Sparsity-inducing Bayesian Causal Forest with Instrumental Variable

Jens Klenke[1,$\bar{\wedge}$], Lennard Maßmann[1,2,*]

[1]Chair of Econometrics, University of Duisburg-Essen
[2]Ruhr Graduate School in Economics

[*]E-mail: lennard.massmann@uni-due.de
[$\bar{\wedge}$]E-mail: jens.klenke@vwl.uni-due.de

September 26, 2024

Abstract

Your text:Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi.

Contents

1 Introduction

?

1.1 Literature

- Using machine learning to infer heterogeneous effects in observational studies often focuses on CATE estimation under regular assignment mechanisms
- Focus in this work: Methods to discover heterogeneous effects in the presence of imperfect compliance
- Methods
  - tree-based
  - ensemble-of-trees
  - deep-learning-based methods
- BCF-IV: Discovers and estimates HTE in an interpretable way
  - BCF: BART-based semi-parametric Bayesian regression model, able to estimate HTE in regular assignment mechanisms, even with strong confounding
  - Use BCF to estimate $\hat{\tau}_C(x)$ and $\widehat{ITT}_Y(x)$ such that the conditional Complier Average Causal Effect $\hat{\tau}^{cace}(x) = \frac{\widehat{ITT}_Y(x)}{\hat{\tau}_C(x)}$
- - BART benefits in general
  - good performance in high-noise settings
  - shrinkage to/emphasize on low-order interactions
  - established software implementations ('BayesTree', 'bartMachine', 'dbarts')
- BART shortcomings
  - non-smooth predictions as BART prior produces stepwise-continuous functions
  - BART prior is overconfident in regions with weak common support

- Research proposal: Rewrite the BCF-IV model with SoftBART instead of BART prior to account for sparsity

1.2 Contribution

We contribute to the literature on machine learning techniques designed to uncover heterogeneous effects while addressing imperfect compliance by developing adaptations of tree-based algorithms. Most of the recent scientific contributions in this area revolve around estimating the Conditional Average Treatment Effect (CATE) by combining machine learning algorithms with the potential outcome framework under a regular assignment mechanism (litertaure overview here). In this work, we focus on the estimation of the conditional Complier Average Causal Effect (cCACE) which is the treatment effect on the subgroup of compliers under the usual four assumptions within an IV framework. Specifically, we generalize Bayesian Instrumental Variable Causal Forest (BCF-IV) from (**?**) to estimate cCACE precisely when there are many covariates. BCF-IV is a semi-parametric Bayesian regression model that builds directly on the Bayesian Additive Regression Trees (BART) algorithm (Chipman et al., 2010). Instead of using BART fur pure predictions of outcomes, BCF-IV is designed to identify and estimate heterogeneous effects within the subpopulation of units that comply with the treatment assignment, known as compliers. Consequently, the estimated effects can be considered doubly local, representing subgroup effects within the compliers subpopulation. BCF-IV identifies heterogeneity through an interpretable tree structure, with each node representing a distinct subgroup. For each leaf node of the generated tree, BCF-IV supports multiple hypothesis test adjustments to control the Type I error rate (familywise error rate) or the false discovery rate.

In our work, we extend BCF-IV in several ways. First, we use the shrinkage prior adaptation of BART (SoftBART) as proposed in (**??**). More precisely, by dividing the conditional Intention-To-Treat (cITT) effects with the corresponding conditional Proportion of Compliers (cPC), one can show to arrive at the conditional Complier Average Causal Effect (cCACE). In the discovery step the original BCF-IV algorithm of (**?**) uses the Bayesian Causal Forest (BCF) (**?**) to estimate the cITT effects. BCF is a nonlinear regression model that builds upon BART and is proposed for estimating heterogeneous treatment effects. It is specifically designed for scenarios characterized by small effect sizes, heterogeneous effects, and significant confounding by observables. First, BCF addresses the issue of highly biased treatment effect estimates in the presence of strong confounding by incorporating an estimate of the propensity function directly into the response model. Thereby, it induces a covariate-dependent prior on the regression function. Second, BCF allows for the separate regularization of treatment effect heterogeneity from the prognostic effect

of control variables. Conventional response surface modeling approaches often fail to adequately model regularization over effect heterogeneity. Instead, BCF enables an informative shrinkage towards homogeneity such that one is able to control the degree of regularization over effect heterogeneity. In our work, we generalize the cITT effect estimation by replacing BCF with the Shrinkage Bayesian Causal Forest (SBCF) proposed by (**?**). SBCF extends BCF by using additional priors proposed in SoftBART that enable to adjust the influence of each covariate based on the number of corresponding splits in the tree ensemble. These priors enhance the model's adaptability to sparse data-generating processes and facilitate fully Bayesian feature shrinkage within the framework for estimating treatment effects. Consequently, it improves to uncover the moderating factors that drive heterogeneity when there is sparsity in the data. Moreover, this method allows the incorporation of prior knowledge regarding relevant confounding covariates and the relative magnitude of their impact on the outcome.

Second, BCF-IV uses a single binary tree based on the CART model of (Breiman1984) to analyse possible heterogeneity patterns within cCACE. In our work, we use posterior splitting probabilities retrieved from SBCF as a measure of variable importance within the `cost`-argument of `rpart`. These are scalings to be applied when considering splits, so the improvement on splitting on a variable is divided by its cost in deciding which split to choose in `rpart`.

Open questions for further contributions. Use Random Forests (Causal Rule Ensemble) instead of single CART for subgroup discovery? More rigorous Bayesian estimation by replacing `ivreg` with `brms` (implementation of credible intervals, estimation error, counterparts to something like weak instrument tests, bayesian model averaging)?

2 Potential outcomes and irregular assignmnet

We follow the setup of **?** and **?** who follow the usual notation of the Rubin's causal model.

Given a set of N indiviudals, indexed by $i = 1, \ldots, N$, we denote with Y_i a generic outcome variable, with W_i a binary treatment indicator, with X an $N \times P$ matrix of P control variables, and with X_i the i-th P-dimensional row vector of covariates. Let us define the pair of potential outcomes $Y_i(W_i)$ by using the Stable Unit Treatment Value Assumption (SUTVA).

Assumption 2.1. Stable unit treatment value assumption (SUTVA).

$$\text{If } W_i = w, \text{ then } Y_i(w) = Y_i^{obs} \ , \forall w \in \{0, 1\} \ , \forall i \in \{1, ..., N\}.$$

Therefore, it holds that $Y_i(W_i = 1) = Y_i(1)$ for an indiviudal i under assignment to the treatment group and $Y_i(W_i = 0) = Y_i(0)$ under assignment to control group. Despite one is unable to observe both potential outcomes at the same time for each individual one is able to observe the potential outcome that aligns with the assigned treatment status such that

$$Y_i^{\text{obs}} = Y_i(1)W_i + Y_i(0)(1 - W_i).$$

Under the strong ignorability assumptions of 2.2 and 2.3 we operate under the regular assignment mechanism which, through 2.2, prevents the existence of unmeasured confounding and, by 2.3, allows for unbiased treatment effect estimation in support of the covariate space.

Assumption 2.2. Unconfoundedness.

$$\begin{aligned} &W_i \perp\!\!\!\perp (Y_i(1), Y_i(0)) \, | X_i), \text{ or equivalently,} \\ &Pr(W_i | (Y_i(1), Y_i(0), X_i) = Pr(W_i | X_i). \end{aligned}$$

Assumption 2.3. Positivity.

$$\epsilon < p(X_i = x) < 1 - \epsilon \text{ with probability 1, } \forall \epsilon > 0, \forall x \text{ in support of } X_i.$$

Using strong ignorability, one can define the CATE to analyse heterogeneous treatment effects by

Definition 2.1. Conditional Average Treatmnet Effect (CATE).

$$\tau(x) = \mathbb{E}[Y_i^{\text{obs}} \mid W_i = 1, X_i = x] - \mathbb{E}[Y_i^{\text{obs}} \mid W_i = 0, X_i = x].$$

In this work, we deviate from this regular assignment mechanism that is implied by assumptions 2.2 and 2.3 and operate under the scenario of an irregular assignment mechanism. This irregular assignment mechanism allows for a violation of compliance between (quasi-)randomized treatment assignment, Z_i, and treatment receipt, W_i, such that just the assignment to treatment is assumed to be unconfounded through a valid choice of Z_i but the treatment receipt might be confounded. A remedy to the issue of confoundedness in treatment receipt is the Instrumental Variable (IV) approach. A proper instrumental variable, Z_i, affects treatment receipt, W_i, while not being allowed to affect Y_i directly such that treatment receipt depends on the treatment assignment by $W_i(Z_i)$. Based on the functional relation between W_i and Z_i and in case of a binary treatment variable, one can define four subgroups of individuals, G_i, such that

Definition 2.2. Subgroups G_i.

$$G_i = \begin{cases} C, & \text{if } W_i(Z_i = 0) = 0, W_i(Z_i = 1) = 1 \\ D, & \text{if } W_i(Z_i = 0) = 1, W_i(Z_i = 1) = 0 \\ AT, & \text{if } W_i(Z_i = 0) = 1, W_i(Z_i = 1) = 1 \\ NT, & \text{if } W_i(Z_i = 0) = 0, W_i(Z_i = 1) = 0 \end{cases}.$$

where C, D, AT and NT are abbreviations for Compliers, Defiers, Always-Takers, Never-Takers. The proportion of individuals that belong to each subgroup is defined as π_{G_i}, i.e. the proportion of compliers reads π_C. Considering the distinction between Z_i and W_i, the Intention-To-Treat (ITT) effect is defined as

Definition 2.3. Intention-To-Treat (ITT) effect.

$$ITT_Y = \mathbb{E}\left[Y_i | Z_i = 1\right] - \mathbb{E}\left[Y_i | Z_i = 0\right],$$

which refers to the instrument's average effect. Based on the subgroups in 2.2 and the proportions π_{G_i}, an IV setting can be formalized by the usual four assumptions (**??**, **??**, **??**, **??**) following **?**:

Assumption 2.4. Classical IV assumptions with a binary treatment.

Exclusion restriction:	$Y_i(0) = Y_i(1),\ \text{for } G_i \in \{AT, NT\}.$
Monotonicity:	$W_i(1) \geq W_i(0) \rightarrow \pi_D = 0.$
Existence of compliers:	$P(W_i(0) < W_i(1)) > 0 \rightarrow \pi_C \neq 0.$
Unconfoundedness of IV:	$Z_i \perp (Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1), W_i(0), W_i(1)).$

If these assumptions **??** - **??** hold, the Complier Average Causal Effect (CACE) is identified as

Definition 2.4. Complier Average Causal Effect (CACE).

$$\tau_{\text{CACE}} = \frac{\text{ITT}_Y}{\pi_C} = \frac{\mathbb{E}[Y_i \mid Z_i = 1] - \mathbb{E}[Y_i \mid Z_i = 0]}{\mathbb{E}[W_i \mid Z_i = 1] - \mathbb{E}[W_i \mid Z_i = 0]},$$

and can be estimated from observational data. The numerator represents the average effect of the instrument, also referred to as the Intention-To-Treat (ITT) effect. The denominator represents the overall proportion of indiviudals that comply with the treatment assignment, also referred to as the proportion of compliers **?**. CACE is also sometimes referred to as Local Average Treatment Effects (see **?**) and represents the estimate of the causal effect of the assignment to treatment on the principal outcome, Y_i, for the subpopulation of compliers **?**. In this work, we follow **?** and consider the conditional version of the CACE,

Definition 2.5. Conditional CACE (cCACE).

$$\tau_{\text{CACE}}(x) = \frac{\text{ITT}_Y(x)}{\pi_C(x)} = \frac{\mathbb{E}[Y_i \mid Z_i = 1, X_i = x] - \mathbb{E}[Y_i \mid Z_i = 0, X_i = x]}{\mathbb{E}[W_i \mid Z_i = 1, X_i = x] - \mathbb{E}[W_i \mid Z_i = 0, X_i = x]}.$$

The cCACE is a straightforward extension of the CACE in 2.4 with **?** providing theorems of a consistent and asymptotic normal conditional Two-Stage-Least-Square (2SLS) estimator which rely on the unconditional theorems of Wooldridge with the additional assumption of sufficient number of observations for every subgroup in which the cCACE is estimated.

3 Sparsity-inducing Bayesian Causal Forest with Instrumental Variable

3.1 Sparse Bayesian Causal Forest

3.2 Sparse BCF-IV

4 Simulation study

Performance criteria according to bargagli-stoffi:

1. Average number of truly discovered heterogeneous subgroups corresponding to the nodes of the generated CART (proportion of correctly discovered subgroups);

2. Monte Carlo estimated bias for the heterogeneous subgroups:

$$\text{Bias}_m(I_{\text{inf}}) = \frac{1}{N_{\text{inf}}} \sum_{i=1}^{N_{\text{inf}}} \sum_{l=1}^{L} \left(\tau_{\text{cace},i}(\ell) - \hat{\tau}_{\text{cace},i}(\ell, \Pi_m, I_{\text{inf}}) \right), \tag{4.0.1}$$

$$\text{Bias}(I_{\text{inf}}) = \frac{1}{M} \sum_{m=1}^{M} \text{Bias}_m(I_{\text{inf}}), \tag{4.0.2}$$

where Π_m is the partition selected in simulation m, L is the number of subgroups with heterogeneous effects (i.e., two for the case of strong heterogeneity and four for the case of slight heterogeneity), and N_{inf} is the number of observations in the inference sample.

3. Monte Carlo estimated MSE for the heterogeneous subgroups:

$$\text{MSE}_m(I_{\text{inf}}) = \frac{1}{N_{\text{inf}}} \sum_{i=1}^{N_{\text{inf}}} \sum_{l=1}^{L} \left(\tau_{\text{cace},i}(\ell) - \hat{\tau}_{\text{cace},i}(\ell, \Pi_m, I_{\text{inf}}) \right)^2, \tag{4.0.3}$$

$$\text{MSE}(I_{\text{inf}}) = \frac{1}{M} \sum_{m=1}^{M} \text{MSE}_m(I_{\text{inf}}); \tag{4.0.4}$$

4. Monte Carlo coverage, computed as the average proportion of units for which the estimated 95% confidence interval of the causal effect in the assigned leaf includes the true value, for the heterogeneous subgroups:

$$C_m(I_{\text{inf}}) = \frac{1}{N_{\text{inf}}} \sum_{i=1}^{N_{\text{inf}}} \sum_{l=1}^{L} \left(\tau_{\text{cace},i}(\ell) \in \hat{\text{CI}}_{95} \left(\hat{\tau}_{\text{cace},i}(\ell, \Pi_m, I_{\text{inf}}) \right) \right), \tag{4.0.5}$$

$$C(I_{\text{inf}}) = \frac{1}{M} \sum_{m=1}^{M} C_m(I_{\text{inf}}). \tag{4.0.6}$$

5 Empirical application

6 Discussion