CAUSAL INFERENCE: BAYESIAN CAUSAL FOREST(BCF), BAYESIAN ADDITIVE REGRESSION TREES(BART) AND GLM WITH HORSESHOE

Authors: Kejing Yan, Ruoyun Yang, Qingyang Cheng, Zhuotong Sheng, Qile Jiang

Affiliation: Brown Data Science Institute

Github Repo: https://github.com/cynthiayry/2020-Final-Project.git

1 Introduction

- This project aims to **estimate heterogeneous treatment effects** (**HTEs**) **under confounding**, using **Bayesian Causal Forests** (**BCF**), **Bayesian Additive Regression Trees** (**BART**), and a **Generalized Linear Model** (**GLM**) **with a horseshoe prior**. We compare the three models in terms of accuracy, uncertainty quantification, and robustness across different data-generating settings.
- Building on Hahn et al. (2020), we use BCF's **decomposition of treatment and prognostic effects** and its integration of the estimated propensity score to improve HTE estimation. BART provides a flexible nonparametric baseline, while the GLM with a horseshoe prior offers a sparse, interpretable alternative.
- We **simulate data** following Section 6.1 of Hahn et al. (2020), varying sample size, treatment effect heterogeneity, and confounding strength. Model performance is evaluated using RMSE, confidence interval width, and empirical coverage of 95% CIs to assess both accuracy and uncertainty.

2 Methods

- **Simulation Design**(**Monte Carlo Repetitions**: 50 runs for 8 design settings (2×2×2 for tau/mu/n):
  - Covariates x1–x5 generated (3 continuous, 1 binary, 1 categorical).
  - Prognostic function μ(x)\mu(x)μ(x): linear or nonlinear.
  - Treatment effect τ(x)\tau(x)τ(x): homogeneous or heterogeneous.
  - Propensity score π(x)\pi(x)π(x) estimated via logistic regression.
  - Outcome model: yi=μ(xi)+τ(xi)⋅zi+ϵi, ϵi N(0,σ2)
- **BCF Implementation**:
  - Inputs: outcome y, treatment z, covariates X, estimated propensity score π^(x)\hat{\pi}(x)π^(x); Uses separate trees for baseline μ\muμ and treatment τ\tauτ
- **BART** (**Bayesian Additive Regression Trees**):
  - Models y f(x,z)y \sim f(x, z)y f(x,z) with no separation between treatment and baseline effects; Posterior treatment effects are inferred by comparing predictions under z=1 vs. z=0.
- **GLM with Horseshoe Prior**:
  - Used to model treatment effect in a high-dimensional sparse linear setting.Horseshoe prior encourages sparsity, shrinking irrelevant covariates.
- **Evaluation Metrics**:
  - Root Mean Squared Error (RMSE) of CATE and ATE; 95% coverage and average credible interval length

3 Simulations & Results

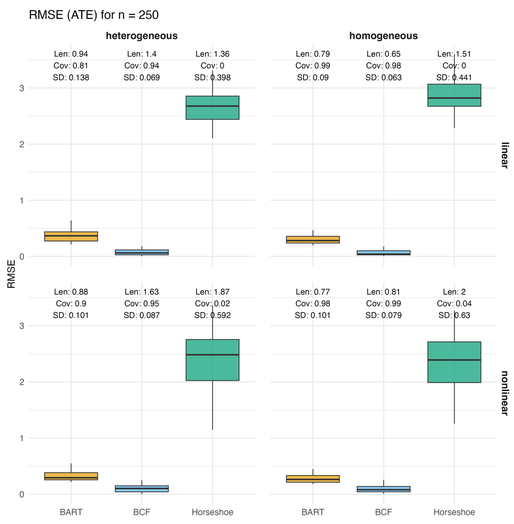

Figure1. Boxplot of RMSE for ATE Estimates Across Settings and Models(n = 250)

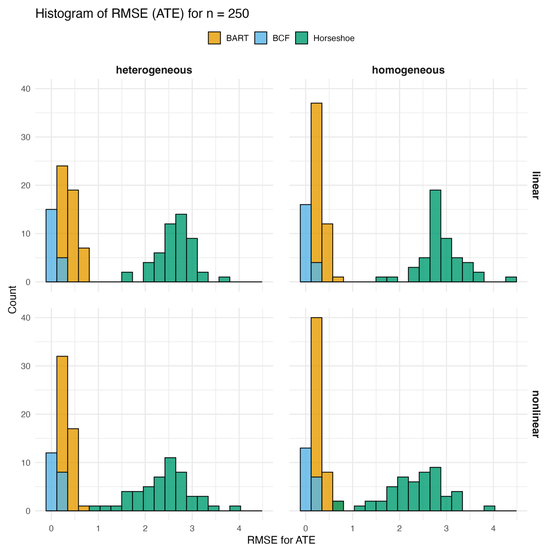

Figure3. Histogram of RMSE for ATE Estimates Across Settings and Models (n = 250)

- **Model Comparison**
  - BCF achieves the lowest RMSE and highest coverage across nearly all settings.
    - Produces generally narrower confidence intervals while maintaining proper uncertainty estimates.
    - Especially effective under strong confounding and complex relationships.
  - BART shows solid performance but tends to:
    - Slightly under-cover the true ATE.
    - Produce wider intervals and more variability in RMSE.
  - GLM with Horseshoe prior underperforms significantly:
    - High RMSE and very low coverage in nonlinear or heterogeneous scenarios.
    - Performs acceptably only in simple, linear settings—highlighting its rigidity.
- **Effect of Sample Size**:
  - All models benefit from larger sample size (n = 500 vs. n = 250).
  - BCF shows the greatest improvement—lower RMSE, reduced SD, and tighter intervals.
  - BART improves modestly; Horseshoe remains inaccurate despite more data.
- **Summary of BCF Across Settings**:
  - Performs best in linear + homogeneous settings, as expected.
  - Maintains competitive accuracy even in nonlinear + heterogeneous scenarios.
- **Overall Takeaway**:
  - BCF is the most robust and well-calibrated method for estimating treatment effects under confounding.
  - It balances flexibility, accuracy, and reliable uncertainty—making it the strongest overall choice.

Figure2. Boxplot of RMSE for ATE Estimates Across Settings and Models (n = 500)

Figure4. Histogram of RMSE for ATE Estimates Across Settings and Models (n = 500)

Figure5. Boxplot of RMSE for ATE Estimates Across Settings for BCF

sample size, prognostic, treatment	X_1	X_2	X_3	X_4	X_5
$n = 250$, homogeneous, linear	0.00	0.02	0.02	0.78	1.00
$n = 250$, homogeneous, nonlinear	0.02	0.02	0.38	0.36	1.00
$n = 250$, heterogeneous, linear	0.02	0.74	0.00	0.54	1.00
$n = 250$, heterogeneous, nonlinear	0.04	0.46	0.40	0.18	1.00
$n = 500$, homogeneous, linear	0.02	0.04	0.00	0.92	1.00
$n = 500$, homogeneous, nonlinear	0.02	0.04	0.52	0.82	1.00
$n = 500$, heterogeneous, linear	0.00	0.98	0.02	0.74	1.00
$n = 500$, heterogeneous, nonlinear	0.04	0.84	0.60	0.54	1.00

Table1. Proportion of Significant Coefficients in Horseshoe Model

Feature Importance

- The coefficient of X_5 is significant in all settings, while that of X_1 is hardly significant.
- In general, a larger sample size yields stronger variable effects.
- Variable X_2 is more important for heterogeneous treatment than homogeneous.
- Variable X_3 is more important for nonlinear prognostic function than linear.

4 Conclusions

Our results demonstrate that BCF provides the most accurate and stable estimates of treatment effects across simulation settings, consistently outperforming BART and the GLM with a horseshoe prior. This supports the findings of Hahn et al. (2020), highlighting the effectiveness of Bayesian models that explicitly separate treatment and prognostic effects and incorporate estimated propensity scores.

By improving uncertainty quantification and maintaining high coverage even under strong confounding, BCF proves especially well-suited for heterogeneous and nonlinear treatment effect settings. Its performance strengthens the case for Bayesian causal forests as a robust tool for individual-level treatment effect estimation, particularly when traditional parametric models like GLMs fall short.

5 Contributions & References

Qile: BART modeling
Qingyang: Horseshoe modeling
Zhuotong: BART modeling
Ruoyun: BCF modeling, poster, plots
Kejing: BCF modeling,poster

REF:
Hahn, P. R., Murray, J. S., & Carvalho, C. M. (2020). Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects. Bayesian Analysis, 15(3), 965–1056. https://doi.org/10.1214/19-BA1175