BAYESIAN INDEX MODELS FOR HETEROGENEOUS TREATMENT EFFECTS ON A BINARY OUTCOME

Hyung Park, Danni Wu, Eva Petkova, Thaddeus Tarpey
Department of Population Health
New York University, New York, NY 10016
parkh15@nyu.edu

R. Todd Ogden
Department of Biostatistics
Columbia University
New York, NY 10032

ABSTRACT

This paper develops a Bayesian model with a flexible link function connecting a binary treatment response to a linear combination of covariates, and a treatment indicator and the interaction between the two. Generalized linear models allowing data-driven link functions are often called "single-index models" and are among popular semi-parametric modeling methods. In this paper, we focus on modeling heterogeneous treatment effects, with the goal of developing a treatment benefit index (TBI) incorporating prior information from historical data. The model makes inference on a composite moderator of treatment effects, summarizing the effect of the predictors within a single variable through a linear projection of the predictors. This treatment benefit index can be useful for stratifying patients according to their predicted treatment benefit levels and can be especially useful for precision health applications. The proposed method is applied to a COVID-19 treatment study.

Keywords Bayesian single index models · Heterogeneous treatment effects · Precision medicine

1 Introduction

In precision medicine, a critical concern is to characterize individuals' heterogeneity in treatment responses in order to enable individual-specific treatment decisions to be made [Murphy, 2003, Robins, 2004, Collins and Varmus, 2015]. Tailoring medical treatments according to individuals' characteristics requires inferring individual-level treatment effects (ITE) (as opposed to inferring the treatment effects on average across the entire population). Then, developing of an individualized treatment rule (ITR) [see, e.g., Qian and Murphy, 2011, Lu et al., 2011, Tian et al., 2014, Shi et al., 2016, Jeng et al., 2018, Zhao et al., 2012, 2015, Song et al., 2015, Laber and Zhao, 2015, Laber and Staicu, 2018, Zhao et al., 2019] naturally follows from drawing inferences about the ITE. As an alternative route to developing an ITR, direct optimization of the population average over a class of ITRs [see, e.g., Zhao et al., 2012, 2015, Liu et al., 2018, Xie et al., 2022] may be considered.

Substantial developments have been made in the statistical methodology on the ITE estimation [see, e.g., review provided in Caron et al., 2022]. Examples within the frequentist paradigm include R-learner [Nie and Wager, 2021], kernel-based multi-task learner [Alaa and van der Schaar, 2017], neural-networks [Shalit et al., 2017], tree-based ensembles [Wager and Athey, 2018] and semi-parametric regression [Guo et al., 2022], among many others. In the Bayesian paradigm, one prominent approach is the causal Bayesian additive regression trees (BART) [see, e.g., Hahn et al., 2020, Klausch et al., 2018, Murray et al., 2018, Logan et al., 2019] among others. Although there are many statistical learning methods that exhibit good performance in capturing complex nonlinear relationship for individual treatment effects, in this paper, we focus on the Bayesian estimation of single-index regression models for ITE, as there has been no specific work designed to model the heterogeneous treatment effects using a single-index model in the Bayesian paradigm, albeit the usefulness that arises from the simplicity in its model formulation.

A single-index model [see, e.g., Brillinger, 1982, Stoker, 1986] is one of the most popular semi-parametric models and provides an efficient way of dealing with multivariate nonparametric regression. Single-index models expand the scope of generalized linear models through a flexible data-driven link function. The model summarizes the effect of the predictors within a single variable through a linear projection of the predictors, called the (single-)index. In the context

of modeling heterogeneous treatment effects, such an index corresponds to a composite moderator of treatment effects, and we will demonstrate that such an index is useful for optimizing treatment decisions and the approach provides a natural way to summarize uncertainties associated with this composite moderator, through its posterior distribution.

There are broadly two lines of research on Bayesian single-index models. One approach employs a spline-based representation of the link [Antoniadis et al., 2004, Poon and Wang, 2013, 2014, Wang, 2009, 2018]. The other line of research employs a Gaussian process-type representation of the link [Choi et al., 2011, Gramacy and Lian, 2012, Yu et al., 2019, Dhara et al., 2020, Liu and Liang, 2022], where the unknown link function is assumed to be a Gaussian process a priori. In this article, we take the former spline-based approach as it allows us to easily incorporate an identifiability constraint on the link function used to model the heterogeneous treatment effect term.

In this paper, we consider a treatment variable A taking a value in $\{0, 1\}$ with the associated randomization probabilities $\{\pi_0, \pi_1\}$ in the context of randomized clinical trials (RCTs), with the corresponding potential outcomes denoted as $\{Y^{(0)}, Y^{(1)}\}$. Depending on A, the observed outcome is $Y = (1 - A)Y^{(0)} + AY^{(1)}$, where the outcome Y is assumed to be a member of the exponential family. Specifically, we focus on a binary outcome $Y \in \{0, 1\}$, where we assume, without loss of generality, that the value of $Y = 0$ is desired so that $Y = 1$ indicates a bad outcome (e.g., death). On the population level, this means that a small value of $h(E[Y])$ is desired, where $h(\cdot)$ denotes the canonical link of the assumed exponential family distribution. For the binary outcome considered here, $h(\cdot)$ is the logit function.

The proposed single-index approach to modeling heterogeneous treatment effects was motivated by an application to a COVID-19 convalescent plasma (CCP) treatment study [Troxel et al., 2022]. One of the primary goals of this RCT was to guide CCP treatment recommendations by providing an estimate of a differential treatment outcome when a patient is treated with CCP vs. without CCP, where a larger differential in favor of CCP would indicate a more compelling reason for recommending CCP. The estimated index, defined as a linear combination of covariates $\boldsymbol{X}$, which is part of the heterogeneous treatment effect term of the model, can be used to discover profiles of patients with COVID-19 associated with different benefit from CCP treatment. Specifically, the covariates $\boldsymbol{X} \in \mathbb{R}^p$ are observed pretreatment measurements and predictors of $\{Y^{(0)}, Y^{(1)}\}$. Our goal is to utilize the information in $\boldsymbol{X}$ to develop an ITR that optimizes the value of $h(E[Y])$ for future patients.

2 Method

2.1 Optimal individualized treatment rules

In this subsection, we define an optimal ITR. The Bayes decision $a^* : \boldsymbol{x} \mapsto \{0, 1\}$ minimizes, over treatment decision (action) $a \in \{0, 1\}$, the patient-specific posterior expected loss for a patient with pretreatment characteristic $\boldsymbol{X} = \boldsymbol{x}$. Let us define the loss function for making treatment decision a as:

$$L(a, \boldsymbol{\theta}, \boldsymbol{x}) = h(E[Y^{(a)}|\boldsymbol{\theta}, \boldsymbol{x}]), \tag{1}$$

where $\boldsymbol{\theta}$ represents the model parameters characterizing the relationship between the potential treatment outcomes $(Y^{(0)}, Y^{(1)})$ and predictors $\boldsymbol{X}$. In (1), $E[Y^{(a)}|\boldsymbol{\theta}, \boldsymbol{x}] = (1 - a)E[Y^{(0)}|\boldsymbol{\theta}, \boldsymbol{x}] + aE[Y^{(1)}|\boldsymbol{\theta}, \boldsymbol{x}]$ is the expected outcome under a particular treatment assignment a. Let us denote the observed clinical trial data as $\mathcal{D}$ consisting of triplets $\mathcal{D}_i = \{\boldsymbol{X}_i, A_i, Y_i\}$ with $i \in \{1, \ldots, n\}$, where $\boldsymbol{X}_i \in \mathbb{R}^p$ is a set of observed pretreatment covariates and $Y_i \in \{0, 1\}$ is an observed outcome for individual i.

Viewing the loss function $L(a, \boldsymbol{\theta}, \boldsymbol{x})$ in (1) as a function of treatment assignment a given a particular $\boldsymbol{x}$, the optimal Bayes decision $a^*(\boldsymbol{x})$ will minimize the posterior expected loss given $\boldsymbol{x}$, i.e.,

$$a^*(\boldsymbol{x}) = \underset{a \in \{0,1\}}{\arg\min} \, E_{\boldsymbol{\theta}|\mathcal{D},\boldsymbol{x}}[L(a, \boldsymbol{\theta}, \boldsymbol{x})],$$

where the expectation is taken with respect to the posterior distribution of $\boldsymbol{\theta}$ given the observed data $\mathcal{D}$. In particular, if we define the loss contrast $\Delta(\boldsymbol{\theta}, \boldsymbol{x}) := L(a = 1, \boldsymbol{\theta}, \boldsymbol{x}) - L(a = 0, \boldsymbol{\theta}, \boldsymbol{x})$, then the above optimal Bayes decision $a^*(\boldsymbol{x})$ is equivalent to:

$$a^*(\boldsymbol{x}) = \mathbb{I}_{(E_{\boldsymbol{\theta}|\mathcal{D},\boldsymbol{x}}[\Delta(\boldsymbol{\theta},\boldsymbol{x})]<0)} \tag{2}$$

(where $\mathbb{I}_{(\cdot)}$ is the indicator function), which we define as the optimal ITR.

We will utilize the following standard causal inference assumptions [see, e.g., Imbens and Rubin, 2015, Rubin, 2005]: 1) no unmeasured confoundedness, $Y^{(a)} \perp A$ given $\boldsymbol{X} = \boldsymbol{x}$, and 2) positivity, $0 < P(A = 1|\boldsymbol{X} = \boldsymbol{x}) < 1$, which together imply $h(E[Y^{(a)}|\boldsymbol{\theta}, \boldsymbol{X} = \boldsymbol{x}]) = h(E[Y|A = a, \boldsymbol{\theta}, \boldsymbol{X} = \boldsymbol{x}])$, and 3) stable unit treatment value assumption (SUTVA). Under these assumptions, we can write the loss contrast function $\Delta(\boldsymbol{\theta}, \boldsymbol{x})$ in (2) as: $\Delta(\boldsymbol{\theta}, \boldsymbol{x}) = h\{E[Y|A = 1, \boldsymbol{\theta}, \boldsymbol{x}]\} - h\{E[Y|A = 0, \boldsymbol{\theta}, \boldsymbol{x}]\}$. Therefore, we can construct the optimal Bayes decision (2) based on posterior inference on the canonical parameter $h\{E[Y|A, \boldsymbol{\theta}, \boldsymbol{X}]\}$ of the exponential family response Y. In the following subsection, we will specify our model for the distributions of $(Y|A, \boldsymbol{\theta}, \boldsymbol{X})$ and $\boldsymbol{\theta}$, to estimate the optimal ITR (2).

2.2 Model and prior specification

2.2.1 Model

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$ be a vector of the treatment outcomes, with each Y_i independently following a specific exponential family distribution with density

$$
\begin{aligned}
f(Y_i | \eta_i, \phi) &= \exp\left\{\phi^{-1}[Y_i \eta_i - b(\eta_i)] + c(Y_i, \phi)\right\} \\
\eta_i &= \boldsymbol{X}_i^\top \boldsymbol{m} + g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, A_i),
\end{aligned}
\tag{3}
$$

where the unknown parameters (which we collectively denote as $\boldsymbol{\theta}$) will be estimated in a Bayesian framework. In (3), $b(\cdot)$ and $c(\cdot)$ are known functions specific to the given member of the exponential family, whereas $g(\cdot, \cdot)$ is an unknown flexible function, and $\phi > 0$ is an unknown dispersion parameter ($\phi = 1$ specializes to a one-parameter exponential family distribution). Throughout we will drop ϕ from (3) because it is fixed at unity in our motivating dataset that has Bernoulli responses.

The canonical parameter $\eta \in \mathbb{R}$ in (3) of the response distribution is related to the treatment decision loss function $L(a, \boldsymbol{\theta}, \boldsymbol{x})$ in (1), through $\eta(\boldsymbol{\theta}, \boldsymbol{x}, a) = h(E[Y | \boldsymbol{\theta}, \boldsymbol{x}, a]) = h(E[Y^{(a)} | \boldsymbol{\theta}, \boldsymbol{x}]) = L(a, \boldsymbol{\theta}, \boldsymbol{x})$, under the standard causal inference assumptions.

Within the specification of η in model (3), the first term $\boldsymbol{X}^\top \boldsymbol{m}$ represents the pretreatment covariates' "main" effect, whereas the second term $g(\boldsymbol{X}^\top \boldsymbol{\beta}, A)$ is the $\boldsymbol{X}$-by-A interaction effect. This interaction effect is characterized by an unspecified treatment a-specific smooth function $g(u, a)$ ($a = 0, 1$) which is a function of a linear projection $u = \boldsymbol{X}^\top \boldsymbol{\beta} \in \mathbb{R}$. The projection vector $\boldsymbol{\beta} \in \mathbb{R}^p$ is subject to $\|\boldsymbol{\beta}\| = 1$, i.e., restricted to $\boldsymbol{\beta} \in \mathbb{S}^{p-1}$, where $\mathbb{S}^{p-1}$ is the $p - 1$ dimensional unit sphere, and the single-index $\boldsymbol{X}^\top \boldsymbol{\beta}$ provides a dimension reduction specifically for the $\boldsymbol{X}$-by-A interaction effect. In (3), for any $\boldsymbol{X}$ and $\boldsymbol{\beta}$, we shall impose the following identifiability condition for the component g

$$
E[g(\boldsymbol{X}^\top \boldsymbol{\beta}, A) | \boldsymbol{X}] = 0,
\tag{4}
$$

which separates the component $g(\boldsymbol{X}^\top \boldsymbol{\beta}, A)$ of interest (the "prescriptive" term representing the heterogeneous treatment effect), from the component $\boldsymbol{X}^\top \boldsymbol{m}$ (the "prognostic" term that does not represent the heterogeneous treatment effect). In general, the covariates, $\boldsymbol{X}$, represented in the terms $g(\boldsymbol{X}^\top \boldsymbol{\beta}, A)$ and $\boldsymbol{X}^\top \boldsymbol{m}$ in (3) are not necessarily the same variates, but for notational simplicity, the same notation for these sets of covariates was employed. As another abuse of notation for simplicity, the treatment A's main effect, which can be represented by $\beta_0 A \in \mathbb{R}$ for some unknown $\beta_0 \in \mathbb{R}$, will be estimated, together with the model intercept, as part of the component $\boldsymbol{X}^\top \boldsymbol{m}$ in (3).

Remark 2.1 *Model (3) with the identifiability condition (4) is more suitable to conduct a posterior inference for heterogeneous treatment effects than the model with $\eta = \boldsymbol{X}^\top \boldsymbol{m} + g(\boldsymbol{X}^\top \boldsymbol{\beta})A$, because this particular parametrization (3) is invariant of the choice of coding of A. In the latter model, the choice of the treatment A coding can meaningfully impact posterior inferences because the two effects $\boldsymbol{X}^\top \boldsymbol{m}$ and $g(\boldsymbol{X}^\top \boldsymbol{\beta})A$ can alias each another, particularly when the same set $\boldsymbol{X}$ enter into the both terms. Under condition (4), the effect captured by the prognostic term, $\boldsymbol{X}^\top \boldsymbol{m}$, and the prescriptive term, $g(\boldsymbol{X}^\top \boldsymbol{\beta}, A)$, can be distinguished regardless of treatment A coding.*

For an individual with pretreatment characteristics $\boldsymbol{x}$, the loss contrast $\Delta(\boldsymbol{\theta}, \boldsymbol{x})$ in (2) under model (3) is

$$
\Delta(\boldsymbol{\theta}, \boldsymbol{x}) = g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 1) - g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 0).
\tag{5}
$$

The loss contrast (3) indicates that only the parameters g and $\boldsymbol{\beta}$ (and not $\boldsymbol{m}$ and ϕ) in model (3) are used to specify the ITR (2), hence g and $\boldsymbol{\beta}$ correspond to the "signal" parameters of interest. Given definition (2), we will define a "treatment benefit index" (TBI) in terms of a (posterior) probability,

$$
\text{TBI}(\boldsymbol{x}) := P(\Delta(\boldsymbol{\theta}, \boldsymbol{x}) < 0 | \mathcal{D}) \in [0, 1],
\tag{6}
$$

that is, the probability of the (active) treatment $A = 1$ providing a greater benefit than the treatment $A = 0$ (for a patient with pretreatment characteristic $\boldsymbol{x}$), in which the probability is evaluated with respect to the posterior distribution of $\boldsymbol{\theta}$. The optimal Bayes decision $a^*(\boldsymbol{x})$ in (2) is then represented by $\mathbb{I}_{(\text{TBI}(\boldsymbol{x}) > 0.5)}$. Since a large value of (6) indicates a large expected "benefit" of taking the active treatment $A = 1$ vs. control $A = 0$, the TBI$(\boldsymbol{x})$ in (6) constructs a "gradient" of the active treatment's benefit that ranges from 0 to 1, as a function of patient characteristic $\boldsymbol{x}$. Furthermore, for each $\boldsymbol{X} = \boldsymbol{x}$, we can obtain a posterior distribution of the treatment a-specific expected outcome $h^{-1}\{\boldsymbol{x}^\top \boldsymbol{m} + g(\boldsymbol{x}^\top \boldsymbol{\beta}, a)\}$ based on the posterior distribution of the parameters $\boldsymbol{\theta} = \{\boldsymbol{m}, g, \boldsymbol{\beta}\}$.

2.2.2   Representation of the link function g

Following Antoniadis et al. [2004], we will use a cubic B-spline basis to represent the flexible function $g(\cdot, a)$ of (3). Using B-splines is appealing because the basis functions are strictly local, as each basis function is only non-zero over the intervals between 5 adjacent knots [Eilers and Marx, 1996]. Using splines allows us to easily incorporate the constraint (4) on the function g, as we describe shortly.

For each fixed $\boldsymbol{\beta}$, the function g is represented as:

$$g(\boldsymbol{\beta}^\top \boldsymbol{x}_i, a_i) = \tilde{\boldsymbol{\psi}}(\boldsymbol{\beta}^\top \boldsymbol{x}_i)^\top \tilde{\boldsymbol{\gamma}}_{a_i} \quad (i = 1, \ldots, n) \tag{7}$$

for some fixed L-dimensional basis $\tilde{\boldsymbol{\psi}}(\cdot) \in \mathbb{R}^L$ (e.g., B-spline basis on evenly-spaced knots on a bounded range of $\{\boldsymbol{\beta}^\top \boldsymbol{x}_i\}_{i=1}^n$) and a set of unknown treatment a-specific spline coefficients $\tilde{\boldsymbol{\gamma}}_a \in \mathbb{R}^L$ ($a = 0, 1$). In our simulation illustration and in the application, we used a cubic B-spline basis with $L = 4 + [n^{1/5.5}]$ where $[n^{1/5.5}]$ denotes the integer part of $n^{1/5.5}$, as recommended by Wang and Yang [2009].

Given representation (7) for g, the identifiability constraint (4) is implied by the linear constraint:

$$\pi_0 \tilde{\boldsymbol{\gamma}}_0 + \pi_1 \tilde{\boldsymbol{\gamma}}_1 = \boldsymbol{\pi} \tilde{\boldsymbol{\gamma}} = \boldsymbol{0}, \tag{8}$$

where $\boldsymbol{\pi} = [\pi_0 \boldsymbol{I}_L; \; \pi_1 \boldsymbol{I}_L]$ with $\pi_a = P(A = a)$ (i.e., the randomization probabilities) is the $L \times 2L$ matrix (in which $\boldsymbol{I}_L$ denotes the $L \times L$ identity matrix) and $\tilde{\boldsymbol{\gamma}} = (\tilde{\boldsymbol{\gamma}}_0^\top, \tilde{\boldsymbol{\gamma}}_1^\top) \in \mathbb{R}^{2L}$, an unknown basis coefficient vector.

To represent (7) in matrix notation, given $\boldsymbol{\beta}$, let the $n \times L$ matrices $\tilde{\boldsymbol{D}}_{\boldsymbol{\beta},a}$ ($a = 0, 1$) denote the evaluation matrices of the basis $\tilde{\boldsymbol{\psi}}(\cdot)$ on $\{\boldsymbol{\beta}^\top \boldsymbol{x}_i\}_{i=1}^n$, specific to the treatment $A = a$ ($a = 0, 1$), whose ith row is the $1 \times L$ vector $\tilde{\boldsymbol{\psi}}(\boldsymbol{\beta}^\top \boldsymbol{x}_i)^\top$ if $A_i = a$, and a row of zeros $\boldsymbol{0}_L^\top$ if $A_i \neq a$. Then, the column-wise concatenation of the design matrices $\tilde{\boldsymbol{D}}_{\boldsymbol{\beta},a}$ ($a = 0, 1$), i.e., the $n \times 2L$ matrix $\tilde{\boldsymbol{D}}_{\boldsymbol{\beta}} = [\tilde{\boldsymbol{D}}_{\boldsymbol{\beta},0}; \tilde{\boldsymbol{D}}_{\boldsymbol{\beta},1}]$, defines the model matrix associated with $\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^{2L}$. Then, we can represent the function g in (7) evaluated on the sample data, by the length-n vector: $\boldsymbol{g} = \tilde{\boldsymbol{D}}_{\boldsymbol{\beta}} \tilde{\boldsymbol{\gamma}} \in \mathbb{R}^n$.

The linear constraint (8) on $\tilde{\boldsymbol{\gamma}}$ can be conveniently absorbed into the model matrix $\tilde{\boldsymbol{D}}_{\boldsymbol{\beta}}$ by reparametrization, as we describe next. We can find a $2L \times L$ basis matrix $\boldsymbol{Q}$ (that spans the *null* space of the linear constraint (8)), such that if we set $\tilde{\boldsymbol{\gamma}} = \boldsymbol{Q}\boldsymbol{\gamma}$ for any arbitrary vector $\boldsymbol{\gamma} \in \mathbb{R}^L$, then the resulting vector $\tilde{\boldsymbol{\gamma}} \in \mathbb{R}^{2L}$ automatically satisfies the constraint (8). Such a basis matrix $\boldsymbol{Q}$ can be constructed by a QR decomposition of the matrix $\boldsymbol{\pi}^\top$ in (8). Then representation $\boldsymbol{g} = \tilde{\boldsymbol{D}}_{\boldsymbol{\beta}} \tilde{\boldsymbol{\gamma}}$ can be reparametrized, in terms of the unconstrained vector $\boldsymbol{\gamma} \in \mathbb{R}^L$, by replacing $\tilde{\boldsymbol{D}}_{\boldsymbol{\beta}}$ with the reparametrized model matrix $\boldsymbol{D}_{\boldsymbol{\beta}} = \tilde{\boldsymbol{D}}_{\boldsymbol{\beta}} \boldsymbol{Q}$ yielding the representation $\boldsymbol{g} = \boldsymbol{D}_{\boldsymbol{\beta}} \boldsymbol{\gamma}$.

Once we have an inferential procedure on $\boldsymbol{\gamma}$, we can also consider inference on the transformed parameter $\tilde{\boldsymbol{\gamma}} = \boldsymbol{Q}\boldsymbol{\gamma}$, from which we can make inference on the treatment a-specific functions $g(\boldsymbol{\beta}^\top \cdot, a) = \tilde{\boldsymbol{\psi}}(\boldsymbol{\beta}^\top \cdot)^\top \tilde{\boldsymbol{\gamma}}_a$ ($a = 0, 1$).

We note that $\tilde{\boldsymbol{\psi}}(\cdot) \in \mathbb{R}^L$ in (7) defines a system of functions specifically chosen to be used as building blocks to represent a (smooth, as implied by penalized splines) link function $g(\cdot, a)$ for each treatment condition a. If a different basis function is used to represent the link function, we may have a different performance. We may need to identify a best suitable basis for the data which depend on the underlying heterogeneous treatment effect (i.e., the underlying log odds ratio function). For example, if there is a reason to expect that the log odds ratio function (5) is a jagged (or, less likely, a cyclic) function over the index $\boldsymbol{x}^\top \boldsymbol{\beta}$, then a wavelet [e.g., Wand and Ormerod, 2011] (or Fourier [e.g., Lenk, 2002]) basis might be more suitable. Although the associated basis coefficient vector, $\tilde{\boldsymbol{\gamma}}$, will still be subject to the identifiability constraint (8) for a different basis system, a specifically tailored prior and penalization would be more appropriate, for example, (L_1-type) Laplace-Zero prior with some hyperparameter determining the sparsity basis [Wand and Ormerod, 2011].

2.2.3   Prior specification

How we specify priors for $\boldsymbol{\beta}$, $\boldsymbol{m}$ and $\boldsymbol{\gamma}$ associated with model (3) is given in this subsection.

- For the distribution of $\boldsymbol{\beta} \in \mathbb{S}^{p-1}$, we will use the von Mises-Fisher with concentration parameter $\kappa > 0$ and modal parameter $\boldsymbol{\beta}_0 \in \mathbb{S}^{p-1}$,

$$P(\boldsymbol{\beta}) \propto \exp(\kappa \boldsymbol{\beta}^\top \boldsymbol{\beta}_0), \tag{9}$$

  which is a probability distribution for $\boldsymbol{\beta}$ on the $(p-1)$-unit-sphere in $\mathbb{R}^p$.
- We will use $\boldsymbol{m} \sim N(\boldsymbol{m}_0, \boldsymbol{R})$, for some vector $\boldsymbol{m}_0 \in \mathbb{R}^p$ and $p \times p$ symmetric positive definite matrix $\boldsymbol{R}$.
- Since the domain of the function g in (3) depends on $\boldsymbol{\beta}$, the prior on g will depend on $\boldsymbol{\beta}$. Conditioning on $\boldsymbol{\beta}$, following Antoniadis et al. [2004], we will use data-dependent prior for $\boldsymbol{\gamma} \in \mathbb{R}^L$,

$$\boldsymbol{\gamma} | \boldsymbol{\beta} \sim N(\boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W} \boldsymbol{Z}, \; \boldsymbol{\Sigma}_0), \tag{10}$$

where

$$\mathbf{\Sigma}_\rho := (\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W} \boldsymbol{D}_{\boldsymbol{\beta}} + \rho \boldsymbol{I})^{-1},$$

and $\mathbf{\Sigma}_0 := (\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W} \boldsymbol{D}_{\boldsymbol{\beta}})^{-1}$ which corresponds to a special case of $\mathbf{\Sigma}_\rho$ at $\rho = 0$. The prior (10) is a Zellner's g-prior, that has the same dispersion matrix as a weighted least squares estimator defined based on the vector of adjusted responses $\boldsymbol{Z}$ and the matrix of weights $\boldsymbol{W}$, which are specified in the next subsection. In the prior (10), $\rho \geq 0$ is a hyperparameter which will be selected via an empirical Bayes procedure with the generalized cross-validation (GCV), as in Antoniadis et al. [2004]. An advantage of using the prior (10) is that it allows us to analytically integrate $\boldsymbol{\gamma}$ out of the joint posterior $P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y})$, facilitating the Gibbs sampling of $\boldsymbol{\beta}$.

2.3   Posterior computation

To conduct posterior inference on $(\boldsymbol{m}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, we will simulate samples from the joint posterior $P(\boldsymbol{m}, \boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{Y})$ (where we use $\boldsymbol{Y} = (y_1, \ldots, y_n)^\top$ to denote the observed treatment outcomes). Since it is difficult to draw samples directly from this joint posterior distribution, we will use a Metropolis-Within-Gibbs algorithm. The Gibbs algorithm will iterate between the following two Steps: *Step 1)* sample $\boldsymbol{m}$ from $P(\boldsymbol{m} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{Y})$; *Step 2)* sample $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ from $P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y})$. Specifically, in *Step 2*, since the joint conditional posterior $P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y})$ does not have a convenient form to directly sample from, we will employ a Metropolis-Hastings step.

2.3.1   Conditional posteriors

1. Derivation of $(\boldsymbol{m} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{Y})$. For fixed $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we will quadratically approximate the log likelihood function of $\boldsymbol{m}$ at its mode, which we denote by $\breve{\boldsymbol{m}}$. To find the mode $\breve{\boldsymbol{m}}$, we will use a Fisher scoring, iteratively updating the center of the quadratic approximation. For fixed $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, at the convergence of the Fisher scoring, we define the *adjusted response* vector $\breve{\boldsymbol{Z}} := (\breve{z}_1, \ldots, \breve{z}_n)^\top \in \mathbb{R}^n$ where $\breve{z}_i := h'(\breve{\mu}_i)(y_i - \breve{\mu}_i) + \breve{\eta}_i$, in which $\breve{\eta}_i = \breve{\boldsymbol{m}}^\top \boldsymbol{x}_i + \boldsymbol{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x}_i)^\top \boldsymbol{\gamma}$ and $\breve{\mu}_i = h^{-1}(\breve{\eta}_i)$, and the $n \times n$ *weight* matrix $\breve{\boldsymbol{W}} = \mathrm{diag}(\breve{w}_i)$, where $\breve{w}_i = 1 / \{(h'(\breve{\mu}_i))^2 V(\breve{\mu}_i)\}$, in which $V(\breve{\mu}_i) = \breve{\mu}_i(1 - \breve{\mu}_i)$. Given each $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the negative log likelihood of $\boldsymbol{m}$ is approximately represented in terms of a weighted least squares (WLS) objective function (up to a constant of proportionality),

$$\sum_{i=1}^n \breve{w}_i(\breve{z}_i - \boldsymbol{m}^\top \boldsymbol{x}_i)^2 = (\breve{\boldsymbol{Z}} - \boldsymbol{X}\boldsymbol{m})^\top \breve{\boldsymbol{W}}(\breve{\boldsymbol{Z}} - \boldsymbol{X}\boldsymbol{m}).$$

   Given the prior $\boldsymbol{m} \sim N(\boldsymbol{m}_0, \boldsymbol{R})$ and the above approximated negative log likelihood, the conditional posterior for $\boldsymbol{m}$ is given by:

$$P(\boldsymbol{m} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{Y}) = N\left((\boldsymbol{R}^{-1} + \boldsymbol{X}^\top \breve{\boldsymbol{W}} \boldsymbol{X})^{-1}(\boldsymbol{R}^{-1}\boldsymbol{m}_0 + \boldsymbol{X}^\top \breve{\boldsymbol{W}} \breve{\boldsymbol{Z}}), \, (\boldsymbol{R}^{-1} + \boldsymbol{X}^\top \breve{\boldsymbol{W}} \boldsymbol{X})^{-1}\right). \quad (11)$$

2. Derivation of $(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y})$.   Given the joint conditional posterior $P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y}) = P(\boldsymbol{\beta} | \boldsymbol{m}, \boldsymbol{Y})P(\boldsymbol{\gamma} | \boldsymbol{\beta}, \boldsymbol{m}, \boldsymbol{Y})$, we will first sample $\boldsymbol{\beta}$ from $P(\boldsymbol{\beta} | \boldsymbol{m}, \boldsymbol{Y})$ and then $\boldsymbol{\gamma}$ from $P(\boldsymbol{\gamma} | \boldsymbol{\beta}, \boldsymbol{m}, \boldsymbol{Y})$. Specifically, following Antoniadis et al. [2004], we will use a Metropolis-Hastings algorithm to sample $\boldsymbol{\beta}$ from $p(\boldsymbol{\beta} | \boldsymbol{m}, \boldsymbol{Y})$. However, this approach employed in Antoniadis et al. [2004] cannot be directly applied to our settings, due to the non-Gaussian likelihood. Thus, we will perform a quadratic approximation of the negative log likelihood function of $\boldsymbol{\gamma}$ at its mode, which we denote by $\breve{\boldsymbol{\gamma}}$. To find $\breve{\boldsymbol{\gamma}}$, as in Step 1, we will conduct a Fisher scoring. For each fixed $\boldsymbol{\beta}$ and $\boldsymbol{m}$, this quadratic approximation at the convergence of the Fisher scoring is summarized in the form of the WLS objective function (up to a constant of proportionality),

$$\sum_{i=1}^n w_i(z_i - \boldsymbol{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x}_i)^\top \boldsymbol{\gamma})^2 = (\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma})^\top \boldsymbol{W}(\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma}), \quad (12)$$

   as a function of $\boldsymbol{\gamma}$, in which $\boldsymbol{Z} := (z_1, \ldots, z_n)^\top \in \mathbb{R}^n$ is the *adjusted response* vector with $z_i := h'(\hat{\mu}_i)(y_i - \hat{\mu}_i) + \hat{\eta}_i$ obtained at the convergence, where $\hat{\eta}_i = \boldsymbol{m}^\top \boldsymbol{x}_i + \boldsymbol{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x}_i)^\top \breve{\boldsymbol{\gamma}}$ and $\hat{\mu}_i = h^{-1}(\hat{\eta}_i)$, and $\boldsymbol{W} = \mathrm{diag}(w_i)$ is the $n \times n$ *weight* matrix with $w_i = 1 / \{(h'(\hat{\mu}_i))^2 V(\hat{\mu}_i)\}$. Given the quadratic approximation (12), we can write the joint conditional posterior $P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y})$:

$$\begin{aligned} P(\boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{m}, \boldsymbol{Y}) &= P(\boldsymbol{Y} | \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{m}) P(\boldsymbol{\gamma} | \boldsymbol{\beta}, \boldsymbol{m}) P(\boldsymbol{\beta}) \\ &\propto \exp\{-\frac{1}{2}(\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma})^\top \boldsymbol{W}(\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma})\} \, P(\boldsymbol{\gamma} | \boldsymbol{\beta}, \boldsymbol{m}) \, \exp(\kappa \boldsymbol{\beta}^\top \boldsymbol{\beta}_0). \end{aligned} \quad (13)$$

Given the conditional prior $P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{m})$ in (10), we can write the terms involving $\boldsymbol{\gamma}$ in (13) as:

$$\propto \exp\left(-\frac{1}{2}\left\{(\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma})^\top \boldsymbol{W}(\boldsymbol{Z} - \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma}) + (\boldsymbol{\gamma} - \boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z})^\top \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\gamma} - \boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z})\right\}\right)$$

$$\propto \exp\left(-\frac{1}{2}\left\{(\boldsymbol{\gamma} - \boldsymbol{\Sigma}_0 \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z})^\top \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\gamma} - \boldsymbol{\Sigma}_0 \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}) + (\boldsymbol{\gamma} - \boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z})^\top \boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\gamma} - \boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z})\right\}\right)$$

$$= \exp\left(-\frac{1}{2}\left\{2\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\gamma} - 2\boldsymbol{\gamma}^\top(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z} + \boldsymbol{Z}^\top \boldsymbol{W}\boldsymbol{Z} + \boldsymbol{Z}^\top \boldsymbol{W}^\top \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\Sigma}_\rho \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}\right\}\right)$$

$$= \exp\left(-\frac{1}{2}\left\{2\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\gamma} - 2\boldsymbol{\gamma}^\top(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z} + S_1(\boldsymbol{\beta})\right\}\right),$$

(14)

where

$$S_1(\boldsymbol{\beta}) = \boldsymbol{Z}^\top \boldsymbol{W}\boldsymbol{Z} + \boldsymbol{Z}^\top \boldsymbol{W}^\top \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\Sigma}_\rho \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho \boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}.$$

Specifically, given (14), we can analytically integrate $\boldsymbol{\gamma}$ out of the joint conditional $P(\boldsymbol{\beta}, \boldsymbol{\gamma}|\boldsymbol{m}, \boldsymbol{Y})$ in (13), which yields

$$P(\boldsymbol{\beta}|\boldsymbol{m}, \boldsymbol{Y}) = \int P(\boldsymbol{\beta}, \boldsymbol{\gamma}|\boldsymbol{m}, \boldsymbol{Y})d\boldsymbol{\gamma}$$

$$\propto \int \exp\left(-\frac{1}{2}\left\{2\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\gamma} - 2\boldsymbol{\gamma}^\top(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z} + S_1(\boldsymbol{\beta})\right\}\right)\frac{1}{|\boldsymbol{\Sigma}_0|^{1/2}}\exp(\kappa\boldsymbol{\beta}^\top \boldsymbol{\beta}_0)d\boldsymbol{\gamma}$$

$$\propto \varphi_{\tilde{\boldsymbol{\gamma}}}\left[(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}\right]\exp\left\{-\frac{1}{2}S_1(\boldsymbol{\beta})\right\}\exp(\kappa\boldsymbol{\beta}^\top \boldsymbol{\beta}_0),$$

(15)

where $\varphi_{\tilde{\boldsymbol{\gamma}}}\left[(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}\right]$ is the moment generating function (MGF) of the variate $\tilde{\boldsymbol{\gamma}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_0/2)$ evaluated at $(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}$. The familiar closed-form expression of the Gaussian MGF allows us to write the last line of (15) as:

$$P(\boldsymbol{\beta}|\boldsymbol{m}, \boldsymbol{Y}) \propto \exp\left(\frac{1}{4}\boldsymbol{Z}^\top \boldsymbol{W}^\top \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\Lambda}\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}\right)\exp\left\{-\frac{1}{2}S_1(\boldsymbol{\beta})\right\}\exp(\kappa\boldsymbol{\beta}^\top \boldsymbol{\beta}_0), \qquad (16)$$

where $\boldsymbol{\Lambda} = (\boldsymbol{I} + \boldsymbol{\Sigma}_\rho\boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{\Sigma}_\rho\boldsymbol{\Sigma}_0^{-1})$. The expression (16) provides a closed form for the approximated $P(\boldsymbol{\beta}|\boldsymbol{m}, \boldsymbol{Y})$ up to a constant of proportionality, which we will use to conduct a random walk Metropolis Markov chain Monte Carlo (MCMC) algorithm. The MCMC algorithm to sample $(\boldsymbol{\beta}|\boldsymbol{m}, \boldsymbol{Y})$ based on (16) is described in the next subsection. Given each $\boldsymbol{m}$ and $\boldsymbol{\beta}$, we can sample $\boldsymbol{\gamma}$ from $P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{m}, \boldsymbol{Y})$,

$$P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{m}, \boldsymbol{Y}) = N(\tfrac{1}{2}\boldsymbol{\Sigma}_0(\boldsymbol{I} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\rho)\boldsymbol{D}_{\boldsymbol{\beta}}^\top \boldsymbol{W}\boldsymbol{Z}, \ \tfrac{1}{2}\boldsymbol{\Sigma}_0), \qquad (17)$$

derived from expression (14).

2.3.2　MCMC algorithm for the posterior sampling

In this subsection, we provide a detailed sampling scheme based on the conditional posterior derived in the previous subsection. First, we initialize the chain with the maximum likelihood estimates of the parameters $(\boldsymbol{m}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ of model (3) with representation (7) for the link g, where the tuning parameter $\rho \geq 0$ is optimized through the generalized cross-validation (GCV) criterion. We will then cycle through the following steps.

1. Sample $\boldsymbol{m}$ from $P(\boldsymbol{m}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{Y})$ in (11) given $(\boldsymbol{\beta}, \boldsymbol{\gamma})$.

2. Sample $\boldsymbol{\beta}$ from $P(\boldsymbol{\beta}|\boldsymbol{m}, \boldsymbol{Y})$ in (16) given $\boldsymbol{m}$, using the Metropolis algorithm. Specifically, given the current state $\boldsymbol{\beta}^{\text{cur}}$ for $\boldsymbol{\beta}$ of the chain, a new value $\boldsymbol{\beta}^{\text{new}}$ is accepted with the acceptance probability $\min\{1, r\}$, where the Metropolis ratio r is given by:

$$r = \frac{P(\boldsymbol{\beta}^{\text{new}}|\boldsymbol{m}, \boldsymbol{Y})}{P(\boldsymbol{\beta}^{\text{cur}}|\boldsymbol{m}, \boldsymbol{Y})},$$

using the conditional posterior (16) given $\boldsymbol{m}$. Here, we provide some more details on this Metropolis procedure.

- The proposal distribution for $\boldsymbol{\beta}^{\text{new}}$ was taken to be von Mises-Fisher with concentration parameter κ to be $\lambda_{prop} > 0$ and direction parameter given by the current value $\boldsymbol{\beta}^{\text{cur}}$. In the simulation example in the next section, we used $\lambda_{prop} = 1000$, which yielded the acceptance probability of around 0.3~0.7

for proposal $\boldsymbol{\beta}^{\text{new}}$, and the sampler appeared to explore the state space for $\boldsymbol{\beta}$ adequately (examining the traceplots of typical MCMC chains did not show any peculiarity). We used the R package movMF [Hornik and Grun, 2022] to generate random samples of $\boldsymbol{\beta} \in \mathbb{S}^{p-1}$ from von Mises-Fisher distributions.

- For the prior distribution of $\boldsymbol{\beta}$ in (9), we can choose $\kappa \geq 0$ (typically in the range of $0 \leq \kappa < 700$ [Antoniadis et al., 2004]) where $\kappa = 0$ corresponds to an uninformative prior, depending on the degree of confidence in the prior direction $\boldsymbol{\beta}_0$.
- $\rho \geq 0$ in (10) is another unknown that controls the smoothness of the data-driven function g, which is crucial to avoid overfitting g. This will be selected via an empirical Bayes procedure using the GCV criterion, at each MCMC update.

3. Sample $\boldsymbol{\gamma}$ from $P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{m}, \boldsymbol{Y})$ in (17) given $(\boldsymbol{\beta}, \boldsymbol{m})$.

To obtain the estimated expected response given a new $\boldsymbol{X} = \boldsymbol{x}$ and a treatment condition $a \in \{0, 1\}$, we take the posterior mean of the expected response $h^{-1}(\eta) = h^{-1}\left(\boldsymbol{m}^\top \boldsymbol{x} + \tilde{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x})\tilde{\boldsymbol{\gamma}}_a\right)$, based on the posterior sampler output. In particular, we construct a treatment decision rule using the posterior distribution of $\tilde{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x})(\tilde{\boldsymbol{\gamma}}_1 - \tilde{\boldsymbol{\gamma}}_0)$. Specifically, we will use the posterior probability $P\left(\tilde{\psi}(\boldsymbol{\beta}^\top \boldsymbol{x})(\tilde{\boldsymbol{\gamma}}_1 - \tilde{\boldsymbol{\gamma}}_0) < 0 \mid \mathcal{D}\right)$ as the TBI$(\boldsymbol{x})$, which we will utilize to obtain a decision rule $a^*(\boldsymbol{x}) = \mathbb{I}_{(\text{TBI}(\boldsymbol{x})>0.5)}$, using the probability threshold of 0.5.

3   Simulation illustration

In this section, via a set of simulation experiments, we compare the performance of the proposed Bayesian single-index approach to modeling heterogeneous treatment effect with an approach that relies on a Bayesian linear model. We replicate the experiment 100 times with sample sizes $n = 500, 1000, 2000$. We use a cubic B-spline for representing g as in Section 2.2.2. The simulation code is available at https://github.com/syhyunpark/bayesSIMML.

3.1   Simulation setting

We independently generated the treatment indicators $A_i \in \{0, 1\}$ from Bernoulli distribution with $P(A_i = 1) = 0.5$ and the vector of covariates $\boldsymbol{X}_i \in \mathbb{R}^p$ from the mean zero multivariate normal distribution with compound symmetry correlation $(= 0.2)$ and the unit variances. Given $(\boldsymbol{X}_i, A_i)$, we generated $Y_i \sim \text{Bernoulli}(P(Y_i = 1))$, where $\text{logit}(P(Y_i = 1)) = m(\boldsymbol{X}_i) + g(\boldsymbol{X}_i, A_i)$, with the following specifications of the functions m (either a "nonlinear" or "linear" $\boldsymbol{X}$ main effect) and g (either a "nonlinear" or "linear" A-by-$\boldsymbol{X}$ interaction effect) :

$$
\begin{aligned}
m(\boldsymbol{X}_i) &= \begin{cases} \frac{1}{2}\sin\left(\frac{\pi}{2}\boldsymbol{m}^\top \boldsymbol{X}_i\right) & \text{"nonlinear } \boldsymbol{X} \text{ main effect"} \\ \frac{\pi}{8}\boldsymbol{m}^\top \boldsymbol{X}_i & \text{"linear } \boldsymbol{X} \text{ main effect"} \end{cases} \\
g(\boldsymbol{X}_i, A_i) &= \begin{cases} 2\left(\exp(-(\boldsymbol{\beta}^\top \boldsymbol{X}_i - 0.5)^2) - 0.6\right)(A_i - 0.5) & \text{"nonlinear } \boldsymbol{X}\text{-by-}A \text{ interaction effect"} \\ \left(0.6\boldsymbol{\beta}^\top \boldsymbol{X}_i\right)(A_i - 0.5) & \text{"linear } \boldsymbol{X}\text{-by-}A \text{ interaction effect"} \end{cases}
\end{aligned}
\tag{18}
$$

In (18), we set $\boldsymbol{m} = (1, 2, 3, 4, 0, \ldots, 0)^\top \in \mathbb{R}^p$ and $\boldsymbol{\beta} = (1, 0.5, 0.25, 0.125, 0, \ldots, 0)^\top \in \mathbb{R}^p$, where each vector was normalized to have unit norm. We considered the cases with $p \in \{5, 10\}$. Throughout the paper, we took $\boldsymbol{m} \sim N(\boldsymbol{0}, 5^2 \boldsymbol{I})$ and an uninformative prior for $\boldsymbol{\beta}$, i.e., set $\kappa = 0$ in (9). As a comparison method, we used the following logistic linear model: $\text{logit}(P(Y_i = 1)) = \alpha + \boldsymbol{m}^\top \boldsymbol{X}_i + (\beta_0 + \boldsymbol{\beta}^\top \boldsymbol{X}_i)(A_i - 0.5)$, with the prior distributions, $\alpha \sim t(df = 3)$ with location$= 0$ and scale$= 8$, $\beta_0 \sim N(0, 5^2)$, and $\boldsymbol{m}, \boldsymbol{\beta} \sim N(\boldsymbol{0}, 5^2 \boldsymbol{I})$.

As an evaluation metric, we first consider the expected deviance measured on an independently generated test set (of size $\tilde{n} = 10^4$) to assess the accuracy of the models in predicting the new data, as defined by:

$$
\text{Deviance}(\tilde{\boldsymbol{y}}, \boldsymbol{\theta}) = -\frac{2}{\tilde{n}} \sum_{i=1}^{\tilde{n}} \log\left(\frac{1}{T}\sum_{t=1}^{T} P(\tilde{y}_i|\boldsymbol{\theta}_t)\right)
\tag{19}
$$

(note that it is scaled by the testing set size $\tilde{n}$), where $T(= 2000)$ is the length of the Markov chain (after a burn-in of 2000) and $\boldsymbol{\theta}_t$ is the tth set of sampled parameter values within the Markov chain obtained from the training set. Accordingly, $\frac{1}{T}\sum_{t=1}^{T} P(\tilde{y}_i|\boldsymbol{\theta}_t)$ in (19) corresponds to the average posterior probability of $\tilde{y}_i$, and the quantity (19) corresponds to a Bayesian version of the deviance, evaluated on the test set (of size $\tilde{n}$). Smaller values of (19) are better, indicating greater average accuracy of the predictive model.

Furthermore, we report the expected outcome under the treatment regime a^*, i.e., $E[Y^{(a^*(\boldsymbol{X}))}]$ (which is called the "value" of the regime a^*) that is Monte-Carlo approximated by $\tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \tilde{y}_i^{(a^*(\tilde{\boldsymbol{x}}_i))}$ based on the test set of size $\tilde{n}$, where

the ITR $a^*(\boldsymbol{x}) = \mathbb{I}_{(\text{TBI}(\boldsymbol{x})>0.5)}$ (see (6) for the definition of the TBI) which is trained based on the training set. We also report the proportion of correct decision (PCD), which is the proportion of the cases such that $a^*(\tilde{\boldsymbol{x}}_i)$ $(i = 1, \ldots, \tilde{n})$ match with the correct optimal treatment assignment under the true data generation model.

3.2   Simulation results

Figure 1 below displays the results from the simulation experiments when we vary $n \in \{500, 1000, 2000\}$, $p \in \{5, 10\}$ and the form of the interaction effect component $g \in \{\text{"nonlinear"}, \text{"linear"}\}$, for the *linear* $\boldsymbol{X}$ main effect (i.e., $m(\boldsymbol{X}) = \frac{\pi}{8}\boldsymbol{m}^\top \boldsymbol{X}$) case. In Appendix, as MCMC convergence diagnostic, we report the Gelman-Rubin potential scale reduction factor (PSRF) [Gelman and Rubin, 1992] computed for each scenario using the method of Vats and Knudson [2021], which provides some assurance that the sampler has performed reasonably. For the nonlinear A-by-$\boldsymbol{X}$ interaction effect scenarios (i.e., the gray panels in Figure 1), the proposed index model that utilizes the flexible link function g clearly outperforms the logistic linear model which assumes a restricted linear model on the interaction term, with respect to the all three criteria (the deviance, PCD and the expected outcome). When there is no nonlinearity in the A-by-$\boldsymbol{X}$ interaction effect term in the underlying model (i.e., the white panels in Figure 1), not surprisingly the logistic linear model outperforms the index model. However, the contrast in the performance between the two models is relatively small, compared to that under the nonlinear A-by-$\boldsymbol{X}$ interaction. This suggests that, in the absence of prior knowledge about the form of the A-by-$\boldsymbol{X}$ interaction effect, the more flexible index model that accommodates nonlinear treatment effect-modifications (i.e., the nonlinear g term) can be a useful alternative to the linear model approach.
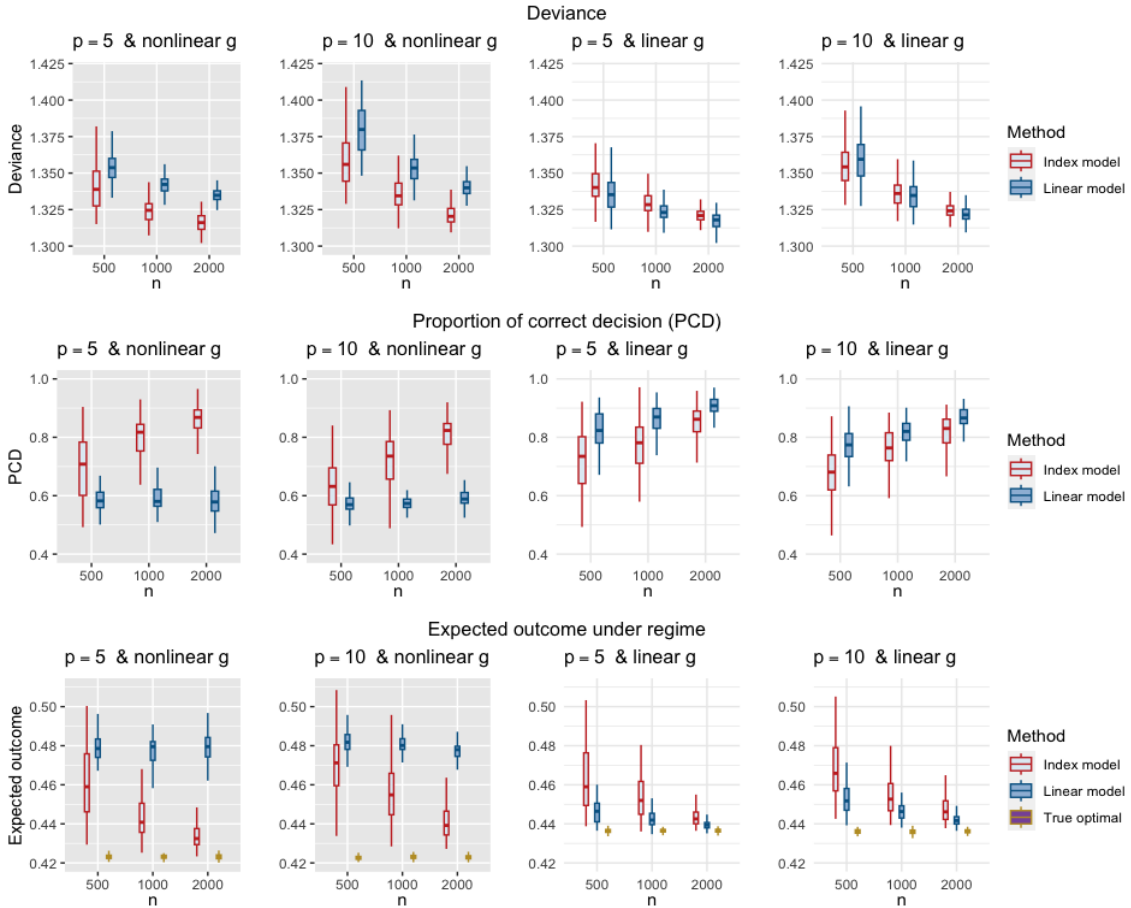

Figure 1:   Results for the *linear* $\boldsymbol{X}$ main effect ($m = $ "linear") case, with varying $g \in \{\text{"nonlinear"}, \text{"linear"}\}$, $p \in \{5, 10\}$, and $n \in \{500, 1000, 2000\}$, comparing the performance of the proposed index model (red) with that of the logistic linear regression model (blue), with respect to the deviance (the first row; a smaller deviance is desired), the proportion of correct decisions (PCD) (the second row; a larger PCD is desired), and the "value" (the expected outcomes under ITRs) (the third row; a smaller value is desired).
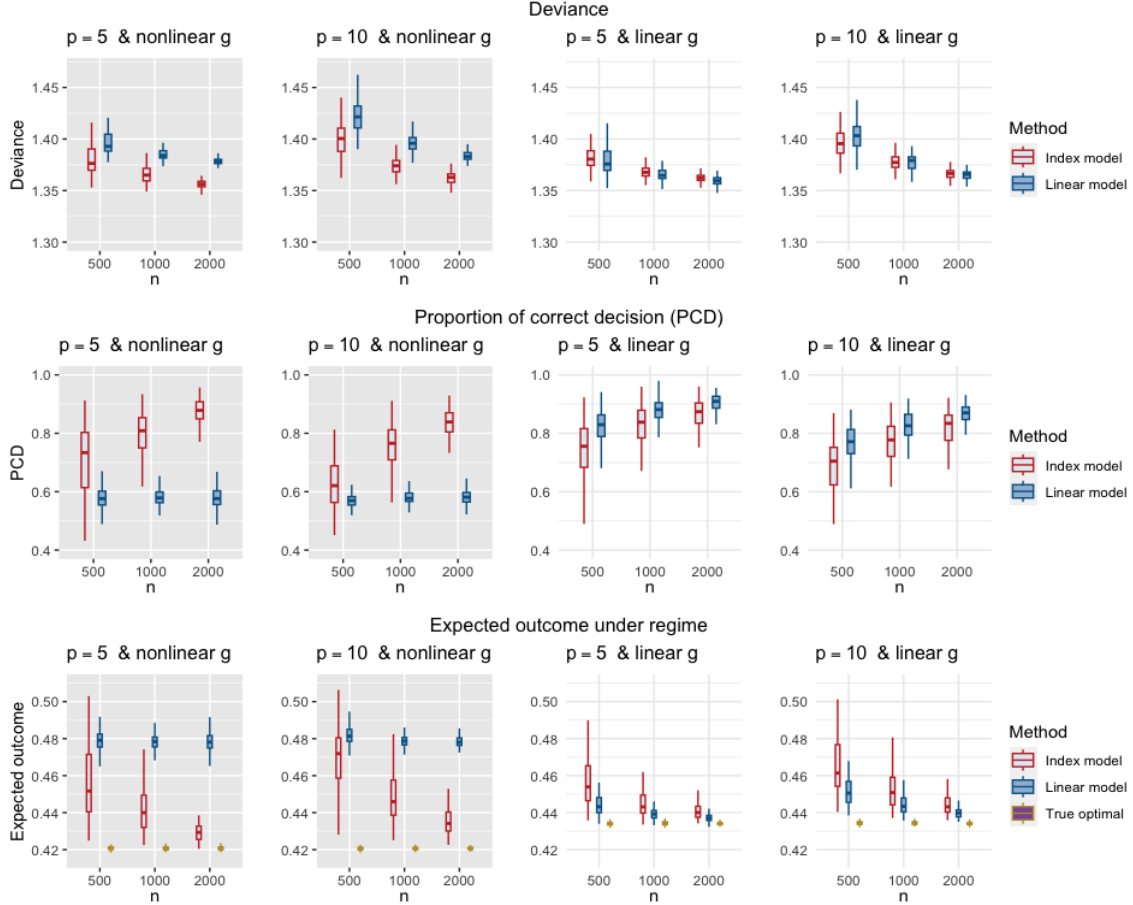
Figure 2: Results for the *nonlinear X* main effect ($m =$ "nonlinear") case, with varying $g \in \{$"nonlinear", "linear"$\}$, $p \in \{5, 10\}$, and $n \in \{500, 1000, 2000\}$, comparing the performance of the proposed index model (red) with that of the logistic linear regression model (blue), with respect to the deviance (the first row; a smaller deviance is desired), the proportion of correct decisions (PCD) (the second row; a larger PCD is desired), and the "value" (the expected outcomes under ITRs) (the third row; a smaller value is desired).

The results under the *nonlinear X* main effect case, i.e., $m(X) = \frac{1}{2} \sin\left(\frac{\pi}{2} m^\top X_i\right)$, are given in Figure 2, and are quite similar to those from the *linear X* main effect case (i.e., the results present in Figure 1), except that the expected deviance of the models (both the index model and the logistic linear model), which is a generalization of the mean squared error that assesses the overall predictive performance, slightly increased due to the X main effect model misspecification. However, optimal ITRs are derived based only on the model's A-by-X interaction effect terms. With respect to the PCD and the expected outcome under ITRs, the results under the linear X main effect and those under the nonlinear X main effect were close to each other, indicating that these models were quite robust to misspecification of the X main effect with respect to the ITR estimation performance. We also note that in both Figures 1 and 2 there was relatively large variability in the ITR estimation performance for the index model, especially when $n = 500$. On the other hand, the variability in the overall predictive performance metric (deviance) was comparable for the two approaches. This reflects the tendency that estimation of optimal ITRs is generally more challenging than making outcome predictions, as the difference (5) between two predictions has to be computed, and its expected error will be invariably larger than for a single prediction.

4   Application

In this section, we illustrate the proposed model on data from a COVID-19 convalescent plasma (CCP) study [Troxel et al., 2022], a meta-analysis of pooled individual patient data from 8 randomized clinical trials. One of the goals of this study was to guide CCP treatment recommendations by providing an estimate of a differential treatment outcome when a patient is treated with CCP vs. without CCP [Park et al., 2022]. A larger differential in favor of CCP would indicate a

more compelling reason for recommending CCP. In this context, we aim to use profiles of patients with COVID-19 associated with different benefits from CCP treatment, to optimize treatment decisions.

The study included 2369 hospitalized adults, not receiving mechanical ventilation at randomization, enrolled from April 2020 through March 2021. We used only complete cases for this analysis. A total of 2287 patients were included, with a mean age of 60.3 (SD 15.2) years and $815(35.6\%)$ women. One of the study's primary outcomes was the binary variable indicating mechanical ventilation or death (hence $Y = 1$ indicates a bad outcome, whose probability we want to minimize) at day 14 post-treatment, where 336 out of 2287 patients (14.7%) experienced the $Y = 1$ case at day 14. The patients were randomized to be treated with either CCP $(A = 1)$ (1190 patients, 52%) or control $(A = 0)$ (i.e., standard of care; 1097 patients, 48%). Pretreatment patient characteristics were collected at baseline. As in [Park et al., 2022], in our application, the baseline variables used to model the covariates "main" effect, i.e., the component associated with the coefficient $\boldsymbol{m}$ in model (3), included age, sex, baseline symptom conditions, age-by-baseline symptom conditions interaction, blood type, the indicators for history of diabetes, pulmonary and cardiovascular disease, and days since the symptoms onset. We also included the RCT-specific intercepts and the patients' enrollment quarters as part of the covariates "main" effect component.

Since our goal in this analysis is to investigate the differential treatment effect explained by the baseline variables $\boldsymbol{X}$, we will focus on reporting the estimation results of the heterogeneous treatment effect (HTE) term $g(\boldsymbol{X}^\top \boldsymbol{\beta}, A)$ in model (3) and the corresponding treatment effect contrast $\Delta(\boldsymbol{x}, \boldsymbol{\theta})$ in (5). The patient characteristics $\boldsymbol{X}$ included in the HTE term, along with the sample proportions, are given in the first column of Table 1. The posterior mean of the index coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_7)^\top$, along with the corresponding 95% posterior credible intervals (CrI), are provided in the second column of Table 1. By examining the posterior CrI, the patient's symptoms severity (oxygen support status) at baseline, blood type, a history of cardiovascular disease and a history of diabetes appear to be important predictors of HTE, as the CrI of the coefficients associated with these variables do not include 0.

In the first panel of Figure 3, we display the exponentiated individualized treatment effect, $\exp(\Delta(\boldsymbol{x}, \boldsymbol{\theta}))$, as a function of the single-index $\boldsymbol{x}^\top \boldsymbol{\beta}$. Specifically, the horizontal axis is the posterior mean of $\boldsymbol{x}^\top \boldsymbol{\beta}$ (a point estimate), where the "observed" posterior mean values $\boldsymbol{x}_i^\top \boldsymbol{\beta}$ $(i = 1, \ldots, n)$ $(n = 2287)$ computed from the posterior sample mean of $\boldsymbol{\beta}$, are represented by the small blue ticks along the axis. The uncertainty in the estimation of the coefficient $\boldsymbol{\beta}$ (as well as that of the coefficient $\boldsymbol{m}$) is also accounted for in the credible bands in Figure 3. For the sake of interpretability, we exponentiated the HTE estimate $\Delta(\boldsymbol{x}, \boldsymbol{\theta}) = g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 1) - g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 0)$, so that the vertical axis in the panel represents the odds ratio (CCP vs. control) for a bad outcome (mechanical ventilation or death). An odds ratio of less than 1 indicates a superior CCP efficacy over the control treatment. As most of the observed values $\boldsymbol{x}_i^\top \boldsymbol{\beta}$ of the single-index fall below the line representing the odds ratio of 1, most of the patients are expected to benefit from CCP treatment, except those with the $\boldsymbol{x}_i^\top \boldsymbol{\beta}$ values greater than 0.45, in which the corresponding expected individualized odds ratios are greater than 1 (about 28% of the observed patients). The U-shaped nonlinear relationship between the expected odds ratio and the single-index of the model suggests that the use of the flexible link function g in (3) is to be preferred over a more restricted linear model for this HTE modeling.

Table 1: Pretreatment patient characteristics $\boldsymbol{X}$ and the corresponding estimated index coefficients $\boldsymbol{\beta}$ (and 95% CrI)

Pretreatment characteristic x_j (sample prevalence, %)	Index coefficient β_j [95% CrI]
Oxygen support by mask or nasal prongs* (1/0) (63%)	0.68 [0.50, 0.80]
Oxygen support by high flow* (1/0) (18%)	0.47 [0.16, 0.61]
Age (dichotomized, ≥ 67) (1/0) (35%)	-0.13 [-0.46,0.04]
Blood type (A or AB vs. O or B) (1/0) (37%)	-0.31 [-0.49, -0.16]
Cardiovascular disease (1/0) (42%)	-0.24 [-0.65,-0.06]
Diabetes (1/0) (34%)	-0.26 [-0.52, -0.08]
Pulmonary disease (1/0) (12%)	0.05 [-0.16,0.22]

* The reference level: hospitalized but no oxygen therapy required.

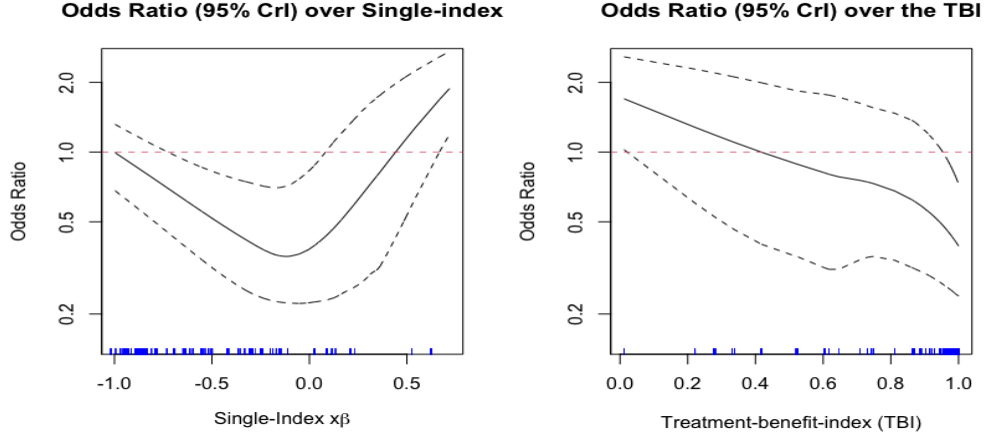**Odds Ratio (95% CrI) over Single-index**   **Odds Ratio (95% CrI) over the TBI**

Figure 3: The left panel displays the exponentiated version of the estimated individualized treatment effect, the posterior mean of $\Delta(\boldsymbol{x}) = g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 1) - g(\boldsymbol{x}^\top \boldsymbol{\beta}, A = 0)$ in (5) (solid curve), along with the corresponding upper and lower 95% credible interval (CrI) (dashed curves), as a function of the posterior mean of $\boldsymbol{x}^\top \boldsymbol{\beta}$. The right panel also displays the expected odds ratio (CCP vs. control) (solid curve) and the corresponding 95% CrI (dashed curves), but the horizontal axis is now the treatment-benefit-index (TBI) (6), $P(\exp(\Delta(\boldsymbol{x})) < 1 \mid \mathcal{D})$, where $\exp(\Delta(\boldsymbol{x}))$ represents the odds ratio, and the TBI probability is evaluated with respect to the posterior distribution of the parameters in Δ. The TBI provides a gradient of benefit that ranges from 0 to 1, with a higher value of the TBI indicating a greater benefit from the CCP treatment, compared to control. The observed values for the quantities on the horizontal axes are represented by the small blue ticks.

Although the first panel of Figure 3 displays an information about the relationship between the (exponentiated) individualized treatment effect $\exp(\Delta(\boldsymbol{x}))$ (i.e., the individualized odds ratio) and the posterior mean of the single-index $\boldsymbol{x}^\top \boldsymbol{\beta}$, this relationship is non-monotonic, which makes it difficult to construct a "gradient" of the treatment benefit of $A = 1$ vs. $A = 0$, as a function of the patient characteristics $\boldsymbol{x}$. Thus, in the second panel of Figure 3, we additionally display the individualized odds ratio $\exp(\Delta(\boldsymbol{x}))$, as a function of TBI$(\boldsymbol{x})$ defined in (6), i.e., TBI$(\boldsymbol{x}) = P(\exp(\Delta(\boldsymbol{x})) < 1|\mathcal{D})$, where the probability is evaluated with respect to the posterior distribution of the parameters involving Δ. As a probability, TBI$(\boldsymbol{x})$ ranges from 0 to 1, where larger values are associated with larger CCP benefit. For example, patients with a large value of TBI$(\boldsymbol{x})$ (i.e., TBI scores near 1) were expected to experience large, clinically meaningful benefits from CCP.

The second panel of Figure 3 displays a monotonically decreasing trend of the expected odds ratio (an increasing CCP benefit), as the TBI score increases from 0 to 1. Some portions of the expected odds ratio as well as the corresponding 95% CrI exceed 1 for very small TBI values (near 0), suggesting the possibility of harm from CCP as the TBI approaches 0, whereas the TBI values close to 1 indicate a substantial benefit from the CCP treatment over the control treatment. We can use the TBI scores to stratify patients according to their predicted treatment benefit levels, by setting the treatment decision rule $\hat{a}^*(\boldsymbol{x}) = \mathbb{I}_{(\text{TBI}(\boldsymbol{x})>0.5)} \in \{0, 1\}$.

To evaluate the performance of the treatment decision rule $\hat{a}^*(\boldsymbol{x})$, we randomly split the data with a ratio of 2:1 into a training set and a testing set (of size $\tilde{n}$), replicated 100 times, each time obtaining $\hat{a}^*(\boldsymbol{x})$ based on the training set, and the corresponding "value" $V(\hat{a}^*) = E[Y^{(\hat{a}^*(\boldsymbol{X}))}]$ by an inverse probability weighted estimator [Murphy, 2005] $\hat{V}(\hat{a}^*) = \frac{\sum_{i=1}^{\tilde{n}} Y_i \mathbb{I}_{(A_i = \hat{a}^*(\boldsymbol{X}_i))}}{\sum_{i=1}^{\tilde{n}} \mathbb{I}_{(A_i = \hat{a}^*(\boldsymbol{X}_i))}}$ computed based on the testing set (of size $\tilde{n}$).

For comparison, we also include two naive rules: treating all patients with Control ("All Control") and treating all patients with CCP ("All CCP"), each regardless of the individual patients' characteristics $\boldsymbol{x}$, in addition to the decision rules based on the Bayesian linear logistic model which was compared with the proposed index model in Section 3. The resulting boxplots obtained from the 100 random splits are illustrated in Figure 4.
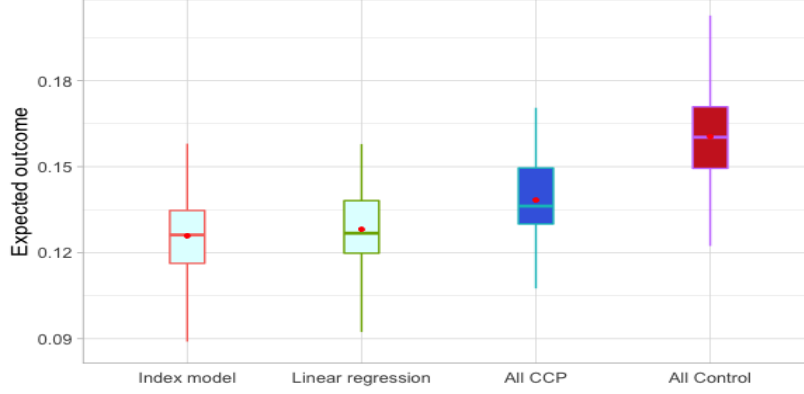
Figure 4: Boxplots of "value," obtained from 100 randomly split testing sets. A smaller "value" is desirable.

The results in Figure 4 demonstrate that the index model and the logistic linear regression perform at a similar level for this dataset, while showing a clear advantage over the naïve rules of giving everyone CCP or giving everyone the control treatment: the averaged proportion of patients with the undesirable outcome (i.e., $Y = 1$) was considerably less for the two regression approaches than the two naïve rules. This suggests that accounting for patient characteristics can help optimizing treatment decisions. Although some of the nonlinearities in the association between the treatment effect and patient characteristics $\boldsymbol{x}$ is captured by the model-implied odds ratio displayed in the first panel of Figure 3, for this dataset, the simpler linear model appears to perform nearly as well as the index model due to its model parsimony. However, as demonstrated in Section 3, the more flexible index model may be preferable to the linear model, as it allows for discovering some key nonlinearities in modeling heterogeneous treatment effects.

5   Discussion

The idea in the Bayesian estimation approach of Antoniadis et al. [2004] was to treat the link function g as another unknown and approximate it by a linear combination of B-spline basis functions. In this article, to estimate heterogeneous treatment effect using a flexible link function, we used the adjusted responses and weights associated with the iteratively re-weighted least squares (IWLS) algorithm in the quadratic approximation of the log-likelihood, for each MCMC sampler. The approximation under the IWLS framework and the specific prior choice (10) allows us to analytically integrate $\boldsymbol{\gamma}$ out of the approximated posterior (13), which simplifies the sampling procedure for $\boldsymbol{\beta}$. Although the sampling was done using approximated conditional posteriors, this approach appears to work reasonably well.

This paper focused on the context of a randomized clinical trial where the treatment $A_i \in \{0, 1\}$ is randomized independently of pretreatment characteristics $\boldsymbol{X}_i$. However, the method can be potentially extended to the case where the treatment assignment depends on $\boldsymbol{X}_i$. To estimate individual treatment effects with observational or non-fully randomized data, we can take a "propensity method" [see, e.g., Imbens and Rubin, 2015, Caron et al., 2022] upon taking an appropriate reparametrization of the proposed model, which we describe below for a more general context of k treatment conditions, in which the treatment A_i takes a value $a \in \{1, \ldots, k\}$ with probability (i.e., propensity score) $P(A_i = a | \boldsymbol{X}_i) = \pi_a(\boldsymbol{X}_i)$ $(a = 1, \ldots, k)$. Let $a = 1$ be the reference (control) treatment.

For each fixed $\boldsymbol{\beta}$, the condition (4) implies: $E[g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, A_i) | \boldsymbol{X}_i] = \sum_{a=1}^k g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) \pi_a(\boldsymbol{X}_i) = 0$, or equivalently, $g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a = 1) = -\sum_{a=1}^k g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) \frac{\pi_a(\boldsymbol{X}_i)}{\pi_1(\boldsymbol{X}_i)}$. Given this representation for $g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a = 1)$, we can reparametrize the canonical parameter of model (3), that is, $\eta_i(\boldsymbol{X}_i, A_i) = \boldsymbol{m}^\top \boldsymbol{X}_i + \sum_{a=1}^k \mathbb{I}_{(A_i = a)} g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a)$, by

$$
\begin{aligned}
\eta_i(\boldsymbol{X}_i, A_i) &= \boldsymbol{m}^\top \boldsymbol{X}_i + \sum_{a=2}^k \mathbb{I}_{(A_i = a)} g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) - \mathbb{I}_{(A_i = 1)} \sum_{a=2}^k g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) \frac{\pi_a(\boldsymbol{X}_i)}{\pi_1(\boldsymbol{X}_i)} \\
&= \boldsymbol{m}^\top \boldsymbol{X}_i + \sum_{a=2}^k g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) w_a(A_i, \boldsymbol{X}_i)
\end{aligned}
\tag{20}
$$

where $w_a(A_i, \boldsymbol{X}_i) = \mathbb{I}_{(A_i = a)} - \frac{\pi_a(\boldsymbol{X}_i)}{\pi_1(\boldsymbol{X}_i)} \mathbb{I}_{(A_i = 1)}$. This parametrization is an unconstrained formulation of model (3) without the constraint (4), where the propensity score $\pi_a(\boldsymbol{X}_i)$ $(a = 1, \ldots, k)$ is incorporated through the subject i- and treatment a-specific weight $w_a(A_i, \boldsymbol{X}_i)$ in the formulation. In model (20), the interaction term $g^*(\boldsymbol{X}_i, A_i) :=$

$\sum_{a=2}^{k} g(\boldsymbol{X}_i^\top \boldsymbol{\beta}, a) w_a(A_i, \boldsymbol{X}_i)$ still satisfies the condition of the form (4), since $E[w_a(A_i, \boldsymbol{X}_i)|\boldsymbol{X}_i] = 0$, indicating that $E[g^*(\boldsymbol{X}_i, A_i))|\boldsymbol{X}_i] = 0$.

In the estimation, as in Section 2.2.2 for the binary treatment condition $A_i \in \{0, 1\}$, we can proceed as follows for the general k treatment conditions with $A_i \in \{1, \ldots, k\}$ and also in the context of an observational or a non-fully randomized study. We can define the $n \times L(k-1)$ design matrix $\boldsymbol{D}_{\boldsymbol{\beta}} = (\boldsymbol{D}_{\boldsymbol{\beta},2}; \ldots; \boldsymbol{D}_{\boldsymbol{\beta},k})$, where each element (the $n \times L$ matrix) $\boldsymbol{D}_{\boldsymbol{\beta},a}$ $(a = 2, \ldots, k)$ that is specific to each treatment condition $A = a$ $(a = 2, \ldots, k)$ denotes the evaluation matrix of the basis $\tilde{\psi}(\cdot)$ on $\{\boldsymbol{\beta}^\top \boldsymbol{X}_i\}_{i=1}^n$ multiplied by the subject i- and treatment condition a-specific weight $w_{ia} = w_a(A_i, \boldsymbol{X}_i)$(the weight is defined on the second line of (20)), so that its ith row corresponds to the $1 \times L$ vector, $w_{ia}\tilde{\psi}(\boldsymbol{\beta}^\top \boldsymbol{X}_i)^\top$, with the weight w_{ia} incorporating the pre-estimated treatment propensity score $\pi_a(\boldsymbol{X})$ $(a = 1, \ldots, k)$. The spline coefficient vector $\boldsymbol{\gamma} \in \mathbb{R}^{L(k-1)}$ associated with the design matrix $\boldsymbol{D}_{\boldsymbol{\beta}}$ can be introduced, which yields the representation $\boldsymbol{g} = \boldsymbol{D}_{\boldsymbol{\beta}}\boldsymbol{\gamma} \in \mathbb{R}^n$ as in Section 2.2.2, and the same estimation procedure of Section 2.2.3 and 2.3 can be employed to conduct posterior inference.

Future work will extend the model to accommodate multiple treatment outcomes to allow for borrowing of strength between the available outcomes in modeling heterogeneous treatment effects.

Acknowledgements

References

S. A. Murphy. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355, May 2003.

J. Robins. *Optimal Structural Nested Models for Optimal Sequential Decisions*. Springer, New York, 2004.

Francis S. Collins and Harold Varmus. A new initative on precision medicine. *New England Journal of Medicine*, 372 (9):793–795, 2015.

M. Qian and S. A. Murphy. Performance guarantees for individualized treatment rules. *The Annals of Statistics*, 39(2): 1180–1210, 2011.

W. Lu, H. Zhang, and D. Zeng. Variable selection for optimal treatment decision. *Statistical Methods in Medical Research*, 22:493–504, 2011.

L. Tian, A. Alizadeh, A. Gentles, and R. Tibshrani. A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*, 109(508):1517–1532, 2014.

C. Shi, R. Song, and W. Lu. Robust learning for optimal treatment decision with np-dimensionality. *Electronic Journal of Statistics*, 10:2894–2921, 2016.

X. Jeng, W. Lu, and H. Peng. High-dimensional inference for personalized treatment decision. *Electronic Journal of Statistics*, 12:2074–2089, 2018.

Y. Zhao, D. Zeng, A. J. Rush, and M. R. Kosorok. Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107:1106–1118, 2012.

Y. Zhao, D. Zheng, E. B. Laber, and M. R. Kosorok. New statistical learning methods for estimating optimal dynamic treatment regimes. *Journal of the American Statistical Association*, 110:583–598, 2015.

R. Song, M. Kosorok, D. Zeng, Y. Zhao, E. B. Laber, and M. Yuan. On sparse representation for optimal individualized treatment selection with penalized outcome weighted learning. *Stat*, 4:59–68, 2015.

E. B. Laber and Y. Zhao. Tree-based methods for individualized treatment regimes. *Biometrika*, 102:501–514, 2015.

E. B. Laber and A. Staicu. Functional feature construction for individualized treatment regimes. *Journal of the American Statistical Association*, 113:1219–1227, 2018.

Y. Zhao, E. Laber, Y. Ning, S. Saha, and B. Sands. Efficient augmentation and relaxation learning for individualized treatment rules using observational data. *Journal of Machine Learning Research*, 20:1–23, 2019.

Y. Liu, Y. Wang, M. R. Kosorok, Y. Zhao, and D. Zeng. Augmented outcome-weighted learning for estimating optimal dynamic treatment regimens. *Statistics in Medicine*, 37:3776–3788, 2018.

S. Xie, T. Tarpey, E. Petkova, and R. T. Ogden. Multiple domain and multiple kernel outcome- weighted learning for estimating individualized treatment regimes. *Journal of Computational and Graphical Statistics*, 2022.

A. Caron, G. Baio, and I. Manolopoulou. Estimating individual treatment effects using non-parametric regression models: a review. *Journal of Royal Statistical Society: Series A*, 185:1115–1149, 2022.

X. Nie and S. Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108:299–319, 2021.

A. M. Alaa and M. van der Schaar. Bayesian inference of individualized treatment effects using multi-task gaussian processes. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, pages 3427–3435, 2017.

Uri Shalit, Fredrik D. Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. *Proceedings of the 34th International Conference on Machine Learning*, 70:3076–3085, 2017.

Stefan Wager and Susan Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113:1228–1242, 2018.

W. Guo, X. ZHou, and S. Ma. Estimation of optimal individualized treatment rules using a covariate-specific treatment effect curve with high-dimensional covariates. *Journal of the American Statistical Association*, 116(533):309–321, 2022.

P. Richard Hahn, Jared S. Murray, and Carlos M. Carvalho. Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects. *Bayesian Analysis*, 15(3):965–1056, 2020.

T. Klausch, Peter van de Ven, Tim van de Brug, Mark A. van de Wiel, and Johannes Berkhof. Estimating bayesian optimal treatment regimes for dichotomous outcomes using observational data. *arXiv*, 2018.

T. Murray, Y. Yuan, and P. F. Thall. A bayesian machine learning approach for optimizing dynamic treatment regimes. *Journal of the American Statistical Association*, 113(523):1255–1267, 2018.

B. R. Logan, R. Sparapani, R. E. McCulloch, and P. W. Laud. Decision making and uncertainty quantification for individualized treatments using bayesian additive regression trees. *Statistical Methods in Medical Research*, 28(4): 1079–1093, 2019.

R. David Brillinger. *A generalized linear model with "Gaussian" regressor variables In A Festschrift for Erich L. Lehman (Edited by P. J. Bickel, K. A. Doksum and J. L. Hodges)*. Wadsworth, New York, 1982.

T. M. Stoker. Consistent estimation of scaled coefficients. *Econometrica*, 54:1461–1481, 1986.

A. Antoniadis, G. Gregoire, and I. McKeague. Bayesian estimation in single-index models. *Statistica Sinica*, 14: 1147–1164, 2004.

W. Poon and H. Wang. Bayesian analysis of generalized partially linear single-index models. *Computational Statistics and Data Analysis*, 68:251–261, 2013.

W. Poon and H. Wang. Multivariate partially linear single-index models: Bayesian analysis. *Journal of Nonparametric Statistics*, 26(4):755–768, 2014.

H. Wang. Bayesian estimation and variable selection for single index models. *Computational Statistics and Data Analysis*, 53:2617–2627, 2009.

H. Wang. A bayesian multivariate partially linear single-index probit model for ordinal responses. *Journal of Statistical Computation and Simulation*, 88:1616–1636, 2018.

T. Choi, J. Shi, and B. Wang. A gaussian process regression approach to a single-index model. *Jounral of Nonparametric Statistics*, 23:21–36, 2011.

R. B. Gramacy and H. Lian. Gaussian process single-index models as emulators for computer experiments. *Technometrics*, 54(1):30–41, 2012.

Y. Yu, Z. Zou, S. Wang, and R. Meyer. Bayesian nonparametric modelling of the link function in the single-index model using a bernstein–dirichlet process prior. *Journal of Statistical Computation and Simulation*, 89:3290–3312, 2019.

Kumaresh Dhara, Stuart Lipsitz, Debdeep Pati, and Debajyoti Sinha. A new bayesian single index model with or without covariates missing at random. *Bayesian Analysis*, 15(3):759–780, 2020.

C. Liu and H. Liang. Bayesian analysis in single-index quantile regression with missing observation. *Communications in Statistics - Theory and Methods*, 2022.

Andrea B. Troxel, Eva Petkova, Keith Goldfeld, Mengling Liu, Thaddeus Tarpey, Yinxiang Wu, Danni Wu, Anup Agarwal, Cristina Avendaño-Solá, Emma Bainbridge, Katherine J. Bar, Timothy Devos, Rafael F. Duarte, Arvind Gharbharan, Priscilla Y. Hsue, Gunjan Kumar, Annie F. Luetkemeyer, Geert Meyfroidt, André M. Nicola, Aparna Mukherjee, Mila B. Ortigoza, Liise-anne Pirofski, Bart J. A. Rijnders, Casper Rokx, Arantxa Sancho-Lopez, Pamela Shaw, Pablo Tebas, Hyun Ah Yoon, Corita Grudzen, Judith Hochman, and Elliott M. Antman. Association of Convalescent Plasma Treatment With Clinical Status in Patients Hospitalized With COVID-19: A Meta-analysis. *JAMA Network Open*, 5(1):e2147331–e2147331, 2022. ISSN 2574-3805. doi:10.1001/jamanetworkopen.2021.47331. URL https://doi.org/10.1001/jamanetworkopen.2021.47331.

G. W. Imbens and D. B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

D. B. Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Paul Eilers and Brian Marx. Flexible smoothing with B-splines and penalties. *Statistical Science*, 11(2):89–121, 1996.

Li Wang and Lijian Yang. Spline estimation of single-index models. *Statistica Sinica*, 19:765–783, 2009.

M.P. Wand and J. T. Ormerod. Penalized wavelets: Embedding wavelets into semiparametric regression. *Electronic Journal of Statistics*, 5:1654–1717, 2011.

P. J. Lenk. Bayesian inference for semiparametric regression using a Fourier representation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61:863–879, 2002.

K. Hornik and B. Grun. movMF: Mixtures of von Mises-Fisher distributions. *R package version 0.2-7*, 2022.

A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4): 457–472, 1992.

Dootika Vats and Christina Knudson. Revisiting the Gelman–Rubin diagnostic. *Statistical Science*, 36(4):518 – 529, 2021.

Hyung Park, Thaddeus Tarpey, Mengling Liu, Keith Goldfeld, Yinxiang Wu, Danni Wu, Yi Li, Jinchun Zhang, Dipyaman Ganguly, Yogiraj Ray, Shekhar Ranjan Paul, Prasun Bhattacharya, Artur Belov, Yin Huang, Carlos Villa, Richard Forshee, Nicole C. Verdun, Hyun ah Yoon, Anup Agarwal, Ventura Alejandro Simonovich, Paula Scibona, Leandro Burgos Pratx, Waldo Belloso, Cristina Avendaño-Solá, Katharine J Bar, Rafael F. Duarte, Priscilla Y. Hsue, Anne F. Luetkemeyer, Geert Meyfroidt, André M. Nicola, Aparna Mukherjee, Mila B. Ortigoza, Liise-anne Pirofski, Bart J. A. Rijnders, Andrea Troxel, Elliott M. Antman, and Eva Petkova. Development and Validation of a Treatment Benefit Index to Identify Hospitalized Patients With COVID-19 Who May Benefit From Convalescent Plasma. *JAMA Network Open*, 5(1):e2147375–e2147375, 2022. ISSN 2574-3805. doi:10.1001/jamanetworkopen.2021.47375. URL https://doi.org/10.1001/jamanetworkopen.2021.47375.

S. A. Murphy. A generalization error for q-learning. *Journal of Machine Learning*, 6:1073–1097, 2005.

Appendix

As MCMC convergence diagnostic for the simulation example in Section 3, we report the Gelman-Rubin potential scale reduction factor (PSRF) computed for each scenario.

Shape of the component m	Shape of the component g	p	n	Pr(GR <1.03)			Pr(GR <1.05)			Pr(GR <1.07)		
				β	γ	m	β	γ	m	β	γ	m
linear	nonlinear	5	500	0.856	1.000	1.000	0.970	1.000	1.000	0.998	1.000	1.000
			1000	0.908	0.999	1.000	0.986	1.000	1.000	0.998	1.000	1.000
			2000	0.916	1.000	1.000	0.984	1.000	1.000	1.000	1.000	1.000
		10	500	0.602	0.997	1.000	0.950	1.000	1.000	1.000	1.000	1.000
			1000	0.744	1.000	1.000	0.981	1.000	1.000	1.000	1.000	1.000
			2000	0.803	1.000	1.000	0.988	1.000	1.000	1.000	1.000	1.000
	linear	5	500	0.844	1.000	1.000	0.984	1.000	1.000	0.998	1.000	1.000
			1000	0.902	1.000	1.000	0.984	1.000	1.000	1.000	1.000	1.000
			2000	0.930	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		10	500	0.485	0.999	1.000	0.925	1.000	1.000	0.999	1.000	1.000
			1000	0.689	0.997	1.000	0.968	0.999	1.000	1.000	1.000	1.000
			2000	0.758	1.000	1.000	0.977	1.000	1.000	1.000	1.000	1.000
nonlinear	nonlinear	5	500	0.950	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
			1000	0.970	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			2000	0.956	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		10	500	0.750	0.996	1.000	0.991	1.000	1.000	1.000	1.000	1.000
			1000	0.888	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
			2000	0.933	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	linear	5	500	0.922	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			1000	0.978	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
			2000	0.968	1.000	1.000	0.996	1.000	1.000	1.000	1.000	1.000
		10	500	0.734	0.999	1.000	0.986	1.000	1.000	1.000	1.000	1.000
			1000	0.884	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
			2000	0.936	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A1: Summary of the Gelman-Rubin (GR) diagnostic over 100 simulation replications for each scenario. The GR potential scale reduction factor (PSRF) was computed for each element of the components β, γ, and m. For each of these components, we computed the proportion of the individual elements' PSRF statistics being less than a certain threshold (1.03, 1.05, 1.07) by averaging the PSRF statistics across the corresponding component's elements and across the 100 simulation replications (for example, for the component $\beta \in \mathbb{R}^p$, the proportion was computed based on $100 * p$ PSRF statistics). The results for β indicate that almost all PSRF statistics were less than 1.07, while those of γ and m were consistently smaller than 1.03 for almost all the scenarios, providing some assurance that the sampler has performed reasonably.