Biometrics 59, ??–??
December 2019
DOI: 10.1111/j.1541-0420.2005.00454.x

A constrained single-index regression for estimating interactions between a treatment and covariates

Hyung Park*, Eva Petkova, Thaddeus Tarpey*****

Division of Biostatistics, Department of Population Health, New York University, New York, NY 10016, USA

**email:* parkh15@nyu.edu

***email:* eva.petkova@nyumc.org

****email:* thaddeus.tarpey@nyumc.org

and

R. Todd Ogden*

Department of Biostatistics, Columbia University, New York, NY 10032, USA

**email:* to166@columbia.edu

Summary: We consider a single-index regression model, uniquely constrained to estimate interactions between a set of pretreatment covariates and a treatment variable on their effects on a response variable, in the context of analyzing data from randomized clinical trials. We represent interaction effect terms of the model through a set of treatment-specific flexible link functions on a linear combination of the covariates (a single index), subject to the constraint that the expected value given the covariates equals zero, while leaving the main effects of the covariates unspecified. We show that the proposed semiparametric estimator is consistent for the interaction term of the model, and that the efficiency of the estimator can be improved with an augmentation procedure. The proposed single-index regression provides a flexible and interpretable modeling approach to optimizing individualized treatment rules based on patients' data measured at baseline, as illustrated by simulation examples and an application to data from a depression clinical trial.

Key words: Individualized treatment rule; Precision medicine; Projection-pursuit regression; Single-index model; Treatment effect modifier

Accepted Article

This paper has been submitted for consideration for publication in *Biometrics*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/biom.13320

This article is protected by copyright. All rights reserved.

Accepted Article

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 (**??**). Furthermore, estimating treatment and pretreatment covariate interactions in the setting of randomized clinical trials may provide valuable information for understanding the factors involved in heterogeneous treatment responses. In this paper, we propose a simple and flexible regression method specifically focused on estimating the interaction effects between a treatment variable and pretreatment covariates on a treatment response.

Since the seminal papers of **?** and **?**, much research has been done on development of individualized treatment rules based on pretreatment covariates. Regression-based methodologies are intended to optimize the individualized treatment rules by estimating treatment-specific mean response functions (e.g., **????**) while attempting to maintain robustness with respect to model misspecification. Machine learning approaches for developing individualized treatment rules are often framed in the context of classification problems (**??**); for example, the outcome weighted learning (e.g., **???**) based on support vector machines, tree-based classification (e.g., **?**), and the methods of **?** based on adaptive boosting, among others. Although the classification approaches are appealing in many settings, here we focus on familiar regression approaches that are frequently utilized in practice, and allow for ready interpretation.

? show that the optimal individualized treatment rules (in terms of maximizing the mean treatment response) depends only on the treatment and pretreatment covariates' interaction effects, and not on the main effects of the pretreatment covariates present in the mean response function. However, if the model inadequately represents the interaction effects, the estimated individualized treatment rule may perform poorly (**??**). The primary focus of this paper is to develop a semiparametric regression method for estimating the interaction effect

This article is protected by copyright. All rights reserved.

term of the mean response function, which reduces concerns regarding misspecification of the interaction effects.

? approximate the mean response function using a rich linear model with a penalized least squares criterion. However, this approach is generally not robust to misspecification of the main effect term of the model and is also restricted to a parametric regression. Also, if the main effect is responsible for a relatively large proportion of the variance in the outcome compared to the interaction effect, consistent estimation of the interaction effect is difficult. Addressing this issue, **?** proposed an approach to consistently estimate the treatment-by-covariate interaction effect without having to specify the main effect. However, their approach is applicable only to the (generalized) linear model framework and only when there are exactly two treatments. In realistic situations, a linear model may be too restrictive to describe complex interactions. **?** proposed a robust approach to estimating an optimal individualized treatment rule, within a class of rules defined by a (possibly misspecified) regression model. However, their method is computationally feasible only in low dimensional settings. **?** proposed a semiparametric regression model to estimate an optimal individualized treatment rule, but their approach is limited to a monotone interaction effect structure and to binary treatment conditions.

A primary contribution of this paper is in generalizing the work of **?** to allow for a semiparametrically defined interaction effect and also for more than two treatments in the context of randomized clinical trials. We do this by extending a single-index model (e.g., **?**) to allow treatment-specific nonparametric link functions (**?**) in order to capture the treatment-by-covariates interaction effects, while allowing for an unspecified main effect of the covariates. The result is a simple and flexible regression model for the interaction effects.

Accepted Article

This article is protected by copyright. All rights reserved.

2. Constrained single-index models

In the context of randomized clinical trials, we consider pretreatment covariates $\boldsymbol{X} \in \mathbb{R}^p$ and a categorical (factor-valued) treatment variable $T \in \{1, \ldots, L\}$ (with L levels) that has associated randomization probabilities $(\pi_1, \ldots, \pi_L)$. We let $Y^{(t)} \in \mathbb{R}$ (t = 1,...,L) be the potential outcome if the patient received treatment $T = t$ $(t = 1, \ldots, L)$; we only observe $Y = Y^{(T)}$, T and $\boldsymbol{X}$. Throughout the paper, we assume, without loss of generality, that $E[Y|T = t] = 0$ $(t = 1, \ldots, L)$, i.e., the main effect for T is centered at 0 (this is only to suppress the treatment-specific intercepts in regression models in order to simplify the exposition, and can be achieved by removing the treatment level t-specific means from Y) and that $\boldsymbol{X}$ is centered at zero.

The focus of this paper is on modeling interactions between $\boldsymbol{X}$ and T on their effects on Y. We assume $Y = E[Y|\boldsymbol{X}, T] + \epsilon$, where ϵ is a zero-mean independent noise with finite variance. Let us assume that the nested mean model associated with the interaction effects of the mean response function $E[Y|\boldsymbol{X}, T]$ has a single-index model structure with a set of treatment t-specific link functions, for a single-index coefficient $\boldsymbol{\alpha}_0 \in \mathbb{R}^p$:

$$E[Y \mid \boldsymbol{X}, T = t] = \underbrace{\mu(\boldsymbol{X})}_{\text{main effect}} + \underbrace{f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})}_{\text{interaction}} \quad (t = 1, \ldots, L), \tag{1}$$

where $\mu(\boldsymbol{X})$ represents an unspecified main effect of $\boldsymbol{X}$. In model (1), the treatment t-specific functions $f_t(\cdot)$ are general smooth univariate functions. To obtain an identifiable representation, without loss of generality, the treatment t-specific functions $(f_1, \ldots, f_L)$ in model (1) are assumed to satisfy a condition: $E[f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})|\boldsymbol{X}] = \sum_{t=1}^{L} \pi_t f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) = 0$ (almost surely). This condition indicates that there are only $L - 1$ unrestricted functions f_t among the L interaction functions $(f_1, \ldots, f_L)$; that is, the Lth function f_L in (1) is identified by the other $(L - 1)$ functions: $f_L(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) = -\pi_L^{-1} \sum_{t=1}^{L-1} \pi_t f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$ (almost surely).

In model (1), the single-index coefficient $\boldsymbol{\alpha}_0$ is identifiable only up to scale and sign due to the nonparametric nature of the link functions f_t $(t = 1, \ldots, L)$ and therefore, without loss

Accepted Article

This article is protected by copyright. All rights reserved.

of generality, we assume $\boldsymbol{\alpha}_0 \in \Theta$, where

$$\Theta := \{\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)^\top \in \mathbb{R}^p : \|\boldsymbol{\alpha}\| = 1, \alpha_1 > 0\}.$$

The semiparametric model (1) captures the variability in $\boldsymbol{X}$ related to the treatment effects via a single index $\boldsymbol{\alpha}_0^\top \boldsymbol{X} \in \mathbb{R}$ and its interactions with the treatment via treatment-specific link functions $(f_1, \ldots, f_L)$. Interaction effects are determined by the distinct shapes of the unspecified functions $(f_1, \ldots, f_L)$. There are several reasons we consider a single-index $\boldsymbol{\alpha}_0^\top \boldsymbol{X}$ in (1) (as opposed to considering treatment-specific L indices). First, the common single index provides a parsimonious 1-dimensional composite treatment effect modifier (defined as a linear combination of $\boldsymbol{X}$) that allows an intuitive visualization for the interaction effect, which can be used to provide clinicians a simple tool to make treatment decisions. Besides its parsimonious appeal, the single-dimensional reduction model (1) naturally and directly extends the linear model-based approach (e.g., **?**) to modeling T-by-$\boldsymbol{X}$ interactions in the $L = 2$ setting. If we consider only two treatments ($L = 2$) and we restrict the unspecified interaction function $f_t(\cdot)$ in model (1) to a prespecified linear form $f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) = (t+\pi_1-2)\boldsymbol{\alpha}_0^\top \boldsymbol{X}$, then the semiparametric model (1) reduces to the modified covariates model assumed in **?** (see also **?????**):

$$E\left[Y \mid \boldsymbol{X}, T = t\right] \;=\; \mu(\boldsymbol{X}) \;+\; \boldsymbol{\alpha}_0^\top \boldsymbol{X}(t + \pi_1 - 2) \quad (t = 1, 2), \tag{2}$$

for some $\boldsymbol{\alpha}_0 \in \mathbb{R}^p$, which assumes a linear form for the T-by-$\boldsymbol{X}$ interaction effects.

For model (1), to estimate the interaction effect terms $f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$ in the presence of the unspecified main effect $\mu(\boldsymbol{X})$, we propose to utilize a working (approximation) model:

$$E[Y \mid \boldsymbol{X}, T = t] \;\approx\; g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) \quad (t = 1, \ldots, L), \tag{3}$$

for $\boldsymbol{\alpha} \in \Theta$, subject to the constraint:

$$E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{X}] = \sum_{t=1}^{L} \pi_t g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = 0 \quad \text{(almost surely)}, \tag{4}$$

Accepted Article

This article is protected by copyright. All rights reserved.

for all $\boldsymbol{\alpha} \in \Theta$. The constraint (4) is imposed on the treatment t-specific smooth link-functions $(g_1, \ldots, g_L)$ of the working model (3). Even if (3) does not generally provide a good approximation to the underlying model (1), in Section 3.3 we will show through the consistency results (Theorem 1 and Corollary 1) that (3) is a useful model for estimating the interaction effect terms $f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$ of model (1).

In a least squares framework of optimizing the working model (3)

$$\begin{aligned} E\big[\big(Y - g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})\big)^2/2\big] \;&\propto\; E\big[Y g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2\big] \quad \text{with respect to } \boldsymbol{\alpha} \text{ and } (g_1, \ldots, g_L) \\ &=\; E\left[\{\mu(\boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})\} g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2\right] \\ &=\; E\big[E\big[\mu(\boldsymbol{X}) g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2 \mid \boldsymbol{X}\big]\big], \end{aligned} \tag{5}$$

the condition (4) ensures that the cross-product term $E\big[\mu(\boldsymbol{X}) g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) | \boldsymbol{X}\big]$ vanishes, and the part relevant to the estimation of the working model (3) is independent of the unspecified main effect $\mu(\boldsymbol{X})$ which is a potentially complicated function. This independence implies that optimization of the unspecified component $\mu(\boldsymbol{X})$ and the working model (3) can be performed separately. As model (3) does not require specification of $\mu(\boldsymbol{X})$, working with (3) sidesteps issues that would arise if $\mu(\boldsymbol{X})$ were to be misspecified.

We call model (3) a *constrained* single-index model with multiple (i.e., treatment t-specific) link functions, which is the main working model of this paper.

3. Estimation

3.1 *A criterion for fitting the model*

To optimize the constrained working model (3), we consider a constrained least squares criterion:

$$\begin{aligned} &\underset{\boldsymbol{\alpha}, (g_1, \ldots, g_L)}{\text{minimize}} \quad E\big[\big(Y - g_T\big(\boldsymbol{\alpha}^\top \boldsymbol{X}\big)\big)^2/2\big] \\ &\text{subject to} \quad E\left[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{X}\right] = 0 \quad \text{(almost surely)}, \end{aligned} \tag{6}$$

where $\boldsymbol{\alpha} \in \Theta$ and each function g_t is in a suitable function space in $L^2(R)$.

Accepted Article

This article is protected by copyright. All rights reserved.

PROPOSITION 1: For each fixed $\boldsymbol{\alpha}$, the minimizer $(g_1, \ldots, g_L)$ of the constrained minimization problem (6) satisfies:

$$g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T = t\right] - E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}\right] \quad (t = 1, \ldots, L), \tag{7}$$

almost surely.

The proof of Proposition 1 is in Section A.1 of the Supporting Information. Proposition 1 suggests that solving (6) to optimize model (3) can be split into the following two iterative steps. First, for a fixed $\boldsymbol{\alpha}$, find the link-functions $(g_1, \ldots, g_L)$ from expression (7). Second, for a fixed $(g_1, \ldots, g_L)$, solve

$$\underset{\boldsymbol{\alpha}\in\Theta}{\operatorname{argmin}}\ E\left[\left(Y - g_T\left(\boldsymbol{\alpha}^\top \boldsymbol{X}\right)\right)^2/2\right]. \tag{8}$$

These two steps can be iterated until convergence to obtain a solution of (6). To obtain a sample counterpart of this population solution, we can insert sample estimates into this population algorithm, as is done in **?**. We will provide the details of this estimation procedure in Section 3.2, and establish the consistency of the estimator in Section 3.3.

3.2 *A cubic spline estimator of the model*

To obtain a sample counterpart of the population solution of (6), we approximate the objective function of (6) based on sample $\{(Y_i, T_i, \boldsymbol{X}_i), i = 1, \ldots, n\}$ (assumed to be independently and identically distributed across i). In particular, we use a nonparametric regression technique to approximate the solution $(g_1, \ldots, g_L)$ in (7) for each fixed $\boldsymbol{\alpha} \in \Theta$. Although other nonparametric regression methods can also be used, in this paper we focus on a cubic B-spline (**?**) representation of the solution $(g_1, \ldots, g_L)$ in (7) for each fixed $\boldsymbol{\alpha}$. Specifically, given each $\boldsymbol{\alpha}$, the functions g_t is represented by:

$$g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = B_t(\boldsymbol{\alpha}^\top \boldsymbol{X})^\top \boldsymbol{\beta}_t \quad (t = 1, \ldots, L), \tag{9}$$

for some vector $\boldsymbol{\beta}_t \in \mathbb{R}^{d_t+4}$, where $B_t(\cdot) \in \mathbb{R}^{d_t+4}$ is a set of $(d_t + 4)$ cubic B-spline basis functions (**?**) defined on the range of the candidate single-index $\{(\boldsymbol{\alpha}^\top \boldsymbol{X}_i), i = 1, \ldots, n\}$. We

Accepted Article

This article is protected by copyright. All rights reserved.

use d_t to denote the number of interior knots (placed with equal distance between neighboring knots). The number d_t depends on the treatment group sample size: $n_t = \sum_{i=1}^{n} 1_{(T_i=t)}$. Furthermore, let us represent the conditional expectation $E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}\right]$ in expression (7) by $E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}\right] = B_0(\boldsymbol{\alpha}^\top \boldsymbol{X})^\top \boldsymbol{\beta}_0$ for some vector $\boldsymbol{\beta}_0 \in \mathbb{R}^{d_0+4}$, where $B_0(\cdot) \in \mathbb{R}^{d_0+4}$ is a set of (d_0+4) cubic B-spline basis functions defined on the range of the candidate single-index $\{(\boldsymbol{\alpha}^\top \boldsymbol{X}_i), i = 1, \ldots, n\}$. We use d_0 number of interior knots (placed with equal distance between neighboring knots), and the number d_0 depends on the sample size n.

Let $\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)}$ $(t = 1, \ldots, L)$ denote the treatment t-specific $n \times d_t$ design matrix, where the ith row is the $1 \times d_t$ vector $B_t(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)^\top$ if $T_i = t$ and is a row of zeros if $T_i \neq t$ $(i = 1, \ldots, n)$ $(t = 1, \ldots, L)$. Let $\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)}$ denote the $n \times d_0$ design matrix in which the ith row is the $1 \times d_0$ vector $B_0(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)^\top$ $(i = 1, \ldots, n)$. Then, for each fixed $\boldsymbol{\alpha} \in \Theta$, we approximate the minimizer $(g_1, \ldots, g_L)$ in (7) by the method of least squares (see Section A.2 of the Supporting Information for a derivation):

$$g_t(\cdot) = B_t(\cdot)^\top (\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top} \boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)})^{-1} \boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top} \big(\boldsymbol{I}_n - \boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)} (\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top} \boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)})^{-1} \boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\big) \boldsymbol{Y} \quad (t = 1, \ldots, L), \quad (10)$$

where $\boldsymbol{Y}$ denotes the $n \times 1$ vector of the observed responses. We define the estimator $\widehat{\boldsymbol{\alpha}}_0$ for $\boldsymbol{\alpha}_0$ of model (1) by

$$\widehat{\boldsymbol{\alpha}}_0 \;=\; \underset{\boldsymbol{\alpha} \in \Theta}{\operatorname{argmin}}\; n^{-1} \sum_{i=1}^{n} \big(Y_i - g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)\big)^2 / 2, \quad (11)$$

where $(g_1, \ldots, g_L)$ are given in (10). Let us define the associated estimators $(\widehat{f}_1, \ldots, \widehat{f}_L)$ for the treatment-specific functions $(f_1, \ldots, f_L)$ of model (1) by

$$(\widehat{f}_1, \ldots, \widehat{f}_L) \;=\; (g_1, \ldots, g_L) \text{ given in (10), computed at } \boldsymbol{\alpha} = \widehat{\boldsymbol{\alpha}}_0. \quad (12)$$

We use iteratively re-weighted least squares to solve (11), repeating the following two steps:

1. Given a current estimate $\boldsymbol{\alpha} \in \Theta$, compute the functions $(g_1, \ldots, g_L)$ in (10).

Accepted Article

This article is protected by copyright. All rights reserved.

2. Given $(g_1, \ldots, g_L)$, approximately solve (11), based on a linear approximation to $g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)$ $(i = 1, \ldots, n)$ at the current estimate of $\boldsymbol{\alpha}$.

The iteration between the two steps continues until convergence of $\boldsymbol{\alpha} \in \Theta$. We next elaborate Step 2 of this iterative procedure. For each (the kth) iterative step, the objective function in (11) is approximated based on a linear approximation of $g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)$ at the current (the kth) iterate, say $\boldsymbol{\alpha}^{(k)} \in \Theta$:

$$\begin{aligned} n^{-1} \sum_{i=1}^{n} \left(Y_i - g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)\right)^2 /2 \;&\approx\; n^{-1} \sum_{i=1}^{n} \left(Y_i - g_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i) - \dot{g}_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i)(\boldsymbol{\alpha} - \boldsymbol{\alpha}^{(k)})^\top \boldsymbol{X}_i\right)^2 /2 \\ &=\; n^{-1} \sum_{i=1}^{n} \left(Y_i^{*(k)} - \boldsymbol{\alpha}^\top \boldsymbol{X}_i^{*(k)}\right)^2 /2, \end{aligned} \tag{13}$$

where the modified responses $Y_i^{*(k)}$ and the modified regressors $\boldsymbol{X}_i^{*(k)}$:

$$\begin{aligned} Y_i^{*(k)} &= Y_i - g_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i) + \dot{g}_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i) \boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i \qquad (i = 1, \ldots, n) \\ \boldsymbol{X}_i^{*(k)} &= \dot{g}_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i) \boldsymbol{X}_i \qquad (i = 1, \ldots, n) \end{aligned} \tag{14}$$

and we minimize the right-hand side of (13) over $\boldsymbol{\alpha} \in \mathbb{R}^p$. The minimizer $\boldsymbol{\alpha}^{(k+1)}$ is then scaled to satisfy $\boldsymbol{\alpha}^{(k+1)} \in \Theta$. The algorithm terminates, when $\|(\boldsymbol{\alpha}^{(k+1)} - \boldsymbol{\alpha}^{(k)})/\boldsymbol{\alpha}^{(k+1)}\|$ is less than a prespecified convergence tolerance.

Remark 1: The objective function of the modified covariates model (2) (**?**) is

$$n^{-1} \sum_{i=1}^{n} \left(Y_i - \boldsymbol{\alpha}^\top \boldsymbol{X}_i (T_i + \pi_1 - 2)\right)^2 /2, \tag{15}$$

in which the terms $\boldsymbol{X}_i(T_i + \pi_1 - 2)$ are called *modified covariates*. By comparing the right-hand side of (13) to (15), the modified regressors, $\boldsymbol{X}_i^{*(k)} = \dot{g}_{T_i}(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}_i) \boldsymbol{X}_i$, take the role of the modified covariates of **?** in updating $\boldsymbol{\alpha}$ for each (the kth) iterative step. For any set of arbitrary functions $(g_1, \ldots, g_L)$ satisfying the constraint (4), we have $E\left[\dot{g}_T(\boldsymbol{\alpha}^{(k)\top} \boldsymbol{X}) \boldsymbol{X} | \boldsymbol{X}\right] = 0$ (almost surely) for any $\boldsymbol{\alpha}^{(k)} \in \Theta$, which is analogous to the condition: $E\left[\boldsymbol{X}(T + \pi_1 - 2) | \boldsymbol{X}\right] = 0$ (almost surely) satisfied by the modified covariates model (2). These conditions, satisfied by the models, make the parts relevant to the optimization of $\boldsymbol{\alpha}$ in (13) and (15) independent

Accepted Article

This article is protected by copyright. All rights reserved.

of the unspecified $\mu(\boldsymbol{X})$ of the underlying models (1) and (2), respectively, as in (5). This *orthogonality* is attractive, since the estimation of the unspecified function $\mu(\boldsymbol{X})$ and the single-index coefficient $\boldsymbol{\alpha}_0$ can be performed separately, independently of each another. If we restrict the treatment t-specific functions g_t to $g_t(u) := u(t + \pi_1 - 2)$, the objective function of (11) reduces to (15).

REMARK 2: Since the weights $\dot{g}_{T_i}(\boldsymbol{\alpha}^{(k)\top}\boldsymbol{X}_i)$ that define the modified regressors $\boldsymbol{X}_i^{(k)*}$ in (14) depend on the values $\boldsymbol{X}_i$ and the shape of the link-functions g_{T_i} (through the first derivatives $\dot{g}_{T_i}$), the iterative procedure of optimizing $\boldsymbol{\alpha}$ that utilizes the right-hand side of (13) accounts for the nonlinear interactions captured by the flexible link-functions $(g_1, \ldots, g_L)$. This is in contrast to the constant weights $(T_i + \pi_1 - 2)$ applied to $\boldsymbol{X}_i$ in defining the modified covariates of **?**.

3.3 *Consistency and asymptotic normality of the estimator*

We establish the consistency of the estimator $\widehat{\boldsymbol{\alpha}}_0$ in (11) for $\boldsymbol{\alpha}_0$ and the estimators $\widehat{f_t}$ in (12) for f_t, where $\boldsymbol{\alpha}_0$ and f_t are given in model (1). The theoretical results rely on those of **?**, where cubic B-splines are used to approximate the link function of their single-index model. In **?**, instead of imposing the true mean function to be a function only of a single-index $\boldsymbol{\theta}_0^\top \boldsymbol{X} \in \mathbb{R}$, i.e., $E[Y|\boldsymbol{X}] = E[Y|\boldsymbol{\theta}_0^\top \boldsymbol{X}]$ for some single-index coefficient $\boldsymbol{\theta}_0 \in \Theta$, the authors develop a root-n consistent cubic spline estimator of the single-index coefficient of a single-index model which is robust against deviations from the exact single-index regression relationship. Specifically, their target single-index coefficient $\boldsymbol{\theta}_0 \in \Theta$ is defined in terms of the optimal L^2 (single-index based) approximation to the response Y: $\boldsymbol{\theta}_0 := \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}}\, E\big[\big(Y - E[Y|\boldsymbol{\theta}^\top \boldsymbol{X}]\big)^2\big]$, rather than in terms of an exact single-index relationship $E[Y|\boldsymbol{X}] = E[Y|\boldsymbol{\theta}_0^\top \boldsymbol{X}]$. In what follows, we adopt the results and assumptions of **?**, and obtain uniformly consistent estimators of the conditional expectations $E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T = t\right]$ $(t = 1, \ldots, L)$ and $E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}\right]$ appearing on the right-hand side of (7) (uniformly over $\boldsymbol{\alpha} \in \Theta$), and establish the root-n consistency of

Accepted Article

This article is protected by copyright. All rights reserved.

$\widehat{\boldsymbol{\alpha}}_0$ in (11) for $\boldsymbol{\alpha}_0$ of model (1); cubic spline smoothing of $Y(=Y^{(t)})$ on $(\widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{X}, T=t)$ results in uniformly consistent estimators $\hat{f}_t$ in (12) for f_t $(t=1,\ldots,L)$. We state our assumptions.

ASSUMPTION 1: The response $Y_i = \mu(\boldsymbol{X}_i)+f_{T_i}(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i)+\epsilon_i$ $(i=1,\ldots,n)$, where $E\left[\epsilon_i|\boldsymbol{X}_i, T_i\right] = 0$ and $E\left[\epsilon_i^2|\boldsymbol{X}_i, T_i\right] = \sigma_{T_i}^2(\boldsymbol{X}_i)$, in which the standard deviation functions $\sigma_t(\cdot)$ $(t=1,\ldots,L)$ are bounded below and above by positive constants, defined on a bounded domain.

ASSUMPTION 2: The function $E\left[(Y - f_T(\boldsymbol{\alpha}^\top \boldsymbol{X}))^2\right]$ is locally convex at $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$.

ASSUMPTION 3: The functions $\mu(\cdot)$ and $f_t(\cdot)$ in (1) have 4th order continuous derivatives.

ASSUMPTION 4: The covariate $\boldsymbol{X}$ is bounded, i.e., $\|\boldsymbol{X}\| \leqslant c$, for some $c > 0$. The density function of $\boldsymbol{X}$ has a 4th order continuous derivative, and is bounded above and below by positive constants, defined on a bounded domain.

ASSUMPTION 5: The number of interior knots d_t used in representing the function $g^{**}_{\boldsymbol{\alpha},t}(u) := E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X} = u, T=t\right]$ satisfies $n_t^{1/6} \ll d_t \ll n_t^{1/5}(\log(n_t))^{-(2/5)}$ $(t=1,\ldots,L)$. The number of interior knots d_0 used in representing the function $g^*_{\boldsymbol{\alpha}}(u) := E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X} = u\right]$ satisfies $n^{1/6} \ll d_0 \ll n^{1/5}(\log(n))^{-(2/5)}$.

Assumption 1 on the standard deviation functions $\sigma_t(\cdot)$ and Assumption 3 on the underlying regression functions are typical in the nonparametric smoothing literature; see for instance, **??**. Assumption 4 on the distribution of $\boldsymbol{X}$ is also assumed in **?**. Assumption 5 gives the requirement for the numbers of interior knots for the cubic spline spaces in approximating the conditional expectations specified on the right-hand side of (7), and is needed to ensure the uniform convergence of the approximated criterion function in (11) to its population counterpart in (8) over $\boldsymbol{\alpha} \in \Theta$. The strong consistency of $\widehat{\boldsymbol{\alpha}}_0$ in (11) to $\boldsymbol{\alpha}_0$ and $\widehat{f_t}$ in (12) to f_t are given below.

THEOREM 1: *Under Assumptions 1–5,* $\widehat{\boldsymbol{\alpha}}_0 \to \boldsymbol{\alpha}_0$*, almost surely.*

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

COROLLARY 1: *Under Assumptions 1–5,*

$$\sup_{u\in[0,1]} \left|\widehat{f}_t(u) - f_t(u)\right| \to 0, \quad \textit{almost surely} \quad (t = 1, \ldots, L). \tag{16}$$

In (16), without loss of generality, we take the domain of the functions $\widehat{f}_t$ and f_t to be $[0, 1]$, since $\boldsymbol{\alpha}^\top \boldsymbol{X}$ $(\boldsymbol{\alpha} \in \Theta)$ is bounded under Assumption 4. We next consider the asymptotic normality of $\widehat{\boldsymbol{\alpha}}_0$. In the working model (3), any vector $\boldsymbol{\alpha} \in \Theta \subset \mathbb{R}^p$ can be expressed as: $\boldsymbol{\alpha} = c(\boldsymbol{\phi}) := (1, \boldsymbol{\phi}^\top)^\top/(1+\|\boldsymbol{\phi}\|^2)^{1/2}$, for some vector $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{p-1})^\top \in \mathbb{R}^{p-1}$. Let $\boldsymbol{J}(\boldsymbol{\phi})$ denote the $p\times(p-1)$ Jacobian transformation matrix from $\boldsymbol{\phi} \in \mathbb{R}^{p-1}$ to $\boldsymbol{\alpha} \in \Theta$, whose (i, j)th element is given by $\partial\alpha_i/\partial\phi_j = -\alpha_i\phi_j/K^2$, for $(i = 1;\ j = 1, \ldots, p-1)$, and $\partial\alpha_i/\partial\phi_j = -\alpha_i\phi_j/K^2 + 1/K$, for $(i = 2, \ldots, p;\ j = 1, \ldots, p-1)$, where $K = (1+\|\boldsymbol{\phi}\|^2)^{1/2}$. Since the relation $\boldsymbol{\alpha} = c(\boldsymbol{\phi})$ is one-to-one, the parameter vector $\boldsymbol{\phi}_0 \in \mathbb{R}^{p-1}$ corresponding to the coefficient $\boldsymbol{\alpha}_0 \in \Theta$ of model (1) can be specified. Let us define the $(p-1)\times(p-1)$ covariance matrix $\boldsymbol{\Sigma}_0 = \text{var}[\Psi(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0)]$, in which $\Psi(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0) := \boldsymbol{J}^\top(\boldsymbol{\phi}_0)\{f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) - Y_i\}\dot{f}_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)\boldsymbol{X}_i$, where $\boldsymbol{\alpha}_0 = c(\boldsymbol{\phi}_0)$. Let $\boldsymbol{A}_0$ denote the $(p-1)\times(p-1)$ matrix of the first derivative of $E[\Psi(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi})]$ with respect to $\boldsymbol{\phi} \in \mathbb{R}^{p-1}$ evaluated at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$. The asymptotic normality of the estimator $\widehat{\boldsymbol{\alpha}}_0$ is given below.

THEOREM 2: *Under Assumptions 1–5,* $n^{1/2}(\widehat{\boldsymbol{\alpha}}_0 - \boldsymbol{\alpha}_0) \to \mathcal{N}(\boldsymbol{0}, \boldsymbol{J}_0\boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}_0(\boldsymbol{J}_0\boldsymbol{A}_0^{-1})^\top)$ *in distribution, where* $\boldsymbol{J}_0$ *is the Jacobian function* $\boldsymbol{J}(\boldsymbol{\phi})$ *evaluated at* $\boldsymbol{\phi} = \boldsymbol{\phi}_0$.

The proofs of Theorem 1 and 2 and Corollary 1 are in Section A.3 of the Supporting Information. Although the convergence rate of the nonparametric component $\widehat{f}_t$ in (12) to f_t $(t = 1, \ldots, L)$ is slower than root-n (see (A.11) of the Supporting Information Section A.3) under Assumption 5 on the numbers of interior knots d_t and d_0, the parametric component $\boldsymbol{\alpha}_0$ can be estimated at a root-n rate by letting the numbers of interior knots of the spline smoothing to increase with the sample size at an appropriate rate (Assumption 5). This indicates that the model can be estimated in two stages: estimation of $\boldsymbol{\alpha}_0$ by the root-n

This article is protected by copyright. All rights reserved.

consistent $\widehat{\boldsymbol{\alpha}}_0$; spline smoothing of Y on $\widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{X}$ for each $T = t$ $(t = 1, \ldots, L)$ to obtain an estimator $\widehat{f_t}$ (as given in (12)) of f_t. Under Assumptions 1–5, Theorem 2 states that root-n rate asymptotic normality for $\widehat{\boldsymbol{\alpha}}_0$ is achievable, and that the estimator is as efficient as if the true nonparametric functions f_t $(t = 1, \ldots, L)$ in model (1) were known and used as the link functions g_t $(t = 1, \ldots, L)$ of the working model (3). However, $\widehat{\boldsymbol{\alpha}}_0$ is generally not the most efficient estimator (although a root-n consistent estimator). This is because $\widehat{\boldsymbol{\alpha}}_0$ is based on a generally misspecified working model (3) that includes only the T-by-$\boldsymbol{X}$ interaction effect component and omits the main effect term. Analogous to the efficiency augmentation of **?**, the efficiency of the estimator can be improved by incorporating a main effect component of $\boldsymbol{X}$ to the estimation of $\boldsymbol{\alpha}_0$ (see Section A.4 of the Supporting Information).

3.4 *An illustration of the consistency of* $\widehat{\boldsymbol{\alpha}}_0$

In this subsection, we illustrate the effect of the constraint (4) on the consistency of $\widehat{\boldsymbol{\alpha}}_0$ for $\boldsymbol{\alpha}_0$ using a simulation experiment. For the purpose of illustration, we consider a simple case of $p = 2$ and $L = 2$. We generate $\{(Y_i, T_i, \boldsymbol{X}_i), i = 1, \ldots n\}$, where T_i takes a value in $\{1, 2\}$ with equal probability, independently of $\boldsymbol{X}_i = (X_{i,1}, X_{i,2})^\top \in \mathbb{R}^2$, where $X_{i,1}, X_{i,2}$ $\sim$ independent unif$[-\pi/2, \pi/2]$. Given T_i and $\boldsymbol{X}_i$, we generate Y_i from model (1), i.e., $Y_i = \mu(\boldsymbol{X}_i) + f_{T_i}(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i) + \epsilon_i$, with additive Gaussian noise ϵ_i, where $\sigma_{T_i}(\boldsymbol{X}_i) = 0.2$ (see Assumption 1). We set $n = 250$. We consider two simulation settings. In setting "A", the T-by-$\boldsymbol{X}$ interaction effect specified in model (1) is *nonlinear*, and it is defined by

$$\text{Setting A:}\ \ f_{T_i}(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i) = (-1)^{T_i}\left(\cos(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i) - 0.5\right) \quad (i = 1, \ldots, n).$$

In setting "B", the interaction effect is *linear*, as defined by

$$\text{Setting B:}\quad f_{T_i}(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i) = (-1)^{T_i}\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i/2.25 \quad (i = 1, \ldots, n).$$

In both settings "A" and "B", we take the main effect component in model (1) to be $\mu(\boldsymbol{X}; \delta) = \delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X})$, where the parameter $\delta \in \{1, 2, 4\}$ regulates the contribution of

Accepted Article

This article is protected by copyright. All rights reserved.

the $\boldsymbol{X}$ main effect to the variance of Y. In both settings, the contribution of $\mu(\boldsymbol{X};\delta)$ to the variance of Y is about 0.85, 3.5, and 14 times greater than that of the interaction effect $f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$, for $\delta = 1$, $\delta = 2$, and $\delta = 4$, respectively. In both settings, $\boldsymbol{\alpha}_0^\top \boldsymbol{X}$ corresponds to the index of interest (since it is associated with the T-by-$\boldsymbol{X}$ interaction effects) and $\boldsymbol{\eta}^\top \boldsymbol{X}$ corresponds to a "nuisance" index (since it is associated only with the $\boldsymbol{X}$ main effects). We set $\boldsymbol{\alpha}_0 := (1,1)^\top/\sqrt{2}$ and $\boldsymbol{\eta} := (1,-1)^\top/\sqrt{2}$. For the purpose of visualization, we parameterize vectors $\boldsymbol{\alpha} \in \Theta \subseteq \mathbb{R}^2$ in terms of an angle $\theta \in [-\pi/2, \pi/2)$ such that $\boldsymbol{\alpha} = (\cos(\theta), \sin(\theta))^\top$. We identify the vectors $\boldsymbol{\alpha}_0$ and $\boldsymbol{\eta}$ (in Cartesian coordinates) with the angles $\theta_1 = \pi/4$ and $\theta_2 = -\pi/4$ (in polar coordinates), respectively.

In this simulation example, as a function of $\theta \in [-\pi/2, \pi/2)$, we illustrate the squared error criterion (i.e., the objective function of (11)) (reparametrized with respect to θ) of the *constrained* working model (3). In addition, we illustrate the squared error criterion of the *unconstrained* working model, which is model (3) but without enforcing the constraint (4). For comparison, we also include the modified covariates squared error criterion (15) (reparametrized with respect to θ, through setting $\boldsymbol{\alpha} = \gamma(\cos(\theta), \sin(\theta))^\top$, in which $\gamma \in \mathbb{R}$ is "profiled out" for each value of θ under the squared error criterion).

We simulate 200 datasets and average the values of these three criterion functions for each value of $\theta \in [-\pi/2, \pi/2)$ (evaluated on a dense grid). Then each of the averaged criterion functions is scaled to have height 1. In Figure 2, the resulting averaged criterion functions are displayed for setting "A" on the top row, and the setting "B" on the bottom row.

[Figure 1 about here.]

In Figure 1, for all three cases of $\delta = 1$, $\delta = 2$, and $\delta = 4$, the *constrained* criterion (the red solid lines) has a "correct" global minimum at $\theta_1 = \pi/4$, implying that the minimization of this constrained criterion would lead to correctly identifying the T-by-$\boldsymbol{X}$ interaction effect coefficient $\boldsymbol{\alpha}_0$. The *unconstrained* criterion (the green dotted lines) has a correct minimum

Accepted Article

This article is protected by copyright. All rights reserved.

at $\theta_1 = \pi/4$ for the case $\delta = 1$ (i.e., when the main effect is relatively small), however, as the main effect intensity parameter δ increases (from $\delta = 1$ to $\delta = 2$ and to $\delta = 4$), the unconstrained criterion function takes its global minimum at the nuisance angle $\theta_2 = -\pi/4$, implying that the minimization of the unconstrained criterion would lead to an inconsistent estimate of $\boldsymbol{\alpha}_0$. Under the linear interaction effect (setting "B"), the constrained single-index regression criterion takes the consistent global minimum at $\theta_1 = \pi/4$ for all three cases of $\delta = 1$, $\delta = 2$, and $\delta = 4$. On the other hand, the unconstrained criterion has its global minimum at the nuisance angle $\theta_2 = -\pi/4$, for the cases $\delta = 2$ and $\delta = 4$ (i.e., when the main effect dominates the interaction effect).

This example has illustrated that the proposed constrained single-index regression criterion consistently takes its global minimum near the "signal" direction $\boldsymbol{\alpha}_0$ associated with the T-by-$\boldsymbol{X}$ interaction effect, even when the interaction effect signal is weak, unlike the case of the unconstrained criterion. In the linear interaction scenario (setting "B"), the modified covariates model (2) also produces a consistent estimate of the "signal" direction $\boldsymbol{\alpha}_0$ since, in this case, the modified covariates model is correctly specified and is a special case of the proposed constrained single-index model (1). However, when the interaction effect is nonlinear (scenario "A"), the modified covariates criterion does not provide relevant information for modeling the T-by-$\boldsymbol{X}$ interaction effects, as it takes its global minimum away from $\boldsymbol{\alpha}_0$.

4. Simulation Studies

In this section, we perform numerical studies to illustrate the performance of the proposed approach to optimizing individualized treatment rules, in comparison with alternative approaches including the modified covariates approach of **?** and the outcome weighted learning method of **?**.

We consider $p = 10$ with $n \in \{250, 500\}$. We generate covariates $\boldsymbol{X}_i \in \mathbb{R}^p$ consisting of independent unif$[-\pi/2, \pi/2]$ variates (a correlated covariate case is considered in Supporting

Accepted Article

This article is protected by copyright. All rights reserved.

Information Section B.5), and the treatment variable T_i takes a value in $\{1, 2\}$ with equal probability, independently of $\boldsymbol{X}_i$. We generate the treatment outcome $Y_i (= \mu(\boldsymbol{X}_i) + f_{T_i}(\boldsymbol{X}_i) + \epsilon_i)$ from 1) the simulation setting "A":

$$Y_i = 0.8\delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X}_i) + (-1)^{T_i} \left\{ e^{-(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i - 0.5)^2 - (\boldsymbol{\alpha}_1^\top \boldsymbol{X}_i - 0.5)^2 \xi} - 0.5 \right\} + \epsilon_i \quad (i = 1, \ldots, n), \tag{17}$$

where the first term $\mu(\boldsymbol{X}; \delta) := 0.8\delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X})$ $(\delta = 1, 2)$ corresponds to the main effect of $\boldsymbol{X}$, and the second term $f_T(\boldsymbol{X}; \xi) := (-1)^T \{ e^{-(\boldsymbol{\alpha}_0^\top \boldsymbol{X} - 0.5)^2 - (\boldsymbol{\alpha}_1^\top \boldsymbol{X})^2 \xi} - 0.5 \}$ $(\xi = 0, 0.5)$ corresponds to the T-by-$\boldsymbol{X}$ interaction effect, determined by a bell-shaped (i.e., Gaussian) surface over two 1-dimensional indices, $\boldsymbol{\alpha}_0^\top \boldsymbol{X}$ and $\boldsymbol{\alpha}_1^\top \boldsymbol{X}$ (if $\xi \neq 0$), which nonlinearly modifies the effect of the variable T on the outcome Y. We also consider 2) the simulation setting "B":

$$Y_i = \delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X}_i) + (-1)^{T_i} \left\{ \cos\left((\boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 I_{(T_i = 2)} \xi)^\top \boldsymbol{X}_i - \pi/8 \right) - 0.5 \right\} + \epsilon_i \quad (i = 1, \ldots, n), \tag{18}$$

where the first term $\mu(\boldsymbol{X}; \delta) := \delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X})$ $(\delta = 1, 2)$ determines the $\boldsymbol{X}$ main effect, and the second term $f_T(\boldsymbol{X}; \xi) := (-1)^T \left\{ \cos\left((\boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 I_{(T=2)} \xi)^\top \boldsymbol{X}_i - \pi/8 \right) - 0.5 \right\}$ $(\xi = 0, 0.5)$ determines the T-by-$\boldsymbol{X}$ interaction effect, which is defined based on different single indices for each treatment, i.e., $\boldsymbol{\alpha}_0^\top \boldsymbol{X}$ for $T = 1$ and $(\boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 \xi)^\top \boldsymbol{X}$ for $T = 2$, when $\xi \neq 0$.

Both models (17) and (18) are indexed by a pair (ξ, δ): the parameter $\xi \in \{0, 0.5\}$ determines whether the T-by-$\boldsymbol{X}$ interaction effect term $f_T(\boldsymbol{X}; \xi)$ in models (17) and (18) has an intrinsic 1-dimensional structure over the single index $\boldsymbol{\alpha}_0^\top \boldsymbol{X}$ $(\xi = 0)$ or whether it deviates from a single-index model structure $(\xi = 0.5)$; the parameter $\delta \in \{1, 2\}$ in $\mu(\boldsymbol{X}; \delta)$ controls the contribution of the $\boldsymbol{X}$ main effect on the variance of Y, where $\delta = 1$ represents a relatively *moderate* $\boldsymbol{X}$ main effect (contributing about the same variance as the interaction effect does) and $\delta = 2$ a relatively *large* $\boldsymbol{X}$ main effect (about 4 times greater than the interaction effect), respectively. In both (17) and (18), we use additive noise ϵ_i (see Assumption 1) that follows the mean zero Gaussian distribution with standard deviation 0.5. In models (17) and (18), we set the vectors $\boldsymbol{\eta} = (-1, 1, -1, 1, -1, 1, 0, 0, 0, 0)^\top$, $\boldsymbol{\alpha}_0 = (1, 0.5, 0.25, 0.125, 0, 0, 0, 0, 0, 0)^\top$

Accepted Article

This article is protected by copyright. All rights reserved.

and $\boldsymbol{\alpha}_1 = (1,1,1,1,1,1,0,0,0,0)^\top$, with each of these length $p(=10)$ vectors normalized to have unit norm. Without loss of generality, we assume that a larger value of Y is preferred.

For the case of a single decision time point, an individualized treatment rule, which we denote as $\mathcal{D}(\boldsymbol{X}) : \mathbb{R}^p \mapsto \{1,\ldots,L\}$, is a rule that maps an individual with (baseline) characteristics $\boldsymbol{X}$ to one of the L available treatment options. One natural measure for the effectiveness of $\mathcal{D}$ is called the value (V) of $\mathcal{D}$ (**?**) defined as the expected mean response when everyone in the population receives treatment according to the rule $\mathcal{D}$, that is: $V(\mathcal{D}) = E[E[Y|\boldsymbol{X}, T = \mathcal{D}(\boldsymbol{X})]]$. The optimal individualized treatment rule, which we denote as $\mathcal{D}^{opt}$, resulting in the largest value of V is: $\mathcal{D}^{opt}(\boldsymbol{X}) = \arg\max_{t\in\{1,\ldots,L\}} E[Y|\boldsymbol{X}, T = t]$. Accordingly, the estimators $\widehat{\mathcal{D}}^{opt}$ for $\mathcal{D}^{opt}$ are: $\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}) = \arg\max_{t\in\{1,\ldots,L\}} \widehat{f}_t(\widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{X})$ and $\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}) = \arg\max_{t\in\{1,\ldots,L\}} \widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{X}(t + \pi_1 - 2)$, for models (1) and (2), respectively. Here, the estimators $\widehat{\boldsymbol{\alpha}}_0$ and $\widehat{f}_t$ for model (1) correspond to the proposed estimators (11) and (12), and the estimator $\widehat{\boldsymbol{\alpha}}_0$ for model (2) corresponds to the minimizer of (15).

The methods for estimating $\mathcal{D}^{opt}$ under comparison include

1. The proposed method of using model (3), estimated through the procedure described in Section 3.2. We use the numbers of interior knots $d_t = [n_t^{1/5.5}]$ $(t = 1, 2)$ and $d_0 = [n^{1/5.5}]$ which satisfy Assumption 5. Here, $[u]$ denotes the integer part of u.
2. The modified covariates method of **?** of using model (2), estimated by minimizing (15).
3. The outcome weighted learning (OWL) method (**?**) based on a linear kernel, implemented in the R-package `DTRlearn`. To improve its efficiency, we employ the augmented outcome weighted learning approach of **?**. The tuning parameter κ in **?** is chosen from the grid of $(0.25, 0.5, 1, 2, 4)$ (the default setting of `DTRlearn`) based on a 5 fold cross-validation.
4. The same approach as in 3 but based on a Gaussian radial basis function kernel instead of a linear kernel. The inverse bandwidth parameter σ_n^2 in **?** is chosen from the grid of

Accepted Article

This article is protected by copyright. All rights reserved.

$(0.01, 0.02, 0.04, \ldots, 0.64, 1.28)$ and κ is chosen from the grid of $(0.25, 0.5, 1, 2, 4)$, based on a 5 fold cross-validation.

5. A penalized additive cubic spline least squares (PLS) approach. We implement this method by estimating $E[Y|\boldsymbol{X}, T = t]$ via an additive regression for each treatment separately. The implementational detail is given in Section B.4 of Supporting Information.

For each simulation run, we estimate $\mathcal{D}^{opt}$ from each of the 5 methods based on a training set (of size n), and for evaluation of these methods, we evaluate the value $V(\widehat{\mathcal{D}}^{opt}) = E[E[Y|\boldsymbol{X}, T = \widehat{\mathcal{D}}^{opt}(\boldsymbol{X})]]$ for each estimate $\widehat{\mathcal{D}}^{opt}$, using a Monte Carlo approximation based on a random sample of size 10^3, denoted as $\widehat{V}(\widehat{\mathcal{D}}^{opt})$. Since we know the true data generating model in simulation studies, the optimal $\mathcal{D}^{opt}$ can be determined for each simulation run. Given each estimate $\widehat{\mathcal{D}}^{opt}$ of $\mathcal{D}^{opt}$, we report $\widehat{V}(\widehat{\mathcal{D}}^{opt}) - \widehat{V}(\mathcal{D}^{opt})$, as the performance measure of $\widehat{\mathcal{D}}^{opt}$. A larger value of the measure indicates better performance.

[Figure 2 about here.]

In Figure 2, we present the boxplots, obtained from 100 simulation runs, of the normalized values $\widehat{V}(\widehat{\mathcal{D}}^{opt})$ (normalized by the optimal values $\widehat{V}(\mathcal{D}^{opt})$) of the decision rules $\widehat{\mathcal{D}}^{opt}$ estimated from the 5 approaches, for each combination of $n \in \{250, 500\}$, $\xi \in \{0, 0.5\}$ (corresponding to *correctly-specified* or *mis-specified* single-index interaction effect models, respectively) and $\delta \in \{1, 2\}$ (corresponding to *moderate* or *large* main effects, respectively), for the simulation setting "A" in the top panels and the setting "B" in the bottom panels.

The results in Figure 2 indicate that the proposed constrained single-index model (CSIM) outperforms all other approaches in estimating $\mathcal{D}^{opt}$. With substantial nonlinearity in the interaction effect term of the models (17) and (18), the modified covariates method, which assumes a restricted linear model on the interaction term, is clearly outperformed by the proposed model that utilizes a set of flexible link functions to accommodate the nonlinear treatment effect modification. The estimated values of the outcome weighted learning with

Accepted Article

This article is protected by copyright. All rights reserved.

a linear kernel and a Gaussian kernel, respectively, are similar to each other, but both are outperformed by the the constrained single-index regression, even when the true interaction model deviates from a single-index model (when $\xi = 0.5$). When $n = 500$ (i.e., with a relatively large sample size) and $\xi = 0.5$ (i.e., when the underlying model deviates from the exact single-index structure), the penalized additive spline approach (PLS), due to its large model space, outperforms the modified covariates method slightly; however, the approach is clearly outperformed by the proposed single-index method that is robust to the main effect model misspecification and also allows nonlinear interactions. When the $\boldsymbol{X}$ main effect dominates the T-by-$\boldsymbol{X}$ interaction effect (i.e., when $\delta = 2$), although the increased magnitude of the main effect dampens the performance of all approaches to optimizing treatment decisions, the constrained single-index regression consistently targets to model the interaction effect-related variability, and its performance is near optimal when $n = 500$.

5. Application to data from a depression randomized clinical trial

The development of the constrained single-index model method was motivated by a randomized clinical trial comparing an antidepressant ($T = 2$) and placebo ($T = 1$) for treating major depressive disorder. The primary purpose of the study is the development of a biosignature, called a differential treatment response index, defined as a combination of multiple biomarkers which can be used for optimization of an individualized treatment rule for patients with major depressive disorder (**?**). In major depressive disorder, each patient characteristic often has at most a weak modifying effect. Therefore, the proposed single-index modeling approach that creates a differential treatment response single-index that collectively exhibits a stronger, and possibly nonlinear, interaction with the treatment is a very clinically significant endeavor.

Of the 166 subjects, 88 were randomized to placebo and 78 to drug. Pretreatment clinical characteristics $\boldsymbol{X} = (X_1, \ldots, X_5)^\top$ include: X_1 = age at evaluation; X_2 = severity of

Accepted Article

This article is protected by copyright. All rights reserved.

depressive symptoms measured by the Hamilton rating scale at baseline; X_3 = logarithm of duration (in month) of the current major depressive episode. In addition, patients underwent neuropsychiatric testing at baseline to assess reaction time, X_4 = median choice reaction time and cognitive control, X_5 = Flanker accuracy, as these behavioral characteristics are believed to correspond to biological phenotypes related to response to antidepressants (**?**). For the purposes of regularization, all pretreatment covariates are centered and scaled to have mean 0 and unit variance. The outcome Y is the improvement in symptoms severity (assessed by the Hamilton rating scale for depression) from baseline to week 8 taken as the difference (week 0 - week 8), and thus larger values of the outcome are considered desirable.

The estimated single-index coefficients $\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_1, \ldots, \widehat{\alpha}_5)^\top$ of the proposed model (1) and their 95% normal approximation bootstrap confidence intervals based on 500 bootstrap replications (see Supporting Information Section B.6 for the coverage proportions of the bootstrap confidence intervals assessed by simulations) are given by: $\widehat{\alpha}_1 = 0.69(0.31, 1.06)$, $\widehat{\alpha}_2 = 0.23(-0.10, 0.57)$, $\widehat{\alpha}_3 = 0.33(0.03, 0.64)$, $\widehat{\alpha}_4 = -0.22(-0.51, 0.08)$ and $\widehat{\alpha}_5 = -0.55(-0.85, -0.25)$, respectively. The estimated treatment t-specific functions $\widehat{f}_t(\cdot)$ $(t = 1, 2)$ (with 95% confidence bands, given $\widehat{\boldsymbol{\alpha}}$) are illustrated in the first two panels of Figure 3.

[Figure 3 about here.]

The right panel of Figure 3 displays the contrast between the two estimated treatment effects (drug - placebo) versus the estimated single-index. This indicates that the superiority of the drug over placebo does not linearly increase with $z = \boldsymbol{\alpha}^\top \boldsymbol{X}$, but rather, it appears to plateau out with some nonlinear patterns to the right of the zero crossing point near $z = -0.7$, and it has another zero crossing point near $z = 2.4$. As implied by the contrast plot in Figure 3, an individualized treatment rule based on the single-index $z = \boldsymbol{\alpha}^\top \boldsymbol{X}$ can be constructed by assigning patients with the index $-0.7 < z < 2.4$ to the active drug.

To evaluate the performance of the individualized treatment rules ($\widehat{\mathcal{D}}^{opt}$) estimated from

Accepted Article

This article is protected by copyright. All rights reserved.

5 different approaches described in Section 4, we randomly split the dataset at a ratio of 5 to 1 into a training set and a testing set (of size $\tilde{n}$), replicated 500 times, each time obtaining $\widehat{\mathcal{D}}^{opt}$ based on the training set and estimating the value of $\widehat{\mathcal{D}}^{opt}$, $V(\widehat{\mathcal{D}}^{opt}) = E[E[Y|\boldsymbol{X}, T = \widehat{\mathcal{D}}^{opt}(\boldsymbol{X})]]$, by an inverse probability weighted estimator (**?**) $\widehat{V}(\widehat{\mathcal{D}}^{opt}) = \sum_{i=1}^{\widetilde{n}} Y_i 1_{(T_i=\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}_i))} / \sum_{i=1}^{\widetilde{n}} 1_{(T_i=\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}_i))}$ based on the testing set (of size $\widetilde{n}$). For the modified covariates method, we use a linear model with covariates $\boldsymbol{X}$ for efficiency augmentation. For comparison, we also include 2 naïve rules: treating all patients with placebo; and treating all patients with the active drug, each regardless of the individual patient's characteristics $\boldsymbol{X}$.

[Figure 4 about here.]

As Figure 4 shows, the proposed constrained single-index regression for estimating $\mathcal{D}^{opt}$ outperforms all other alternatives in terms of the average estimated values. In particular, the approach outperforms the modified covariates method and the outcome weighted learning with a linear kernel, illustrating the utility of the flexible treatment-specific link functions in approximating the nonlinear interactions. The method also outperforms the penalized additive spline least squares approach, suggesting that estimating and utilizing an optimal linear combination (a single-index $\boldsymbol{\alpha}^\top \boldsymbol{X}$) of biomarkers that collectively exhibits a stronger (and possibly nonlinear) interaction with the treatment is practically an appealing approach to optimizing treatment decision rules. In this example, the outcome weighted learning with a Gaussian kernel does not perform well.

The proposed single-index regression provides a visualization of the estimated single-index as shown in the panels of Figure 3, and the relative importance of each pretreatment covariate in terms of characterizing the heterogeneous treatment responses can be indicated by the coefficients $(\alpha_1, \ldots, \alpha_5)^\top$. The practical utility of the proposed methodology is highlighted here by noting that the difference between the values of the treatment decision rule based

Accepted Article

This article is protected by copyright. All rights reserved.

on the new method and the values of the naïve rule that treats everyone with the drug is almost twice as large as the difference between the efficacies of the drug and placebo.

6. Discussion

The proposed method is primarily designed to analyze data from randomized clinical trials. A limitation of the proposed method can occur when applying it to an observational study, where the covariates and treatment assignment can be correlated, in which case the estimator $\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}) = \arg\max_{t\in\{1,\ldots,L\}} \widehat{f}_t(\widehat{\boldsymbol{\alpha}}_0^\top \boldsymbol{X})$ might not yield the optimal decision rule. However, the working model (3), with the link functions $(g_1, \ldots, g_L)$ as defined in (7), can still be useful in fitting the T-by-$\boldsymbol{X}$ interaction effect term of model (1). If there are estimators $(g_1, \ldots, g_L)$, for each fixed $\boldsymbol{\alpha}$, that asymptotically satisfy (7), then, in the objective function (6), the part relevant to the estimation of the coefficient $\boldsymbol{\alpha}_0$ of model (1) is asymptotically separated from the $\boldsymbol{X}$ main effect term $\mu(\boldsymbol{X})$ of model (1), as in (5), resulting in robustness against misspecification of the $\boldsymbol{X}$ main effects in estimating the T-by-$\boldsymbol{X}$ interactions. We can utilize an iterative optimization procedure to optimize both $\boldsymbol{\alpha}$ and $(g_1, \ldots, g_L)$. For each fixed $\boldsymbol{\alpha}$, estimators of $(g_1, \ldots, g_L)$ that asymptotically satisfy (7) are relatively easy to obtain. For example, we can first compute unconstrained estimators, denoted as $\widehat{g}_t^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$, of the conditional expectations $g_t^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X}) = E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T = t]$ $(t = 1, \ldots, L)$, based on a 1-dimensional (along the axis $\boldsymbol{\alpha}^\top \boldsymbol{X}$) nonparametric smoother for each treatment $T = t$ $(t = 1, \ldots, L)$, and then remove the component in the fitted $\widehat{g}_T^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X})$ corresponding to the main effect of $\boldsymbol{\alpha}^\top \boldsymbol{X}$, by fitting a 1-dimensional nonparametric smoother (along the axis $\boldsymbol{\alpha}^\top \boldsymbol{X}$), denoted as $\widehat{g}^*(\boldsymbol{\alpha}^\top \boldsymbol{X})$, to the fitted $\widehat{g}_T^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X})$. Then, for each fixed $\boldsymbol{\alpha}$, we can take $\widehat{g}_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = \widehat{g}_t^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X}) - \widehat{g}^*(\boldsymbol{\alpha}^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$ as such estimators $(g_1, \ldots, g_L)$, which approximately satisfy (7). See Supporting Information Section A.6 for a justification for the robustness of this procedure against misspecification of $\mu(\cdot)$. However, when T depends on

Accepted Article

This article is protected by copyright. All rights reserved.

$\boldsymbol{X}$, the fitted T-by-$\boldsymbol{X}$ interaction effect term might result in biased causal effect estimates, as described in Supporting Information Section A.7.

In many applications, only a subset of measurements may be useful in determining an optimal treatment decision rule. Also, high-dimensional settings can lead to instabilities and issues of overfitting. Forthcoming work will introduce a regularization method that can both avoid overfitting and choose among multiple potential covariates by obtaining a sparse estimate of the single-index coefficient $\boldsymbol{\alpha}_0$. In this paper, the theoretical results are developed with a B-spline basis approximation, with the number of knots used as the tuning parameters. In finite samples, the choice of the number of knots can be crucial and delicate. At present an ad-hoc choice of $d_t = \lfloor n_t^{1/5.5} \rfloor$ and $d_0 = \lfloor n^{1/5.5} \rfloor$ is used for the number of knots, which is likely to be sub-optimal in practice (note, one can set $d_t = C\lfloor n_t^{1/5.5} \rfloor$ or $d_0 = C\lfloor n^{1/5.5} \rfloor$ with an arbitrary $C > 0$ while still achieving the requirements in Assumption 5). In practice, a penalized spline approximation, e.g., P-splines (**?**), can be considered, which is relatively robust to the choice of the number of knots. Although in Section 3.3, Assumption 4 does not allow discrete-valued covariates, our simulation experiment (see Supporting Information Section B.2) where the covariates $\boldsymbol{X}$ include a binary variable suggests that estimation in practice is rather insensitive to departure from this assumption. Generally, multiple local optima exist in the squared error criterion in (11) (and also its population counterpart in (6) with respect to $\boldsymbol{\alpha}$, where g_t is defined in (7)), and in such cases the estimates can be sensitive to the choice of initialization. We assume (in Assumption 1) the local convexity of the criterion function near $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, and this implies that the algorithm will converge if the initial estimate is close to $\boldsymbol{\alpha}_0$. Otherwise, the estimate $\widehat{\boldsymbol{\alpha}}_0$ can be sub-optimal. One way to mitigate this problem is to consider multiple initial points for $\boldsymbol{\alpha}_0$ in estimation (as in Supporting Information Section B.2 and B.3). In this paper, we use a linear estimate based on the modified covariates linear model (2) (scaled to be in Θ) as an initial estimate, and the

Accepted Article

This article is protected by copyright. All rights reserved.

estimate $\widehat{\boldsymbol{\alpha}}_0$, optimized to incorporate possibly nonlinear interactions, provides a significant improvement over the modified covariate initial estimate, as illustrated by the simulations in Section 4.

Future directions of this work also include an extension of the proposed regression to a multiple-index regression for modeling interactions. For example, when $L = 2$, model (1) can be extended to a partially-linear single-index model (e.g., **??**) by adding a modified covariate (**?**) linear component to the single-index interaction component. We will also investigate the incorporation of functional covariates and longitudinal outcomes.

Acknowledgements

We are grateful to the editor, the associate editor, and two referees for their insightful comments and suggestions. One referee specifically suggested the direction that clarifies a limitation of the proposed method that can occur when applying it to an observational study setting. This work was supported by National Institute of Health (NIH) grant 5 R01 MH099003.

Data Availability Statement

The data that support the findings in this paper are available from the corresponding author upon reasonable request.

References

de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.

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

Hardle, W., Hall, P., and Ichimura, H. (1993). Optimal smoothing in single-index models. *Annals of Statistics* **21,** 157–178.

Accepted Article

This article is protected by copyright. All rights reserved.

Hastie, T. and Tibshirani, R. (1999). *Generalized Additive Models*. Chapman & Hall Ltd.

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

Kang, C., Janes, H., and Huang, Y. (2014). Combining biomarkers to optimize patient treatment recommendations. *Biometrics* **70,** 696–707.

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

Lian, H. and Liang, H. (2016). Separation of linear and index covariates in partially linear single-index models. *Journal of Multivariate Analysis* **143,** 56–70.

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

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

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

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

Park, H., Petkova, E., Tarpey, T., and Ogden, R. T. (2019). simml: A single-index model with multiple-links. *R package version 0.1.0* .

Park, H., Petkova, E., Tarpey, T., and Ogden, R. T. (2020). A single-index model with multiple-links. *Journal of Statistical Planning and Inference* **205,** 115–128.

Petkova, E., Park, H., Ciarleglio, A., Ogden, R., and Tarpey, T. (2019). Optimising treatment decision rules through generated effect modifiers: a precision medicine tutorial. *BJPsych Open* **6,** 1–7.

Accepted Article

This article is protected by copyright. All rights reserved.

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

R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

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

Shi, C., Fan, A., Song, R., and Lu, W. (2018). High-dimensional A-learning for optimal dynamic treatment regimes. *The Annals of Statistics* **46,** 925–957.

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

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

Song, R., Luo, S., Zeng, D., Zhang, H., Lu, W., and Li, Z. (2017). Semiparametric single-index model for estimating optimal individualized treatment strategy. *Electronic Journal of Statistics* **11,** 364–384.

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

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

Trivedi, M., McGrath, P., Fava, M., Parsey, R., Kurian, B., Phillips, M., Pquendo, M., Bruder, G., Pizzagalli, D., Toups, M., Cooper, C., Adams, P., Weyandt, S., Morris, D., Grannemann, B., Ogden, R., Buckner, R., McInnis, M., Kraemer, H., Petkova, E., Carmody, T., and Weissman, M. (2016). Establishing moderators and biosignatures of

Accepted Article

This article is protected by copyright. All rights reserved.

antidepressant response in clinical care: Rationale and design. *Journal of Psychiatric Research* **78,** 11–23.

Wang, L. and Yang, L. (2009). Spline estimation of single-index models. *Statistica Sinica* **19,** 765–783.

Xia, Y., Tong, H., and Li, W. (1999). On extended partially linear single-index models. *Biometrika* **86,** 831–842.

Xia, Y., Tong, H., Li, W., and Zhu, L. (2002). An adaptive estimation of dimension reduction space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology* **64,** 363–410.

Zhang, B., Tsiatis, A. A., Davidian, M., Zhang, M., and Laber, E. (2012). Estimating optimal treatment regimes from classification perspective. *Stat* **1,** 103–114.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68,** 1010–1018.

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

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

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

Supporting Information

Web Appendices A (Technical details of mathematical results) and B (Results from simulation studies) referenced in Sections 3, 4, 5 and 6 are available with this paper at the Biometrics

Accepted Article

This article is protected by copyright. All rights reserved.

website on Wiley Online Library. An R code demonstrating the method described in this article is also available there. The R package `simml` (**?**) available on CRAN (**?**) provides an implementation of the proposed method.

Received September 2019. *Revised September* 2019.

Accepted September 2019.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Figure 1. The averaged criterion functions of $\theta \in [-\pi/2, \pi/2)$ averaged over 200 simulated datasets. The vector $\boldsymbol{\alpha}_0$ corresponds to the angle $\theta_1 = \pi/4$ which is indicated by the grey dashed vertical line, and the "nuisance" vector $\boldsymbol{\eta}$ corresponds to the angle $\theta_2 = -\pi/4$ indicated by the grey dotted vertical line. The criteria and their corresponding line styles: 1) the constrained single-index criterion (the red solid curves); 2) the unconstrained single-index criterion (the green dotted curves); 3) the modified covariates criterion (the blue dashed curves).

This article is protected by copyright. All rights reserved.

Figure 2. Boxplots comparing 5 approaches to estimating $\mathcal{D}^{opt}$, given each scenario indexed by $\xi \in \{0, 0.5\}$ and $\delta \in \{1, 2\}$, for the simulation setting "A" in the top panels and the setting "B" in the bottom panels. For each scenario, from left to right, estimation approaches for $\mathcal{D}^{opt}$: 1) the constrained single-index model (red); 2) the modified covariates model (green); 3) the outcome weighted learning with a linear kernel (violet); 4) the outcome weighted learning with a Gaussian kernel (purple); 5) the penalized spline least squares approach (dark purple). The case with $\xi = 0$ (or $\xi = 0.5$) corresponds to the correctly-specified (or mis-specified) single-index interaction model scenario; $\delta = 1$ (or $\delta = 2$) corresponds to the moderate (or large) main effect scenario. The dotted horizontal line represents the optimal value corresponding to $\mathcal{D}^{opt}$.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Figure 3. Depression randomized clinical trial: scatter plots of the outcome against the estimated single index $z = \boldsymbol{\alpha}^\top \boldsymbol{X}$, for the placebo group ($T = 1$, in the left panel) and the drug group ($T = 2$, in the middle panel); the estimated treatment-specific curve (with 95% confidence bands) for each group is overlaid (the red solid curve). In the right panel, the contrast between the estimated two treatment effects (drug - placebo) as a function of the estimated single-index is displayed.

This article is protected by copyright. All rights reserved.

Accepted Article

Figure 4. Depression randomized clinical trial: boxplots of the estimated values of the individualized treatment rules estimated from 7 approaches, obtained from 500 randomly split testing sets. From left to right, estimation approaches to $\mathcal{D}^{opt}$: 1) the constrained single-index model; 2) the modified covariates model; 3) outcome weighted learning with a linear kernel; 4) outcome weighted learning with a Gaussian kernel; 5) the penalized spline least squares approach; 6) treating all patients with placebo; 7) treating all patients with the active drug.

This article is protected by copyright. All rights reserved.