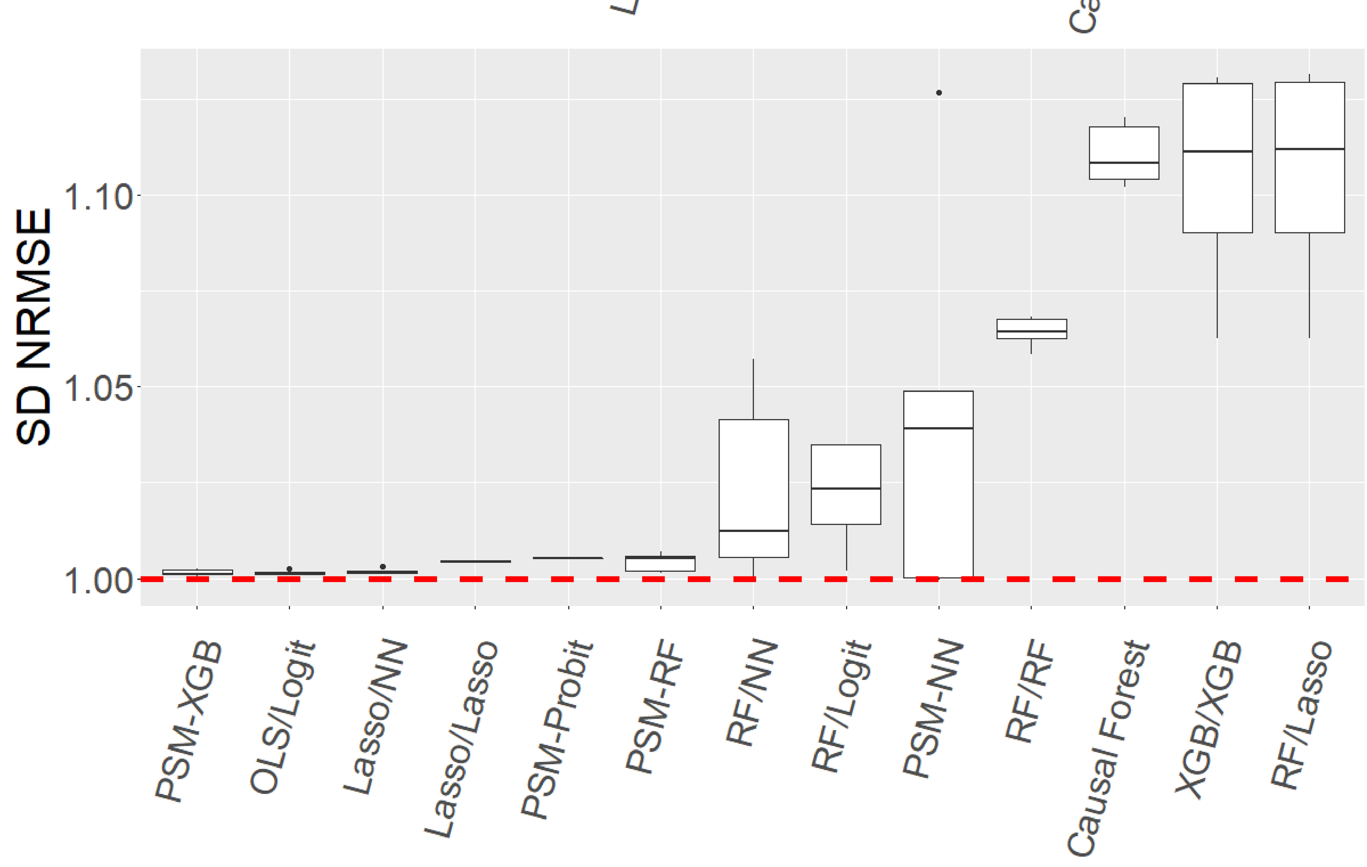
- Traditional regression techniques may produce biased estimates due to dimensionality, heterogeneous treatment effects, and functional form misspecification.
- In big data sets, confoundedness is more likely.
- Three popular ML approaches to address these issues are:
  - Double machine learning or DML (Chernozhukov et al. 2018) utilizes the non-inversion feature to learn the ATE & ATT under unconfoundedness. DML also corrects for bias through orthogonalization and cross-validation.
  - Causal forest or CF (Wager and Athey 2018) utilizes ML's classification power to maximize difference across recursively generated data partitions in the relationship between outcome and treatment, thus uncovering heterogeneity in causal effect.
  - A matching method, e.g., using ML for propensity matching (Ikezawa et al. 2022) to minimize the distance between treatment and control groups.
- Many studies have adopted these methods for more flexible and informative causal analysis. However, no study so far has compared the effectiveness of these models.

DATA & METHODS

- We study and compare three approaches: (1) causal forest (CF), (2) double machine learning (DML), and (3) propensity score matching with ML (PSM-ML) for causal analysis.
- Simulation part: Synthetic data generation with a continuous dependent variable and 10 continuous independent variables and an exogenous binary treatment.
- Empirical part: Data from (Dhurandhar et al. 2014) to obtain the impact of breakfast recommendations on weight loss. Data from (Bryan et al., 2014) to obtain the impact of cash transfer to promote out-migration on seasonal food security.

SYNTHETIC DATA: ATE=1





KEY RESULTS

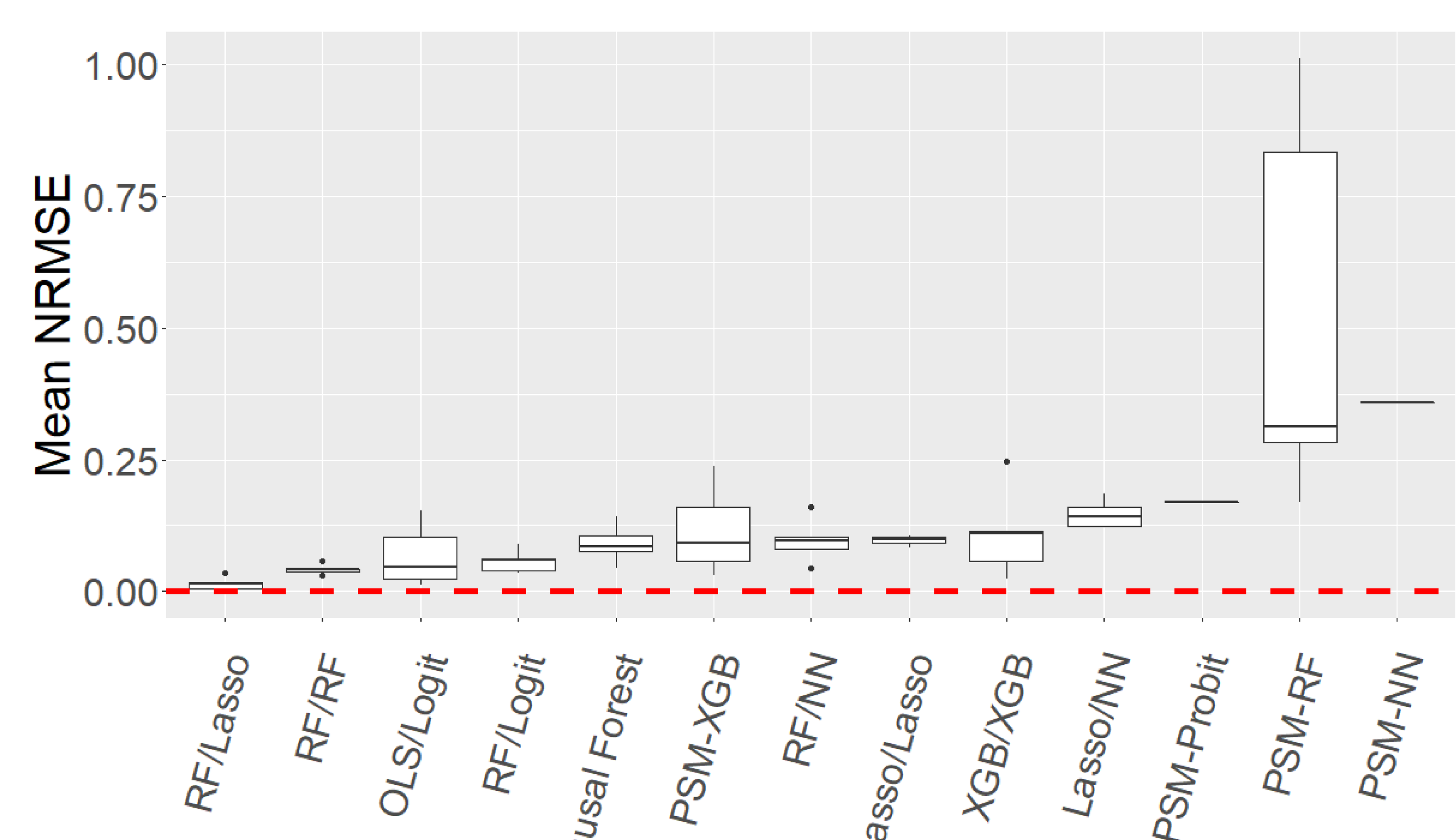
- Using PSM-ML is only advantageous in smaller data sets, especially under linear specifications, followed by DML and CF.
- DML has a relatively lower rate of error and exhibits better estimation accuracy as the data dimension grows.
- Results indicate that dimension reduction and a more flexible functional form with ML might help find the causal estimate previously undetected by simple regressions.
- For instance, we found a significant association between breakfast recommendations and weight loss for Dhurandhar et al. (2014) using Lasso/Lasso approach of DML.

DHURANDHAR ET AL: ATE=NULL

	RF/ Logit	RF/ NN	RF/ Lasso	RF/ RF	Lasso/ NN	XGB/ XGB	OLS/ Logit	Lasso/ Lasso
ATE	-0.26	-0.16	-0.20	-0.11	-0.23	-0.20	-0.26	-0.12
(SD)	(0.44)	(0.45)	(0.44)	(0.45)	(0.46)	(0.40)	(0.44)	(0.04)

	Causal Forest	PSM-Probit	PSM-NN	PSM-RF	PSM-XGB
ATE	-0.09	-0.27	0.15	-0.08	-0.18
(SD)	(0.47)	(0.49)	(0.62)	(0.51)	(0.44)

BRYAN ET AL: ATE=44.43



REMARKS

- Online marketing and weather-related data sets are often high dimensional and involve variables related to product features, consumer behavior, market, climate, etc. that are closely related, and controlling for one's movement implicitly controls for others due to parametric assumptions.
- We show that the ML approach may offer a viable alternative to traditional ATE regression for its flexible, data-driven nature, predictive accuracy, and dimension reduction capabilities.

Works cited:

1. Bryan, G., Chowdhury, S. and Mobarak, A.M., 2014. Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. Econometrica, 82(5), pp.1671-1748.
2. Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters.
3. Dhurandhar, E.J. et al., 2014. The effectiveness of breakfast recommendations on weight loss: a randomized controlled trial. The American Journal of Clinical Nutrition, 100(2), pp.507-513.
4. Ikezawa, K., Hirose, M., Maruyama, T., Yuji, K., Yabe, Y., Kanamori, T., Kaide, N., Tsuchiya, Y., Hara, S. and Suzuki, H., 2022. Effect of early nutritional initiation on post-cerebral infarction discharge destination: A propensity-matched analysis using machine learning. *Nutrition & Dietetics*, *79*(2), pp.247-254.
5. Wager, S. and Athey, S., 2018. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), pp.1228-1242.