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

Hyung G. Park[a*], Eva Petkova[a], Thaddeus Tarpey[a], R. Todd Ogden[b]

[a] *Division of Biostatistics, Department of Population Health, New York University, New York, NY 10016, USA*
[b] *Department of Biostatistics, Columbia University, New York, NY 10032, USA*

Abstract

We provide proofs for the theoretical results presented in the main manuscript of the paper in Web Appendix A, and additional numerical examples including an illustration of the proposed regression approach to a $L = 3$ treatment level scenario and comparisons to the doubly robust estimation approach to optimizing individualized treatment rules in Web Appendix B.

Web Appendix A: Technical details of mathematical results

A.1. Proof of Proposition 1

In the main manuscript, we assume $Y = E[Y|\boldsymbol{X}, T] + \epsilon$, in which ϵ is a zero-mean independent noise with finite variance and

$$E[Y|\boldsymbol{X}, T = t] = \mu(\boldsymbol{X}) + f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \quad (t = 1, \dots, L), \tag{S.1}$$

where we assume, without loss of generality, $E[f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})|\boldsymbol{X}] = \sum_{t=1}^{L} \pi_t f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) = 0$, for an identifiable representation of (S.1) and $\boldsymbol{\alpha}_0 \in \Theta = \{\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^\top \in \mathbb{R}^p : \|\boldsymbol{\alpha}\| = 1, \alpha_1 > 0\}$. To estimate the components $f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$ $(t = 1, \dots, L)$ of model (S.1), we utilize the following working model

$$E[Y|\boldsymbol{X}, T = t] \;\approx\; g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) \quad (t = 1, \dots, L), \tag{S.2}$$

for some $\boldsymbol{\alpha} \in \Theta$, subject to the constraint on the treatment-specific smooth link-functions $(g_1, \dots, g_L)$:

$$E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{X}] = \sum_{t=1}^{L} \pi_t g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = 0 \quad \text{(almost surely)}, \tag{S.3}$$

for all $\boldsymbol{\alpha} \in \Theta$.

In a least squares framework of optimizing model (S.2), $\underset{\boldsymbol{\alpha},(g_1,\dots,g_L)}{\operatorname{argmin}} E\big[(Y - g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}))^2\big]/2 =$

[0]To whom correspondence should be addressed; parkh15@nyu.edu

$\underset{\boldsymbol{\alpha},(g_1,\ldots,g_L)}{\text{argmax}}\ E\big[Y g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2\big]$, in which

$$\begin{aligned}
E\left[Y g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2)\right] &= E\left[(\mu(\boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X}))g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2\right] \\
&= E\big[E\big[\mu(\boldsymbol{X})g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2 \mid \boldsymbol{X}\big]\big] \\
&= E\big[E\big[f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2 \mid \boldsymbol{X}\big]\big] \\
&= E\Big[\sum_{t=1}^{L} \pi_t \big\{ f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_t^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2 \big\}\Big] \\
&= \sum_{t=1}^{L} \pi_t E\big[E\big[f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X})g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_t^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2 \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big]\big] \\
&= \sum_{t=1}^{L} \pi_t E\big[E\big[f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_t^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2\big],
\end{aligned}$$

where the third line follows from the constraint (S.3) imposed on $(g_1, \ldots, g_L)$. For the notational simplicity, let us write $Q(g_1, \ldots, g_L, \boldsymbol{\alpha}) = E\left[Y g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) - g_T^2(\boldsymbol{\alpha}^\top \boldsymbol{X})/2)\right]$. Conditioning on $\boldsymbol{X}$ and for a fixed $\boldsymbol{\alpha}$, we have

$$\frac{\partial Q(g_1, \ldots, g_L, \boldsymbol{\alpha})}{\partial g_t} = E\big[f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] - g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) \quad (t = 1, \ldots, L). \tag{S.4}$$

Therefore, the stationary point of $Q(g_1, \ldots, g_L, \boldsymbol{\alpha})$, for each fixed $\boldsymbol{\alpha}$, can be formulated as:

$$g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = E\big[f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \quad (t = 1, \ldots, L) \quad \text{(almost surely).} \tag{S.5}$$

Since $f_t(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) = E[Y|\boldsymbol{X}, T = t] - \mu(\boldsymbol{X})$ by the true model (S.1), the right-hand side of (S.5) can be expressed as, for each $t = 1, \ldots, L$,

$$\begin{aligned}
g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) &= E\big[E[Y|\boldsymbol{X}, T = t] - \mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\
&= E\big[E[Y|\boldsymbol{X}, T = t] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] - E\big[\mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\
&= E\big[E[Y|\boldsymbol{X}, T = t] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}, T = t\big] - E\big[\mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\
&= E\big[Y \mid \boldsymbol{\alpha}^\top \boldsymbol{X}, T = t\big] - E\big[\mu(\boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\
&= E\big[Y \mid \boldsymbol{\alpha}^\top \boldsymbol{X}, T = t\big] - E\big[Y \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big],
\end{aligned} \tag{S.6}$$

where the fourth line follows from: $E\big[f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] = E\big[E[f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X})|\boldsymbol{X}] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] = 0$, which comes from the identifiability condition of model (S.1). Expression (S.6) gives the desired result of Proposition 1.

A.2. Derivation of the approximate solutions for the t-specific functions g_t

Given a $n \times d$ matrix $\boldsymbol{D}$, let us use $\mathcal{S}(D)$ to denote the linear subspace (in $\mathbb{R}^n$) spanned by the columns of $\boldsymbol{D}$. Let us define the vector spaces: $V_t := \mathcal{S}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)})$ $(t = 1, \ldots, L)$, $V_0 := \mathcal{S}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)})$ and $V := \mathcal{S}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(1)}, \ldots, \boldsymbol{D}_{\boldsymbol{\alpha}}^{(L)})$, where the matrices $\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)}$ $(t = 1, \ldots, L)$ and $\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)}$ are defined in the main manuscript. Note that $V = V_1 \oplus \cdots \oplus V_L$. Therefore, $\mathcal{P}_V(\boldsymbol{Y}) = \sum_{t=1}^{L} \mathcal{P}_{V_t}(\boldsymbol{Y})$, if we use $\mathcal{P}_V : \mathbb{R}^n \to V$ to denote the projection operator onto V, defined in $\mathbb{R}^n$. On the right-hand side of (S.6) (that is, the right-hand side of expression (7) of the main manuscript), the second term $-E\left[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}\right]$ amounts to removing the main effect of $\boldsymbol{\alpha}^\top \boldsymbol{X}$ that corresponds to the subspace V_0 in $\mathbb{R}^n$. We can decompose $\mathbb{R}^n$ as $\mathbb{R}^n = V_0^\perp \oplus V_0$, where $V_0^\perp$ denotes the orthogonal perpendicular subspace of V_0. Accordingly, using the least squares projection, we can decompose the vector $\boldsymbol{Y} \in \mathbb{R}^n$ of the observed outcomes as: $\boldsymbol{Y} = \mathcal{P}_{V_0^\perp}(\boldsymbol{Y}) + \mathcal{P}_{V_0}(\boldsymbol{Y}) = \mathcal{P}_V(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})) + \mathcal{P}_{V^\perp}(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})) + \mathcal{P}_{V_0}(\boldsymbol{Y}) = \sum_{t=1}^{L} \mathcal{P}_{V_t}(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})) + \mathcal{P}_{V^\perp}(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})) + \mathcal{P}_{V_0}(\boldsymbol{Y})$. The first term of the decomposition, $\sum_{t=1}^{L} \mathcal{P}_{V_t}(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y}))$, correspond to projecting $\boldsymbol{Y}$ on to the subspace $V_0^\perp \cap V$, which is the subspace representing the right-hand side of (S.6) in $\mathbb{R}^n$, in two steps: 1) projecting $\boldsymbol{Y}$ on $V_0^\perp$, and 2) projecting the projection on V. Since $\mathcal{P}_{V_0^\perp}(\boldsymbol{Y}) =$

$\Big(\boldsymbol{I}_n - \boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\Big)\boldsymbol{Y}$ and $\mathcal{P}_{V_t}(\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})) = \boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\mathcal{P}_{V_0^\perp}(\boldsymbol{Y})$, the right-hand side of (S.6) can be represented in $\mathbb{R}^n$ by $\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\Big(\boldsymbol{I}_n - \boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)}(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\Big)\boldsymbol{Y}$ $(t = 1, \dots, L)$. This gives expression (10) of the main manuscript for the approximation functions $(g_1, \dots, g_L)$.

A.3. Proof of Theorem 1 and 2 and Corollary 1

Under model (S.1),

$$\begin{aligned} E\left[Y \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] &= E\left[\mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] + E\left[f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] \quad (t=1,\dots,L) \\ &= E\left[\mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] + E\left[f_t(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] \quad (t=1,\dots,L) \end{aligned} \tag{S.7}$$

By the second line in (S.7), we can write

$$\begin{aligned} E\left[f_t(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] &= E\left[Y \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] - E\left[\mu(\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] \quad (t=1,\dots,L) \\ &= E\left[Y \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] - E\left[\mu(\boldsymbol{X}) + f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] \quad (t=1,\dots,L) \\ &= E\left[Y \mid \boldsymbol{\alpha}^\top\boldsymbol{X}, T=t\right] - E\left[Y \mid \boldsymbol{\alpha}^\top\boldsymbol{X}\right] \quad (t=1,\dots,L) \\ &= g^{**}_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}) - g^*_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^\top\boldsymbol{X}) \quad (t=1,\dots,L) \end{aligned} \tag{S.8}$$

where the second equality follows from the identifiability condition on $(f_1, \dots, f_L)$ in model (S.1), and the last line from the definitions of $g^{**}_{\boldsymbol{\alpha},t}$ and $g^*_{\boldsymbol{\alpha}}$ (see Assumption 5). For each fixed $\boldsymbol{\alpha} \in \Theta$, let us define the functions $f_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}) := g^{**}_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}) - g^*_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^\top\boldsymbol{X})$ $(t = 1, \dots, L)$, which, if $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, reduce to the true link functions $f_t(\boldsymbol{\alpha}_0^\top\boldsymbol{X})$ $(t = 1, \dots, L)$ of model (S.1). Under Assumption 3, $u = \boldsymbol{\alpha}^\top\boldsymbol{X}$ is bounded, and without loss of generality, we assume the domain of these functions to be $[0, 1]$. For the functions $g^{**}_{\boldsymbol{\alpha},t}(u)$ and $g^*_{\boldsymbol{\alpha}}(u)$, we introduce the associated estimators $\widehat{g}^{**}_{\boldsymbol{\alpha},t}(u) = B_t(u)^\top(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(t)\top}\boldsymbol{Y}$ $(t = 1, \dots, L)$ and $\widehat{g}^*_{\boldsymbol{\alpha}}(u) = B_0(u)^\top(\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)})^{-1}\boldsymbol{D}_{\boldsymbol{\alpha}}^{(0)\top}\boldsymbol{Y}$, respectively, and let $\widehat{f}_{\boldsymbol{\alpha},t}(u) := \widehat{g}^{**}_{\boldsymbol{\alpha},t}(u) - \widehat{g}^*_{\boldsymbol{\alpha}}(u)$ $(t = 1, \dots, L)$, $u \in [0, 1]$. Note, for each fixed $\boldsymbol{\alpha}$, these estimators $\widehat{f}_{\boldsymbol{\alpha},t}$ $(t = 1, \dots, L)$ correspond to the approximation functions g_t $(t = 1, \dots, L)$ that appear in (10) of the main manuscript.

Under Assumptions 1–4 of the main manuscript, by Proposition $A.1$ in the Supplementary Material of Wang and Yang (2009), we have

$$\sup_{\boldsymbol{\alpha}\in\Theta}\sup_{u\in[0,1]} \left|\widehat{g}^{**}_{\boldsymbol{\alpha},t}(u) - g^{**}_{\boldsymbol{\alpha},t}(u)\right| = O(n_t^{-1/2}d_t^{1/2}\log n_t + d_t^{-4}) \quad (t = 1, \dots, L), \tag{S.9}$$

almost surely. Similarly, under Assumptions 1–4 of the main manuscript, by Proposition $A.1$ in the Supplementary Material of Wang and Yang (2009), we have

$$\sup_{\boldsymbol{\alpha}\in\Theta}\sup_{u\in[0,1]} |\widehat{g}^*_{\boldsymbol{\alpha}}(u) - g^*_{\boldsymbol{\alpha}}(u)| = O(n^{-1/2}d_0^{1/2}\log n + d_0^{-4}), \tag{S.10}$$

almost surely.

Then, (S.9) and (S.10) entail that

$$\sup_{\boldsymbol{\alpha}\in\Theta}\sup_{u\in[0,1]} \left|\widehat{f}_{\boldsymbol{\alpha},t}(u) - f_{\boldsymbol{\alpha},t}(u)\right| = O\big(n_t^{-1/2}d_t^{1/2}\log n_t + d_t^{-4} + n^{-1/2}d_0^{1/2}\log n + d_0^{-4}\big) \quad (t = 1, \dots, L), \tag{S.11}$$

almost surely. For each $t = 1, \dots, L$, (S.11) implies that

$$\sup_{\boldsymbol{\alpha}\in\Theta} n^{-1}\sum_{i=1}^n \left|\widehat{f}_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i) - f_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\right| 1_{(T_i=t)} = O\big(n_t^{-1/2}d_t^{1/2}\log n_t + d_t^{-4} + n^{-1/2}d_0^{1/2}\log n + d_0^{-4}\big) \tag{S.12}$$

and

$$\sup_{\boldsymbol{\alpha}\in\Theta} n^{-1}\sum_{i=1}^{n}\big\{\widehat{f}_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},t}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big\}^2 1_{(T_i=t)} = O\big\{\big(n_t^{-1/2}d_t^{1/2}\log n_t+n^{-1/2}d_0^{1/2}\log n\big)^2+\big(d_t^{-4}+d_0^{-4}\big)^2\big\}, \tag{S.13}$$

almost surely.

Let us denote the empirical criterion function as $\widehat{Q}(\boldsymbol{\alpha}) := n^{-1}\sum_{i=1}^{n}\big(Y_i-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2 = n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i+f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2$, and consider the decomposition:

$$\begin{aligned}
\widehat{Q}(\boldsymbol{\alpha}) =& n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i+f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)+f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2\\
=& n^{-1}\sum_{i=1}^{n}(\mu(\boldsymbol{X}_i)+\epsilon_i)^2+n^{-1}\sum_{i=1}^{n}\big(f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2\\
&+n^{-1}\sum_{i=1}^{n}\big(f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2+2n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i\big)\big(f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)\\
&+2n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i+f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)\big(f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big).
\end{aligned}$$

Let us denote the corresponding population criterion function by $Q(\boldsymbol{\alpha}) := E[(Y-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X}))^2] = E[(\mu(\boldsymbol{X})+f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})+\epsilon-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X}))^2] = E[(\mu(\boldsymbol{X})+\epsilon)^2+(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X}))^2+2(\mu(\boldsymbol{X})+\epsilon)(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X})]$. Then, we can write

$$\sup_{\boldsymbol{\alpha}\in\Theta}\Big|\widehat{Q}(\boldsymbol{\alpha})-Q(\boldsymbol{\alpha})\Big| \le I_1+I_2+I_3+I_4+I_5 \quad \text{(almost surely)}, \tag{S.14}$$

in which

$$I_1 = \sup_{\boldsymbol{\alpha}\in\Theta}\left|n^{-1}\sum_{i=1}^{n}\big(f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2\right|,$$

$$I_2 = \sup_{\boldsymbol{\alpha}\in\Theta}\left|2n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i+f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)\big(f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)\right|,$$

$$I_3 = \sup_{\boldsymbol{\alpha}\in\Theta}\left|n^{-1}\sum_{i=1}^{n}\big(f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2-E\big[\big(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X})\big)^2\big]\right|,$$

$$I_4 = \sup_{\boldsymbol{\alpha}\in\Theta}\left|2n^{-1}\sum_{i=1}^{n}\big(\mu(\boldsymbol{X}_i)+\epsilon_i\big)\big(f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)-2E[(\mu(\boldsymbol{X})+\epsilon)(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})-f_{\boldsymbol{\alpha},T}(\boldsymbol{\alpha}^\top\boldsymbol{X})]\right|,$$

$$I_5 = \left|2n^{-1}\sum_{i=1}^{n}(\mu(\boldsymbol{X}_i)+\epsilon_i)^2-2E\left[(\mu(\boldsymbol{X})+\epsilon)^2\right]\right|.$$

The strong law of large numbers implies that $I_3+I_4+I_5=o(1)$, almost surely. Also,

$$\begin{aligned}
I_1 &\le \sum_{t=1}^{L}\sup_{\boldsymbol{\alpha}\in\Theta}\left|n^{-1}\sum_{i=1}^{n}\big(f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\big)^2 1_{(T_i=t)}\right|\\
&= \sum_{t=1}^{L} O\big\{\big(n_t^{-1/2}d_t^{1/2}\log n_t+n^{-1/2}d_0^{1/2}\log n\big)^2+\big(d_t^{-4}+d_0^{-4}\big)^2\big\},
\end{aligned} \tag{S.15}$$

almost surely, by (S.13). And,

$$I_2 \le \sup_{\boldsymbol{\alpha}\in\Theta} 2n^{-1}\sum_{i=1}^{n}\left|f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)-\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)\right| \times \sup_{\boldsymbol{\alpha}\in\Theta} n^{-1}\sum_{i=1}^{n}\left|\mu(\boldsymbol{X}_i)+\epsilon_i+f_{T_i}(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i)-f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i)\right|$$
$$= \sum_{t=1}^{L} O\big(n_t^{-1/2}d_t^{1/2}\log n_t + d_t^{-4} + n^{-1/2}d_0^{1/2}\log n + d_0^{-4}\big)\times O(1), \tag{S.16}$$

almost surely, by (S.12) and the strong law of large numbers. Since we choose the numbers of interior knots, d_t $(t=1,\dots,L)$ and d_0, for the cubic spline smoothing under Assumption 5 of the main manuscript, we obtain

$$\sup_{\boldsymbol{\alpha}\in\Theta}\left|\widehat{Q}(\boldsymbol{\alpha})-Q(\boldsymbol{\alpha})\right| \to 0, \tag{S.17}$$

almost surely. Now, we prove the consistency of $\widehat{\boldsymbol{\alpha}}_0$ to $\boldsymbol{\alpha}_0$. Denote by $(\Omega,\mathcal{F},\mathcal{P})$ the probability space on which all $\{Y_i,T_i,\boldsymbol{X}_i^\top\}_{i=1}^{\infty}$ are defined. By (S.17), for any $\delta>0$, $\omega\in\Omega$, there is an integer $n^*(\omega)$, such that $\widehat{Q}(\boldsymbol{\alpha}_0,\omega)-Q(\boldsymbol{\alpha}_0)<\delta/2$, whenever $n>n^*(\omega)$. Since $\widehat{\boldsymbol{\alpha}}_0(\omega)$ is the minimizer of $\widehat{Q}(\boldsymbol{\alpha},\omega)$, we have $\widehat{Q}(\widehat{\boldsymbol{\alpha}}_0(\omega),\omega)-Q(\boldsymbol{\alpha}_0)<\delta/2$. Also, by (S.17), there exists an integer $n^{**}(\omega)$, such that $Q(\widehat{\boldsymbol{\alpha}}_0(\omega),\omega)-Q(\widehat{\boldsymbol{\alpha}}_0(\omega),\omega)<\delta/2$, whenever $n>n^{**}(\omega)$. Therefore, whenever $n>\max(n^*(\omega),n^{**}(\omega))$, we have $\widehat{Q}(\widehat{\boldsymbol{\alpha}}_0(\omega),\omega)-Q(\boldsymbol{\alpha}_0)<\delta$. The strong consistency $\widehat{\boldsymbol{\alpha}}_0\to\boldsymbol{\alpha}_0$ follows from the local convexity of Assumption 2 of the main manuscript.

The proof of Corollary 1 follows from (S.11) that, under Assumptions 1–4 of the main manuscript, $\sup_{u\in[0,1]}\left|\widehat{f}_{\boldsymbol{\alpha},t}(u)-f_{\boldsymbol{\alpha},t}(u)\right| = O\big(n_t^{-1/2}d_t^{1/2}\log n_t + d_t^{-4} + n^{-1/2}d_0^{1/2}\log n + d_0^{-4}\big)$, almost surely, for each $t=1,\dots,L$, for any $\boldsymbol{\alpha}\in\Theta$. Taking $\boldsymbol{\alpha}=\boldsymbol{\alpha}_0$ and the strong consistency $\widehat{\boldsymbol{\alpha}}_0\to\boldsymbol{\alpha}_0$ (Theorem 1) imply Corollary 1, under Assumptions 1–5 of the main manuscript.

For the proof of Theorem 2, we first note that $\boldsymbol{\alpha}\in\Theta\subset\mathbb{R}^p$ of the working model (S.2) can be expressed as: $\boldsymbol{\alpha}(=c(\boldsymbol{\phi}))=(1,\boldsymbol{\phi}^\top)^\top/(1+\|\boldsymbol{\phi}\|^2)^{1/2}$, for a $p-1$ dimensional vector $\boldsymbol{\phi}=(\phi_1,\dots,\phi_{p-1})^\top\in\mathbb{R}^{p-1}$, as stated in the main manuscript of the paper. Let $\boldsymbol{J}(\boldsymbol{\phi})$ denote the $p\times(p-1)$ Jacobian 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 $(i=1;\ j=1,\dots,p-1)$, and $\partial\alpha_i/\partial\phi_j=-\alpha_i\phi_j/K^2+1/K$, for $(i=2,\dots,p;\ j=1,\dots,p-1)$, in which $K=(1+\|\boldsymbol{\phi}\|^2)^{1/2}$. The relation $\boldsymbol{\alpha}=c(\boldsymbol{\phi})$ is one-to-one, hence the parameter $\boldsymbol{\phi}_0\in\mathbb{R}^{p-1}$ that corresponds to $\boldsymbol{\alpha}_0\in\Theta$ can be specified.

The estimator $\widehat{\boldsymbol{\alpha}}_0$ $(=c(\widehat{\boldsymbol{\phi}}_0))$ in (11) of the main manuscript corresponds to the solution of the set of score estimating equations:

$$\frac{\partial}{\partial\alpha_j}\widehat{Q}(\boldsymbol{\alpha}) = 2n^{-1}\sum_{i=1}^{n}\{\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-Y_i\}\frac{\partial}{\partial\alpha_j}\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)=0 \quad (j=1,\dots,p) \tag{S.18}$$

subject to the identifiability constraint $\boldsymbol{\alpha}\in\Theta$. In (S.18), again, $\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)=g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)$ for each fixed $\boldsymbol{\alpha}$, where the nonparametrically-defined functions $g_t(\cdot)$ $(t=1,\dots,L)$ are given by (10) of the main manuscript, under Assumption 5 on the numbers of interior knots for the cubic spline approximation. In particular, under Assumptions 1 and 3–5 of the main manuscript, by Lemma A.15 of Wang and Yang (2007), we have:

$$\sup_{\boldsymbol{\alpha}\in\Theta}\sup_{1\le j\le p}\left|\frac{\partial}{\partial\alpha_j}\left\{\widehat{Q}(\boldsymbol{\alpha})-Q(\boldsymbol{\alpha})\right\}-n^{-1}\sum_{i=1}^{n}\xi_{\boldsymbol{\alpha},i,j}\right| = o(n^{-1/2}), \tag{S.19}$$

almost surely, where

$$\xi_{\boldsymbol{\alpha},i,j} := 2\{f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-Y_i\}\frac{\partial}{\partial\alpha_j}f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-\frac{\partial}{\partial\alpha_j}Q(\boldsymbol{\alpha}).$$

Note, $E[\xi_{\boldsymbol{\alpha},i,j}]=E\left[2\{f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)-Y_i\}\frac{\partial}{\partial\alpha_j}f_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\right]-\frac{\partial}{\partial\alpha_j}Q(\boldsymbol{\alpha})=0$, and thus $\xi_{\boldsymbol{\alpha},i,j}$ $(i=1,\dots,n)$ are mean-zero independent random variables, for each $j=1,\dots,p$. If $\frac{\partial}{\partial\alpha_j}\left\{\widehat{Q}(\boldsymbol{\alpha})-Q(\boldsymbol{\alpha})\right\}$ is evaluated at the minimum $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$ (i.e., $\boldsymbol{\phi} = \boldsymbol{\phi}_0$), then by the local convexity of Assumption 2, we have $\left\{\frac{\partial}{\partial \alpha_j} Q(\boldsymbol{\alpha})\right\}\big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} = 0$. Thus, at $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, (S.19) gives:

$$\sup_{1\leq j\leq p} \left| \frac{\partial}{\partial \alpha_j} \widehat{Q}(\boldsymbol{\alpha}_0) - n^{-1}\sum_{i=1}^{n} \xi_{\boldsymbol{\alpha}_0,i,j} \right| = o(n^{-1/2}), \tag{S.20}$$

almost surely, under Assumptions 1–5. Recall for model identifiability, we restrict $\boldsymbol{\alpha} \in \Theta$. By incorporating the identifiability constraint through the "delete-one-component" reparametrization, the score estimating equations (S.18), $\frac{\partial}{\partial \boldsymbol{\alpha}^\top}\widehat{Q}(\boldsymbol{\alpha}) = \mathbf{0}$ subject to $\boldsymbol{\alpha} \in \Theta$, are written by:

$$n^{-1}\sum_{i=1}^{n} \widehat{\Psi}\left(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}\right) = \mathbf{0} \tag{S.21}$$

in which $\widehat{\Psi}(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}) = 2\boldsymbol{J}^\top(\boldsymbol{\phi})\left\{\widehat{f}_{\boldsymbol{\alpha},T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i) - Y_i\right\}\frac{\partial}{\partial \boldsymbol{\alpha}^\top}\widehat{f}_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)$, where $\boldsymbol{\alpha} = c(\boldsymbol{\phi})$. If evaluated at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$, the left-hand side of (S.21), i.e., $n^{-1}\sum_{i=1}^{n}\widehat{\Psi}\left(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0\right) = 2\boldsymbol{J}^\top(\boldsymbol{\phi}_0)\frac{\partial}{\partial \boldsymbol{\alpha}^\top}\widehat{Q}(\boldsymbol{\alpha}_0)$, is a linear transformation of the length-p gradient vector $\frac{\partial}{\partial \boldsymbol{\alpha}^\top}\widehat{Q}(\boldsymbol{\alpha}_0) = \left(\frac{\partial}{\partial \alpha_1}\widehat{Q}(\boldsymbol{\alpha}_0), \frac{\partial}{\partial \alpha_2}\widehat{Q}(\boldsymbol{\alpha}_0), \ldots, \frac{\partial}{\partial \alpha_p}\widehat{Q}(\boldsymbol{\alpha}_0)\right)^\top$, which is represented, up to $o(n^{-1/2})$ almost surely, by a sum of mean-zero independent random vectors $\boldsymbol{\xi}_{\boldsymbol{\alpha}_0,i} := \left(\xi_{\boldsymbol{\alpha}_0,i,1}, \xi_{\boldsymbol{\alpha}_0,i,2}, \ldots, \xi_{\boldsymbol{\alpha}_0,i,p}\right)^\top \in \mathbb{R}^p$ $(i = 1, \ldots, n)$, by (S.20). Cramér-Wold device and the central limit theorem entail that $n^{-1/2}\sum_{i=1}^{n}\widehat{\Psi}(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\Sigma}_0 = \boldsymbol{J}^\top(\boldsymbol{\phi}_0)\mathrm{var}(\boldsymbol{\xi}_{\boldsymbol{\alpha}_0,i})\boldsymbol{J}(\boldsymbol{\phi}_0)$, with $\boldsymbol{\xi}_{\boldsymbol{\alpha}_0,i} = 2\{f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) - Y_i\}\frac{\partial}{\partial \boldsymbol{\alpha}^\top}f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) = 2\{f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) - Y_i\}\dot{f}_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)\boldsymbol{X}_i \in \mathbb{R}^p$. Let $\Psi(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0) := \boldsymbol{J}^\top(\boldsymbol{\phi}_0)\boldsymbol{\xi}_{\boldsymbol{\alpha}_0,i} = 2\boldsymbol{J}^\top(\boldsymbol{\phi}_0)\{f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) - Y_i\}\dot{f}_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)\boldsymbol{X}_i$, where $\boldsymbol{\alpha}_0 = c(\boldsymbol{\phi}_0)$, and let $\boldsymbol{A}_0$ be the $(p-1)\times(p-1)$ matrix of the first derivative of $E\left[\Psi(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi})\right]$ with respect to $\boldsymbol{\phi} \in \mathbb{R}^{p-1}$, evaluated at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$. Taking the first-order Taylor series expansion of the left-hand side of the equations, $n^{-1/2}\sum_{i=1}^{n}\widehat{\Psi}(Y_i, T_i, \boldsymbol{X}_i|\widehat{\boldsymbol{\phi}}_0) = \mathbf{0}$, at $\widehat{\boldsymbol{\phi}}_0 = \boldsymbol{\phi}_0$ and rearranging leads to:

$$n^{1/2}(\widehat{\boldsymbol{\phi}}_0 - \boldsymbol{\phi}_0) = -\left[\frac{\partial}{\partial \boldsymbol{\phi}^\top}\{n^{-1}\sum_{i=1}^{n}\widehat{\Psi}(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi})\}\bigg|_{\boldsymbol{\phi}=\boldsymbol{\phi}^*}\right]^{-1} n^{-1/2}\sum_{i=1}^{n}\widehat{\Psi}(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0), \tag{S.22}$$

where $\boldsymbol{\phi}^*$ is between $\boldsymbol{\phi}_0$ and $\widehat{\boldsymbol{\phi}}_0$. The uniform consistency of the observed Hessian in (S.22) to $\boldsymbol{A}_0$ is obtained by Lemma A.15 of Wang and Yang (2007) under Assumptions 1 and 3–5, $\sup_{\boldsymbol{\alpha}\in\Theta}\sup_{1\leq q,j\leq p}\left|\frac{\partial^2}{\partial \alpha_q \partial \alpha_j}\{\widehat{Q}(\boldsymbol{\alpha}) - Q(\boldsymbol{\alpha})\}\right| = o(1)$ (almost surely), and that $\widehat{\boldsymbol{\phi}}_0 \to \boldsymbol{\phi}_0$ (almost surely) implied by Theorem 1 of the main manuscript. Through Slutsky's theorem, (S.22) implies that $n^{1/2}(\widehat{\boldsymbol{\phi}}_0 - \boldsymbol{\phi}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}_0\boldsymbol{A}_0^{-1\top})$. Then an application of the multivariate delta method with mapping $\boldsymbol{\alpha}_0 = c(\boldsymbol{\phi}_0)$ entails that $n^{1/2}(\widehat{\boldsymbol{\alpha}}_0 - \boldsymbol{\alpha}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{J}_0\boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}_0\boldsymbol{A}_0^{-1\top}\boldsymbol{J}_0^\top)$, where $\boldsymbol{J}_0 = \boldsymbol{J}(\boldsymbol{\phi}_0)$.

A.4. Efficiency augmentation

As referenced in Section 3.3 of the main manuscript, in this subsection we consider an approach that is analogous to the efficiency augmentation of Tian et al. (2014) which improves the original estimator $\widehat{\boldsymbol{\alpha}}_0$ of $\boldsymbol{\alpha}_0$. We consider an augmented estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$ for $\boldsymbol{\alpha}_0$ of model (S.1):

$$\widehat{\boldsymbol{\alpha}}_{0,m} = \underset{\boldsymbol{\alpha}\in\Theta}{\mathrm{argmin}}\; n^{-1}\sum_{i=1}^{n}\left[\left\{Y_i - g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\right\}^2/2 \;+\; g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)m(\boldsymbol{X}_i)\right], \tag{S.23}$$

where the original objective function in (11) of the main manuscript (Section 3.3) is augmented by the term $n^{-1}\sum_{i=1}^{n} g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)m(\boldsymbol{X}_i)$, in which the function $m(\cdot) : \mathbb{R}^p \mapsto \mathbb{R}$ represents an arbitrary continuous function. The strong consistency of the augmented estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$ for $\boldsymbol{\alpha}_0$ of model (1) in the main manuscript (i.e., of model (S.1)) is given in the following Corollary.

Corollary 2. *Under Assumptions 1–5 in the main manuscript, $\widehat{\boldsymbol{\alpha}}_{0,m} \to \boldsymbol{\alpha}_0$, almost surely, for any continuous function $m(\cdot) : \mathbb{R}^p \mapsto \mathbb{R}$.*

Given the consistency of the estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$ to $\boldsymbol{\alpha}_0$ of model (S.1), we now consider selecting an optimal augmentation function $m(\cdot)$ in (S.23), where the associated augmented estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$ may have a smaller asymptotic variance than that of the original estimator $\widehat{\boldsymbol{\alpha}}_0$ given in (11) of the main manuscript.

Corollary 3. *Under Assumptions 1–5 in the main manuscript, the optimal choice of the function $m(\cdot)$ that results in the smallest asymptotic variance of $\widehat{\boldsymbol{\alpha}}_{0,m}$ in (S.23) satisfies $m(\boldsymbol{X}) = E[Y|\boldsymbol{X}]$, almost surely.*

The proofs of Corollary 2 and 3 are given in Section A.5. Corollary 3 indicates that the optimal choice of $m(\cdot)$ is the function $\mu(\cdot)$ associated with the unspecified $\boldsymbol{X}$ main effect in the true model (S.1).

Next, we describe how to construct an augmented estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$. The optimal augmentation term $m(\boldsymbol{X}) = E[Y|\boldsymbol{X}]$ $(= \mu(\boldsymbol{X}))$ can be approximated by a (possibly misspecified) regression model $E[Y|\boldsymbol{X}] \approx B(\boldsymbol{X})^\top \boldsymbol{\eta}$, where $B(\boldsymbol{X})$ and $\boldsymbol{\eta}$ are a design function and a vector of coefficients, respectively. As given by Corollary 2, the appealing feature of $\widehat{\boldsymbol{\alpha}}_{0,m}$ in (S.23) is that the estimator is robust to the misspecification of the optimal augmentation function $m(\boldsymbol{X}) = E[Y|\boldsymbol{X}]$, in terms of its consistency to $\boldsymbol{\alpha}_0$. In the special case of taking $m(\boldsymbol{X}) = 0$, the augmented estimator, $\widehat{\boldsymbol{\alpha}}_{0,m}$, which provides a means of incorporating the term $\mu(\boldsymbol{X})$ in model (S.1) to the estimation of $\boldsymbol{\alpha}_0$ through a specification of $m(\boldsymbol{X})$ in (S.23), reduces to the original estimator $\widehat{\boldsymbol{\alpha}}_0$ in (11) of the main manuscript.

Let us use $\widehat{\boldsymbol{\eta}}$ to denote an estimate for $\boldsymbol{\eta}$ of model $E[Y|\boldsymbol{X}] \approx m(\boldsymbol{X}) = B(\boldsymbol{X})^\top \boldsymbol{\eta}$, estimated based on the pairs $\{(Y_i, \boldsymbol{X}_i), i = 1, \ldots, n\}$. We can obtain an estimate $\widehat{\boldsymbol{\eta}}$, before fitting the working model (S.2). Given an estimate $\widehat{\boldsymbol{\eta}}$, the objective function of (S.23) becomes:

$$n^{-1} \sum_{i=1}^{n} \left[\left\{ Y_i - g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i) \right\}^2 /2 + g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i) B(\boldsymbol{X}_i)^\top \widehat{\boldsymbol{\eta}} \right] \;\propto\; n^{-1} \sum_{i=1}^{n} \left[\left\{ Y_i - B(\boldsymbol{X}_i)^\top \widehat{\boldsymbol{\eta}} - g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i) \right\}^2 \right].$$

The form of the right-hand side indicates that the same iterative procedure described in Section 3.2 of the main manuscript can be used to obtain the augmented estimator $\widehat{\boldsymbol{\alpha}}_{0,m}$ in (S.23), by using the residualized responses $\widetilde{Y}_i = Y_i - B(\boldsymbol{X}_i)^\top \widehat{\boldsymbol{\eta}}$ $(i = 1, \ldots, n)$, instead of the original responses Y_i $(i = 1, \ldots, n)$.

A.5. Proofs of Corollary 2 and 3

We first prove Corollary 2. Let $\widehat{Q}_m(\boldsymbol{\alpha})$ and $\widehat{Q}(\boldsymbol{\alpha})$ denote the objective functions of (S.23) and (11) in the main manuscript, respectively. For a fixed set of the link-functions $(g_1, \ldots, g_L)$ satisfying (S.3), $\widehat{Q}_m(\boldsymbol{\alpha})$ in (S.23) converges pointwise to $E\big[\widehat{Q}(\boldsymbol{\alpha})\big]$ for each $\boldsymbol{\alpha} \in \Theta$, almost surely, by the law of large numbers and the fact that $E\big[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) m(\boldsymbol{X}) | \boldsymbol{X}\big] = m(\boldsymbol{X}) E\big[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) | \boldsymbol{X}\big] = 0$, almost surely, which follows from the constraint (S.3) on $(g_1, \ldots, g_L)$. Furthermore, we have

$$\left\| \frac{\partial}{\partial \boldsymbol{\alpha}^\top} \widehat{Q}_m(\boldsymbol{\alpha}) \right\| \;=\; \left\| n^{-1} \sum_{i=1}^{n} \left\{ g_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i) + m(\boldsymbol{X}_i) - Y_i \right\} \dot{g}_{T_i}(\boldsymbol{\alpha}^\top \boldsymbol{X}_i) \boldsymbol{X}_i \right\| \;\le\; C \qquad \text{(S.24)}$$

for some constant $C > 0$ and large n, since $\boldsymbol{X}$ is bounded, $\boldsymbol{\alpha}$ is in a compact set Θ, the link-functions $(g_1, \ldots, g_L)$ are smooth functions with continuous second derivatives, and the function m is a continuous function. Therefore, $\widehat{Q}_m(\boldsymbol{\alpha}) \to E\big[\widehat{Q}(\boldsymbol{\alpha})\big]$ uniformly over $\boldsymbol{\alpha} \in \Theta$, almost surely, for a fixed set of the link-functions $(g_1, \ldots, g_L)$ satisfying (S.3). Furthermore, $E\big[\widehat{Q}(\boldsymbol{\alpha})\big] \to Q(\boldsymbol{\alpha})$ uniformly over $\boldsymbol{\alpha} \in \Theta$ almost surely under Assumptions 1–5 of the main manuscript, by the uniform convergence rate (S.11) of the nonparametric component $g_{\boldsymbol{\alpha},t} (= \widehat{f}_{\boldsymbol{\alpha},t})$ to $f_{\boldsymbol{\alpha},t}$. Therefore, $\widehat{Q}_m(\boldsymbol{\alpha}) \to Q(\boldsymbol{\alpha})$ uniformly over $\boldsymbol{\alpha} \in \Theta$, almost surely. By using the same argument as in the last part of the proof of Theorem 1, we obtain $\widehat{\boldsymbol{\alpha}}_{0,m} \to \boldsymbol{\alpha}_0$, almost surely.

Next we prove Corollary 3. The augmented estimator, $\widehat{\boldsymbol{\alpha}}_{0,m} (= c(\widehat{\boldsymbol{\phi}}_{0,m}))$ which appears on the left-hand side of (S.23), can be viewed as the solution to the following augmented estimating equations:

$n^{-1}\sum_{i=1}^{n}\widehat{\Psi}_m\left(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}\right) = \mathbf{0}$ (corresponding to the first-order condition), in which

$$\begin{aligned}\widehat{\Psi}_m(Y_i, T_i, \boldsymbol{X}_i \mid \boldsymbol{\phi}) &= \boldsymbol{J}^\top(\boldsymbol{\phi})\big\{g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i) - Y_i\big\}\dot{g}_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\boldsymbol{X}_i + \boldsymbol{J}^\top(\boldsymbol{\phi})m(\boldsymbol{X}_i)\dot{g}_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\boldsymbol{X}_i \\ &= \boldsymbol{J}^\top(\boldsymbol{\phi})\big\{g_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i) + m(\boldsymbol{X}_i) - Y_i\big\}\dot{g}_{T_i}(\boldsymbol{\alpha}^\top\boldsymbol{X}_i)\boldsymbol{X}_i,\end{aligned}$$

where $\boldsymbol{\alpha} = c(\boldsymbol{\phi})$. The same argument as in the proof of Theorem 2 is applied to the left-hand side of the estimation equations $n^{-1}\sum_{i=1}^{n}\widehat{\Psi}_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}) = \mathbf{0}$, but now $Y_i - m(\boldsymbol{X}_i)$ taking the role of Y_i in (S.21). Cramér-Wold device and the central limit theorem imply that $n^{-1/2}\sum_{i=1}^{n}\widehat{\Psi}_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{0,m})$, where $\boldsymbol{\Sigma}_{0,m} = \mathrm{var}\big(\Psi_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0)\big)$, in which $\Psi_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0) := \boldsymbol{J}^\top(\boldsymbol{\phi}_0)\big\{f_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i) + m(\boldsymbol{X}_i) - Y_i\big\}\dot{f}_{T_i}(\boldsymbol{\alpha}_0^\top\boldsymbol{X}_i)\boldsymbol{X}_i$, where $\boldsymbol{\alpha}_0 = c(\boldsymbol{\phi}_0)$. Let $\boldsymbol{A}_{0,m}$ denote the matrix of the first derivative of the function $E\left[\Psi_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi})\right]$ (with respect to $\boldsymbol{\phi}$) evaluated at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$. Further, let $\ddot{f}_t(\cdot)$ denote the second derivative of $f_t(\cdot)$ (with respect to $\cdot$). Note, $E\left[\boldsymbol{J}^\top(\boldsymbol{\phi}_0)m(\boldsymbol{X})\ddot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})\boldsymbol{X}|\boldsymbol{X}\right] = \boldsymbol{J}^\top(\boldsymbol{\phi}_0)m(\boldsymbol{X})\boldsymbol{X}E\left[\ddot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})|\boldsymbol{X}\right] = 0$, almost surely, by the identifiability condition of model (S.1) imposed on the link functions $(f_1, \ldots, f_L)$, with the result that the Hessian matrix $\boldsymbol{A}_{0,m}$ does not depend on the augmentation function $m(\cdot)$, and in fact $\boldsymbol{A}_{0,m} = \boldsymbol{A}_0$. By taking the first-order Taylor series expansion on the left-hand side of the equations, $n^{-1/2}\sum_{i=1}^{n}\widehat{\Psi}_m(Y_i, T_i, \boldsymbol{X}_i|\widehat{\boldsymbol{\phi}}_{0,m}) = \mathbf{0}$, at $\widehat{\boldsymbol{\phi}}_{0,m} = \boldsymbol{\phi}_0$, we obtain that $n^{1/2}(\widehat{\boldsymbol{\phi}}_{0,m} - \boldsymbol{\phi}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}_{0,m}\boldsymbol{A}_0^{-1\top})$. An application of the multivariate delta method with mapping $\boldsymbol{\alpha}_0 = c(\boldsymbol{\phi}_0)$ entails that $n^{1/2}(\widehat{\boldsymbol{\alpha}}_{0,m} - \boldsymbol{\alpha}_0)$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{J}_0\boldsymbol{A}_0^{-1}\boldsymbol{\Sigma}_{0,m}\boldsymbol{A}_0^{-1\top}\boldsymbol{J}_0^\top)$, where $\boldsymbol{J}_0 = \boldsymbol{J}(\boldsymbol{\phi}_0)$. It follows that selecting the asymptotically optimal augmentation term in (S.23) corresponds to finding a function $m(\cdot)$ that minimizes the variance of $\Psi_m(Y_i, T_i, \boldsymbol{X}_i|\boldsymbol{\phi}_0)$, i.e., $\boldsymbol{\Sigma}_{0,m}$. We can decompose the matrix $\boldsymbol{\Sigma}_{0,m}$ by

$$\begin{aligned}\boldsymbol{\Sigma}_{0,m} &= \boldsymbol{J}^\top(\boldsymbol{\phi})E\big[E\big[\big\{\big(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) + m(\boldsymbol{X}) - Y\big)\dot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})\boldsymbol{X}\big\}^{\otimes 2} \mid \boldsymbol{X}, T\big]\big]\boldsymbol{J}(\boldsymbol{\phi}) \\ &= \boldsymbol{J}^\top(\boldsymbol{\phi})E\big[E\big[\big\{\big(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) + m_0(\boldsymbol{X}) - Y\big)\dot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})\boldsymbol{X}\big\}^{\otimes 2}|\boldsymbol{X}, T\big]\big]\boldsymbol{J}(\boldsymbol{\phi}) \\ &\quad + \boldsymbol{J}^\top(\boldsymbol{\phi})E\big[\big\{\big(m(\boldsymbol{X}) - m_0(\boldsymbol{X})\big)\dot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})\boldsymbol{X}\big\}^{\otimes 2}\big]\boldsymbol{J}(\boldsymbol{\phi}),\end{aligned} \tag{S.25}$$

where $m_0(\cdot)$ corresponds to a function that satisfies:

$$E\left[\big(f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) + m_0(\boldsymbol{X}) - Y\big)\eta\big(\dot{f}_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X})\boldsymbol{X}\big) \mid \boldsymbol{X}, T\right] = 0 \quad \text{(almost surely)}, \tag{S.26}$$

for any arbitrary square-integrable function $\eta(\cdot)$. The second term in representation (S.25) implies that the optimal $m(\cdot)$ that minimizes the variance must satisfy $m(\boldsymbol{X}) = m_0(\boldsymbol{X})$, almost surely. Condition (S.26) implies that the function $m_0(\cdot)$, given $\boldsymbol{X}$ and T, must satisfy:

$$m_0(\boldsymbol{X}) = E\left[Y|\boldsymbol{X}, T\right] - f_T(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) \quad \text{(almost surely)}. \tag{S.27}$$

Integrating the both sides of (S.27) with respect to the distribution of T (given $\boldsymbol{X}$) gives $m_0(\boldsymbol{X}) = \sum_{t=1}^{L}\pi_t E\left[Y|T = t, \boldsymbol{X}\right] - \sum_{t=1}^{L}\pi_t f_t(\boldsymbol{\alpha}_0^\top\boldsymbol{X}) = E\left[Y|\boldsymbol{X}\right]$ (almost surely), where the second equality comes from the identifiability condition on the link functions $(f_1, \ldots, f_L)$ of model (S.1). It follows that $m_0(\boldsymbol{X}) = E\left[Y|\boldsymbol{X}\right]$ corresponds to the optimal augmentation function.

A.6. Robustness with respect to misspecified $\boldsymbol{X}$ main effects when T depends on $\boldsymbol{X}$

In observational studies, the treatment assignment and covariates are generally correlated. In this subsection, we provide a justification for the utility of the working model (S.2) in estimating the T-by-$\boldsymbol{X}$ interaction effects, even when there is a correlation between the covariates and the treatment assignment. As defined in (7) of the main manuscript, for each fixed $\boldsymbol{\alpha} \in \Theta$, let the link functions $g_t(\cdot)$ $(t = 1, \ldots, L)$ of the working model (S.2) satisfy:

$$g_T(\boldsymbol{\alpha}^\top\boldsymbol{X}) = E[Y|\boldsymbol{\alpha}^\top\boldsymbol{X}, T] - E[Y|\boldsymbol{\alpha}^\top\boldsymbol{X}]. \tag{S.28}$$

Here, we illustrate that the function (S.28) is designed to satisfy the constraint $E[g_T(\boldsymbol{\alpha}^\top\boldsymbol{X})|\boldsymbol{X}] = 0$ that appears in the constrained optimization (6) of the main manuscript, and therefore, the associated estimator of the single-index coefficient $\boldsymbol{\alpha}_0$, i.e., $\text{argmin}_{\boldsymbol{\alpha}\in\Theta} E[(Y - g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}))^2]$ where the link functions $(g_1, \ldots, g_L)$ are defined according to (S.28), is still robust to the misspecification of $\mu(\boldsymbol{X})$ of the underlying model (S.1). Consider the following equality:

$$E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})|\boldsymbol{X}] = E[E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}]|\boldsymbol{X}] \quad (\forall \boldsymbol{\alpha} \in \Theta), \tag{S.29}$$

where we apply the iterated expectation rule to condition on $\boldsymbol{\alpha}^\top \boldsymbol{X}$ on the right-hand side of (S.29). Equation (S.29) implies that the condition $E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})|\boldsymbol{\alpha}^\top \boldsymbol{X}] = 0$ $(\forall \boldsymbol{\alpha} \in \Theta)$ is a sufficient condition for the constraint $E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})|\boldsymbol{X}] = 0$ $(\forall \boldsymbol{\alpha} \in \Theta)$.

On the other hand, for the function $g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})$ defined in (S.28), we have:

$$\begin{aligned} E\big[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] &= E\big[E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T] - E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\ &= E\big[E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}] - E[E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}] \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] \\ &= E\big[Y \mid \boldsymbol{\alpha}^\top \boldsymbol{X}] - E[Y \mid \boldsymbol{\alpha}^\top \boldsymbol{X}\big] = 0, \end{aligned}$$

i.e., $E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \mid \boldsymbol{\alpha}^\top \boldsymbol{X}] = 0$ $(\forall \boldsymbol{\alpha} \in \Theta)$. Therefore, from the right-hand side of (S.29), we obtain the desired constraint $E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})|\boldsymbol{X}] = 0$ $(\forall \boldsymbol{\alpha} \in \Theta)$. This constraint implies $E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})\mu(\boldsymbol{X})] = E[E[g_T(\boldsymbol{\alpha}^\top \boldsymbol{X})|\boldsymbol{X}]\mu(\boldsymbol{X})] = 0$ $(\forall \boldsymbol{\alpha} \in \Theta)$, i.e.,

$$g_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) \perp \mu(\boldsymbol{X}) \quad (\forall \boldsymbol{\alpha} \in \Theta). \tag{S.30}$$

Thus, in the iterative optimization procedure, for each candidate $\boldsymbol{\alpha} \in \Theta$, the estimator $\widehat{g}_T(\boldsymbol{\alpha}^\top \boldsymbol{X}) = \widehat{g}_T^{**}(\boldsymbol{\alpha}^\top \boldsymbol{X}) - \widehat{g}^*(\boldsymbol{\alpha}^\top \boldsymbol{X})$ (as defined in the discussion section of the main manuscript) for the link function (S.28) can be used as the working link functions g_t $(t = 1, \ldots, L)$ in (11) of the main manuscript. Then, by the orthogonality (S.30), the associated profile minimizer $\widehat{\boldsymbol{\alpha}}_0$ of the objective (11) is (asymptotically) robust to the misspecification of $\mu(\boldsymbol{X})$ of the underlying model (S.1). Therefore, the working model (S.2) can still be useful in fitting the interaction effect term of model (S.1) even when $\boldsymbol{X}$ and T are correlated, as it side-steps the issue that would arise (i.e., severe inconsistency of the estimators of the interaction term) when the $\boldsymbol{X}$ main effect $\mu(\boldsymbol{X})$ is misspecified. The simulation example in Section B.3 considers a scenario where the covariates and the treatment assignment are correlated, and a close-to-optimal performance of the proposed regression approach to optimizing individualized treatment rules is reported. However, the approach generally results in biased causal effect estimates and sub-optimal individualized treatment rules if T depends on $\boldsymbol{X}$, as discussed in the next subsection.

A.7. Suboptimality of the individualized treatment rules when T depends on $\boldsymbol{X}$

Let $Y^{(t)} \in \mathbb{R}$ be the potential outcome under treatment $T = t$ (as defined in Section 2 of the main manuscript), and suppose the standard assumptions of causal inference (Rubin, 1978) hold: Assumption 1) consistency, i.e., $T = t$ implies $Y = Y^{(t)}$; Assumption 2) no unmeasured confounders, i.e., T is independent of $Y^{(t)}$ given $\boldsymbol{X}$; Assumption 3) positivity, i.e., for every covariate $\boldsymbol{X}$, the probability of receiving every level of treatment is positive.

Under such assumptions, the functions $g_t(\boldsymbol{\alpha}^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$ on the right-hand side of (3) of the main manuscript can be defined in terms of the potential outcome framework, as:

$$g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) := E[Y^{(t)}|\boldsymbol{\alpha}^\top \boldsymbol{X}] - E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}] \quad \text{(almost surely)} \quad (t = 1, \ldots, L). \tag{S.31}$$

If the treatment T is independent of $\boldsymbol{X}$ (as can happen in randomized studies), (S.31) reduces to (7) of the main manuscript, i.e.,

$$g_t(\boldsymbol{\alpha}^\top \boldsymbol{X}) = E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}, T = t] - E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}] \quad \text{(almost surely)} \quad (t = 1, \ldots, L), \tag{S.32}$$

which one can consistently estimate from observed data, for example, using spline smoothing, given each fixed $\boldsymbol{\alpha}$. However, if T depends on $\boldsymbol{X}$ (as in observational studies), the expression for $g_t(\boldsymbol{\alpha}^\top \boldsymbol{X})$ $(t = 1, \ldots, L)$

on the right-hand side of (S.32) is generally not valid. To elaborate on this, under the consistency assumption (Assumption 1), the right-hand side of (S.32) can be written as:

$$E[Y^{(t)}|\boldsymbol{\alpha}^\top \boldsymbol{X}, T=t] - E[Y|\boldsymbol{\alpha}^\top \boldsymbol{X}] \quad \text{(almost surely)} \quad (t=1,\ldots,L). \tag{S.33}$$

Although the no unmeasured confounder assumption (Assumption 2) implies that $Y^{(t)} \perp T$ given $\boldsymbol{X}$, in general, $Y^{(t)}$ and T need not be independent each other, given only $\boldsymbol{\alpha}^\top \boldsymbol{X}$ (as in the case of (S.33)). Therefore, expression (S.33) is not, in general, equal to the right-hand side of (S.31). It follows that, in observational studies, even if one could consistently estimate the right-hand side of (S.32), the estimators would not be generally consistent for the functions (S.31). Thus, the associated treatment decision rules are potentially suboptimal in the context of observational studies.

Web Appendix B: Results from simulation studies

B.1. Constrained single-index regression for $L = 3$ treatment level case

In this section, we provide additional simulation results to investigate the performance of constrained single-index model for estimating optimal treatment decision rules when the number of treatment options $L = 3$. We consider $n \in \{250, 500\}$ and $p \in \{10, 20\}$, with a varying main effect intensity $\delta \in \{1, 2\}$. 100 training datasets are simulated for each scenario. We generate covariates $\boldsymbol{X}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_p)$, and treatments $T_i \in \{1,2,3\}$ with equal probability at random, independently of $\boldsymbol{X}_i$. We generate outcomes from model (S.1): $Y_i = \mu(\boldsymbol{X}_i) + f_{T_i}(h(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i)) + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, 0.4^2)$. We set $\boldsymbol{\alpha}_0 = (1, 0.5, 0.25, 0.125, 0, \ldots, 0)^\top \in \mathbb{R}^p$ normalized to have unit L^2 norm. The treatment-specific functions $f_t(u)$ $(t = 1, 2, 3)$, $u \in [0, 1]$ are set to be

$$\begin{cases} f_1(u) &= u^1(1-u)^4/B(2,5) - f_0(u) \\ f_2(u) &= u^1(1-u)^1/B(2,4) - f_0(u) \\ f_3(u) &= u^4(1-u)^0/B(5,1) - f_0(u), \end{cases} \tag{S.34}$$

where $B(a,b) = (\Gamma(a)\Gamma(b))/\Gamma(a+b)$ is a Beta function, and $f_0(u) := \{u^1(1-u)^4/B(2,5) + u^1(1-u)^1/B(2,4) + u^4(1-u)^0/B(5,1)\}/3$. The functions in (S.34) are illustrated in Figure S.1.

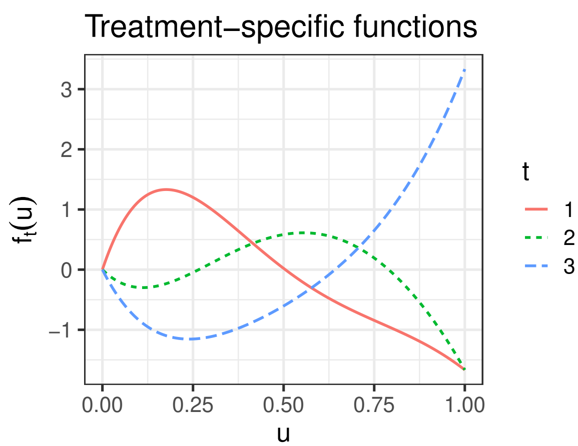

Figure S.1: Illustration of the treatment t-specific functions $f_t(u)$ $(t = 1, 2, 3)$ used to generate the treatment t-specific responses in the $L = 3$ treatment group case.

The "centering" function $f_0(u)$ is introduced to satisfy the identifiability condition on the treatment-specific functions f_t of model (S.1). In (S.34), $u = h(\boldsymbol{\alpha}_0^\top \boldsymbol{X})$, where h is the cumulative distribution function of a re-scaled and centered $B(3,3)$, that is, $h(s) = 0.9375 \int_{-1}^{s/r} (1-s^2)^2 ds$, $s \in [-r, r]$, in which r is the maximum of $\{|\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i|, i = 1, \ldots, n\}$. This transformation h makes the distribution of u relatively uniformly distributed in $[0, 1]$. Figure S.1 displays the treatment-specific functions $f_t(u)$ $(t = 1, 2, 3)$, $u \in [0, 1]$ of (S.34). The main effect component $\mu(\boldsymbol{X})$ in model (S.1) is set to be

$$\mu(\boldsymbol{X}; \delta) = 2\delta \cos(\boldsymbol{\eta}^\top \boldsymbol{X}), \tag{S.35}$$

where the scaling parameter $\delta \in \{1, 2\}$ regulates the contribution of the $\boldsymbol{X}$ main effect on the variance of Y, in which $\delta = 1$ represents a relatively *moderate* main effect (contributing about the same variance as the interaction effect does) and $\delta = 2$ a relatively *big* main effect case (about 4 times larger), respectively. In (S.35), we set the coefficient $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{10}, 0, \ldots, 0)^\top \in \mathbb{R}^p$, in which the vector $(\eta_1, \ldots, \eta_{10})^\top \in \mathbb{R}^{10}$ is generated independently from a multivariate Gaussian distribution, and is then rescaled to have unit L^2 norm, for each simulation run. Without loss of generality, we assume that higher values of Y are preferred.

When $L > 2$, a common regression approach to model the interaction effects between T and $\boldsymbol{X}$ on an outcome is to fit a regression model separately for each of the L treatment groups, as functions of $\boldsymbol{X}$. For instance, one can fit a linear model (or a single-index model) to estimate the functions $E[Y|\boldsymbol{X}, T = t]$ $(t = 1, \ldots, L)$, separately for each treatment group t. The corresponding estimators, $\widehat{\mathcal{D}}^{opt}$, of $\mathcal{D}^{opt}$ can be set to be $\widehat{\mathcal{D}}^{opt}(\boldsymbol{X}) = \arg\max_{t\in\{1,\ldots,L\}} \widehat{E}[Y|\boldsymbol{X}, T = t]$, in which each of the functions $\widehat{E}[Y|\boldsymbol{X}, T = t]$ $(t = 1, \ldots, L)$ is estimated from a treatment t-specific linear regression or a treatment t-specific single-index regression. These two treatment-specific regression-based estimators are compared to the proposed estimator of $\mathcal{D}^{opt}$ that utilizes the constrained single-index model (S.1).

For each simulation run, we estimate $\mathcal{D}^{opt}$ from each of the 3 methods based on a training set (of size n), and for evaluation of these methods, we estimate the value $V(\widehat{\mathcal{D}}^{opt})$ of each estimate $\widehat{\mathcal{D}}^{opt}$ by

$$\widehat{V}(\widehat{\mathcal{D}}^{opt}) = \sum_{i=1}^{\widetilde{n}} Y_i 1_{(T_i = \widehat{\mathcal{D}}^{opt}(\boldsymbol{X}_i))} \Big/ \sum_{i=1}^{\widetilde{n}} 1_{(T_i = \widehat{\mathcal{D}}^{opt}(\boldsymbol{X}_i))}, \tag{S.36}$$

computed based on a testing set of size $\widetilde{n} = 10^5$. Since we know the true data generating model in simulation studies, the optimal individualized treatment rule $\mathcal{D}^{opt}$ can be determined for each simulation run. Given each estimate $\widehat{\mathcal{D}}^{opt}$ for $\mathcal{D}^{opt}$, we report $\widehat{V}(\widehat{\mathcal{D}}^{opt}) - \widehat{V}(\mathcal{D}^{opt})$ calculated based on (S.36), as the performance measure of the estimator $\widehat{\mathcal{D}}^{opt}$. A larger value of the measure indicates a better performance.

Figure S.2 displays the boxplots of the value ratios of the individualized treatment rules estimated from the 3 methods (the constrained single-index model, the L separate linear regression models, and the L separate single-index models), for each combination of $n \in \{250, 500\}$, $p \in \{10, 20\}$ and the main effect intensity parameter $\delta \in \{1, 2\}$.

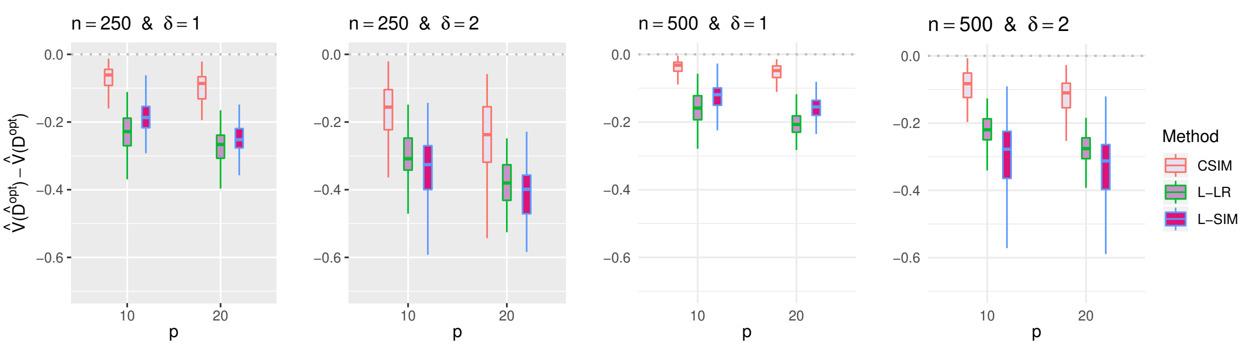

Figure S.2: Boxplots comparing 3 approaches to estimating $\mathcal{D}^{opt}$, given each scenario indexed by $\delta \in \{1, 2\}$ and $p \in \{10, 20\}$, for the $L = 3$ treatment level cases. For each scenario, from left to right, estimation approaches for $\mathcal{D}^{opt}$: 1) the constrained single-index model (red); 2) the L separate linear models (green); 3) the L separate single-index models (violet). The case with $\delta = 1$ (or $\delta = 2$) corresponds to the moderate (or large) main effect scenario; $p = 10$ (or $p = 20$) corresponds to the dimension of $\boldsymbol{X}$. The dotted horizontal line represents the optimal value corresponding to $\mathcal{D}^{opt}$.

The boxplots indicates that the proposed constrained single-index model outperforms all other methods, in all cases. In particular, when $L > 2$, estimating L separate regression models lacks parsimony and interpretability, whereas the constrained single-index model provides a single projection $\boldsymbol{\alpha}^\top \boldsymbol{X}$ that captures the variability in $\boldsymbol{X}$ related to the treatment effect-modification, in the presence of an unspecified main effects of $\boldsymbol{X}$. When the $\boldsymbol{X}$ main effect dominates the T-by-$\boldsymbol{X}$ interaction effect (when $\delta = 2$), these L separate regressions tend to focus more on capturing the main effect of $\boldsymbol{X}$ and therefore missing the important T-by-$\boldsymbol{X}$ interaction effect. On the other hand, the constrained single-index model consistently targets the interaction effect. As a result, in Figure S.2, although the increased magnitude of the main effect affects the performance of all methods, it has least effect on the constrained single-index model, and more effect on the L separate regression approaches.

B.2. Comparison to the doubly robust estimation

In this subsection, we compare our method to the augmented inverse probability weighted estimator (AIPWE) of Zhang et al. (2012). Here, we follow the exact simulation scenario (with slightly different notation) reported in Appendix D of Zhang et al. (2012) and compare their results to the results of using constrained single-index model (S.1). We generate $n = 500$ observations $(Y_i, T_i, \boldsymbol{X}_i)$, $i = 1, \ldots, n$. The correct model for the outcome in their setting is $Y_i = \mu_c(\boldsymbol{X}_i, T_i) + \epsilon_i$, for ϵ_i standard normal and

$$\mu_c(\boldsymbol{X}, T) = E\left[Y|\boldsymbol{X}, T\right] = \exp\left\{(2 - 1.5X_1^2 - 1.5X_2^2 + 3X_1X_2 + (T-1)(-0.1 - X_1 + X_2 + 0.2X_3)\right\},$$

where the treatment random variable T takes a value in $\{1, 2\}$ with equal probability, independently of the 3 covariates $\boldsymbol{X} = (X_1, X_2, X_3)^\top$ with $X_1, X_2 \sim \text{unif}[-1.5, 1.5]$ and $X_3 \sim \text{Bernoulli}(0.5)$. Note, this simulation setting includes a discrete-valued covariate, which violates Assumption 4 of the main manuscript. Further, in this setting, the proposed model (S.1) is misspecified, since the $\boldsymbol{X}$ main effect term and the T-by-$\boldsymbol{X}$ interaction effect term of the data generating model $\mu_c(\boldsymbol{X}, T)$ are associated in a multiplicative manner, rather than in an additive manner as in model (S.1) (notice the $\exp\{\cdot\}$ function in the data generating model). The optimal individualized treatment rule according to the true model μ_c is $\mathcal{D}^{opt}(\boldsymbol{X}) = 1_{(-0.1 - X_1 + X_2 + 0.2X_3 > 0)} + 1$, and the corresponding optimal value (obtained by (S.36) based on Monte Carlo simulation using 10^6 replicates) is $\widehat{V}(\mathcal{D}^{opt}) = 3.95$.

Zhang et al. (2012) compared several estimators of the optimal treatment decision: i) a direct regression estimator (RGμ_c) based on the correct model μ_c; ii) a direct regression estimator using a misspecified linear model for the outcome (RGμ_m), containing only the main effects of the $p = 3$ covariates; iii) the inverse probability weighted estimator (IPWE); iv) the AIPWE with the correct model (AIPWEμ_c); and v) the AIPWE with a misspecified regression linear model (AIPWEμ_m) which will be used in the usual case of the absence of prior knowledge. In Table S.1, these methods are compared to the approaches of using the constrained single-index model (S.1) and the modified covariates (MC) model (2) in the main manuscript to estimating $\mathcal{D}^{opt}$, with respect to the values $\widehat{V}(\widehat{\mathcal{D}}^{opt})$ defined in (S.36), computed based on an independently generated Monte Carlo simulation using 10^6 replicates of $(Y_i, T_i, \boldsymbol{X}_i)$. The quantities reported in Zhang et al. (2012) are indicated by †. In this example, to avoid being trapped into a local minimum in optimizing the working model (S.2), we used 12 different starting points for $\boldsymbol{\alpha}$ by setting an initial candidate solution $\boldsymbol{\alpha}_{\text{ini}}^{(k)} = (1, u_1, u_2)$, normalized to have unit L^2 norm, with $u_1, u_2 \sim \text{unif}[-1, 1]$, for each $k = 1, \ldots, 12$. We then chose among these 12 fits the model giving the minimal residual sum of squares.

The results in Table S.1 indicate that constrained single-index model (CSIM) performance is close to optimal (with the estimated value 3.90) and outperforms AIPWEμ_m (the last column), which will be used in the usual case of the absence of prior knowledge on the outcome models. Importantly, the constrained single-index model is computationally fast and can handle high dimensional settings whereas the AIPWE approach quickly becomes infeasible as the dimension increases using the grid search described in Zhang et al. (2012).

B.3. Comparison to the doubly robust estimation for observational study setting

In this subsection, we compare the constrained single-index model to AIPWE of Zhang et al. (2012), by conducting a simulation that mimics an observational study in which the treatment assignment mechanism

Table S.1: Comparison of the constrained single-index model and the modified covariates model with AIPWE, with respect to the value of the individualized treatment rules. The optimal value is $\widehat{V}(\mathcal{D}^{opt}) = 3.95$. AIPWE$\mu_m$ is the AIPWE method that will be used in practice.

	CSIM	MC	RGμ_c	RGμ_m	IPWE	AIPWEμ_c	AIPWEμ_m
$\widehat{V}(\widehat{\mathcal{D}}^{opt})$ (SD)	3.90 (0.02)	3.65 (0.10)	†3.95 (0.00)	†3.66 (0.10)	†3.84 (0.08)	†3.94 (0.01)	†3.89 (0.07)

depends on $\boldsymbol{X}$. As in Section B.2, we follow the exact simulation scenario (with slightly different notation) reported in Section 4 (Simulation Studies) of Zhang et al. (2012) and compare their results to the results of using the constrained single-index model (S.1). We generate $n = 500$ independent observations $(Y_i, T_i, \boldsymbol{X}_i)$, $i = 1, \ldots, n$. The true model for the outcome in their setting is $Y_i = \mu_c(\boldsymbol{X}_i, T_i) + \epsilon_i$, for ϵ_i standard normal and

$$\mu_c(\boldsymbol{X}_i, T_i) = \exp\left\{2 - 1.5X_{i,1}^2 - 1.5X_{i,2}^2 + 3X_{i,1}X_{i,2} + (T_i - 1)(-0.1 - X_{i,1} + X_{i,2})\right\},$$

where the 2 covariates $\boldsymbol{X}_i = (X_{i,1}, X_{i,2})^\top$ with $X_{i,1}, X_{i,2} \sim$ unif$[-1.5, 1.5]$, and the treatment $T_i \sim$ Bernoulli$(\pi_i) + 1 \in \{1, 2\}$, where $\pi_i = \exp(-1 + 0.8X_{i,1}^2 + 0.8X_{i,2}^2)/\{1 + \exp(-1 + 0.8X_{i,1}^2 + 0.8X_{i,2}^2)\}$. The optimal individualized treatment rule according to the true model μ_c is $\mathcal{D}^{opt}(\boldsymbol{X}_i) = 1_{(X_{i,2} > X_{i,1} + 0.1)} + 1$, and the corresponding optimal Value (obtained by (S.36) based on Monte Carlo simulation using 10^6 replicates) is $\widehat{V}(\mathcal{D}^{opt}) = 3.71$.

We report several estimators of the optimal individualized treatment rule $\mathcal{D}^{opt}$ presented in Zhang et al. (2012): i) the inverse probability weighted estimator (IPWE); ii) the AIPWE with the correctly specified outcome model (AIPWEμ_c); and iii) the AIPWE with a misspecified (linear) outcome model (AIPWEμ_m). For these three estimators, results are shown using both correct and incorrect models for the propensity score (PS) of treatment assignments: a) the correctly specified model ("PS correct"); b) a misspecified logistic regression linear model ("PS incorrect").

In Table S.2, these methods are compared to the approaches of using the constrained single-index model (S.1) and the modified covariates model (2) in the main manuscript to estimating $\mathcal{D}^{opt}$, with respect to the values $\widehat{V}(\widehat{\mathcal{D}}^{opt})$ defined in (S.36), computed based on an independently generated Monte Carlo simulation using 10^6 replicates of $(Y_i, T_i, \boldsymbol{X}_i)$. For the modified covariates method, we treat T_i as independent of $\boldsymbol{X}_i$, and π_1 in model (2) of the main manuscript is estimated by $n^{-1}\sum_{i=1}^n 1_{(T_i=1)}$. As discussed in Section 6 of the main manuscript, the constrained single-index regression does not require to postulate a model for the propensity score, however, it requires estimators $(g_1, \ldots, g_L)$ that asymptotically satisfy (7) of the main manuscript (for each fixed $\boldsymbol{\alpha}$). We use the estimators (10) in the main manuscript for such $(g_1, \ldots, g_L)$.

To avoid being trapped into a local minimum in optimizing the working model (S.2), as in Section , we used 12 different starting points for $\boldsymbol{\alpha}$ by setting an initial candidate solution $\boldsymbol{\alpha}_{\text{ini}}^{(k)} = (1, u)$, normalized to have unit L^2 norm, with $u \sim$ unif$[-1, 1]$ for each $k = 1, \ldots, 12$, and then chose among these 12 fits the model giving the minimal residual sum of squares. The quantities reported in the Table 1 of Zhang et al. (2012) are indicated by †.

The results in Table S.2 indicate that the constrained single-index model performance (with the estimated value 3.66) is close to optimal, and performs similarly or outperforms AIPWEμ_m (the 5th and 8th columns in Table S.2), which will be used in the usual case of the absence of prior knowledge on the outcome models. The results suggest that utilizing model (S.2) in estimating the interaction effects can still be effective in the context of observational studies.

Table S.2: Comparison of the constrained single-index model and the modified covariates model with AIPWE, with respect to value of individualized treatment rules. The optimal value is $\widehat{V}(\mathcal{D}^{opt}) = 3.71$. AIPWE$\mu_m$ is the AIPWE method that will be used in practice.

			propensity score correct			propensity score incorrect		
	CSIM	MC	IPWE	AIPWEμ_c	AIPWEμ_m	IPWE	AIPWEμ_c	AIPWEμ_m
$\widehat{V}(\widehat{\mathcal{D}}^{opt})$ (SD)	3.66 (0.09)	3.38 (0.16)	†3.63 (0.07)	†3.70 (0.01)	†3.66 (0.07)	†3.42 (0.20)	†3.70 (0.01)	†3.57 (0.20)

B.4. Implementational detail of the penalized additive cubic spline least squares approach

In this subsection, as referenced in Section 4 of the main manuscript, we provide the implementational detail of the penalized additive cubic spline least squares (PLS) approach considered in Section 4 and 5 of the main manuscript. As indicated in Section 4 of the main manuscript, we implement this method by estimating $E[Y|\boldsymbol{X}, T = t]$ via an additive regression for each treatment separately. We use a set of additive cubic spline bases $\{b(X_{ij}^{(t)}), j = 1, \ldots, p\}$ $(t = 1, 2)$, where $b(\cdot)$ denotes a 8-dimensional cubic spline basis defined over the range of its argument, and $X_{ij}^{(t)}$ represents the jth covariate for the ith subject assigned with the tth treatment. We use the integrated squared second derivative cubic spline penalty, with the smoothing parameter (for each treatment) estimated by the generalized cross-validation (GCV). For each treatment, each additive term is subject to a sum-to-zero identifiability constraint and an additional intercept term is included (which is not penalized). The method is implemented through the R-package `mgcv` (Wood, 2019).

B.5. Correlated covariates case

In Section 4 of the main manuscript, all covariates are generated independently. In this section, we consider the case where there is a substantial amount of correlation among the covariates $\boldsymbol{X}$. We consider the same settings as Section 4 of the main manuscript, i.e., the simulation settings "A" and "B", but with dependent covariates $\boldsymbol{X}$. To generate correlated $\boldsymbol{X}$, first, we independently generate a set of n p-dimensional normal random vectors from the zero mean and unit variance, with pairwise correlation 0.5. Next, we transform each of the p variables by using the standard normal cumulative distribution function (CDF), and obtain a set of n p-dimensional standard uniform correlated random variables; then these are shifted and scaled to give a set of n p-dimensional correlated unif$[-\pi/2, \pi/2]$ variables $\boldsymbol{X}_i \in \mathbb{R}^p$ $(i = 1, \ldots, n)$. In the both simulation settings "A" and "B" in Section 4 of the main manuscript, the main effect function $\mu(\boldsymbol{X}, \delta)$ is a function of $\boldsymbol{\eta}^\top \boldsymbol{X}$, where we set $\boldsymbol{\eta} = (-1, 1, -1, 1, -1, 1, 0, 0, 0, 0)^\top$. In this correlated $\boldsymbol{X}$ setting, however, there is a substantial amount of positive correlation among the covariates $\boldsymbol{X}$. Therefore, the variance of $\boldsymbol{\eta}^\top \boldsymbol{X}$ for the correlated $\boldsymbol{X}$ case is substantially smaller than that of $\boldsymbol{\eta}^\top \boldsymbol{X}$ for the independent $\boldsymbol{X}$ case (due to substantial cancelation by the particular linear combination $\boldsymbol{\eta}$ that consists of -1 and 1 that appear consecutively to each other), resulting in a much smaller magnitude of the $\boldsymbol{X}$ main effect as compared to that of the settings considered in Section 4 of the main manuscript. Thus, in this section with correlated $\boldsymbol{X}$, we randomly generate the first 6 elements $\eta_j \sim$ unif$[-1, 1]$ $(j = 1, \ldots, 6)$ and take $(\eta_1, \ldots, \eta_6, 0, 0, 0, 0)^\top \in \mathbb{R}^{10}$, which is then scaled to have a unit L^2 norm, as the index vector $\boldsymbol{\eta} \in \mathbb{R}^{10}$ associated with the $\boldsymbol{X}$ main effect, for each simulation run. All other parameters are set as in Section 4 of the main manuscript.

As in Section 4 of the main manuscript, we present the boxplots, obtained from 100 simulation runs, of the (centered) values $\widehat{V}(\widehat{\mathcal{D}}^{opt}) - \widehat{V}(\mathcal{D}^{opt})$ of the individualized treatment rules $\widehat{\mathcal{D}}^{opt}$ estimated from the 5 approaches, for each combination of $n \in \{250, 500\}$, $\delta \in \{1, 2\}$ (corresponding to *moderate* or *large* main effects, respectively) and $\xi \in \{0, 0.5\}$ (corresponding to *correctly-specified* or *mis-specified* single-index interaction effect models, respectively), for the simulation set "A" in the top panels and the set "B" in the bottom panels.

The results in Figure S.3 indicate that the proposed constrained single-index model (CSIM) outperforms all other approaches in estimating $\mathcal{D}^{opt}$, except in simulation set A for the case of a relatively small sample size $(n = 250)$, particularly when the underlying interaction effect model deviates from the single-index model $(\xi = 0.5$, i.e., the model is misspecified); however, as the sample size increases to $n = 500$, the proposed method outperforms all other approaches. With substantial nonlinearity in the interaction effect term in the both settings A and B, the modified covariates model, which assumes a restricted linear model on the interaction term, is clearly outperformed by the proposed approach that utilizes a set of flexible link functions to accommodate the nonlinear treatment effect modification. When $n = 500$ (i.e., with a relatively large training sample size) and $\xi = 0.5$ (i.e., when the underlying model deviates from the single-index structure), the penalized additive spline approach (PLS), due to its large model space, outperforms the modified covariates approach; however, the penalized additive spline approach is outperformed by the proposed constrained single-index regression method, which is robust to the $\boldsymbol{X}$ main effect model misspecification. When the $\boldsymbol{X}$ main effect dominates the T-by-$\boldsymbol{X}$ interaction effect (i.e., when $\delta = 2$), although the increased magnitude of

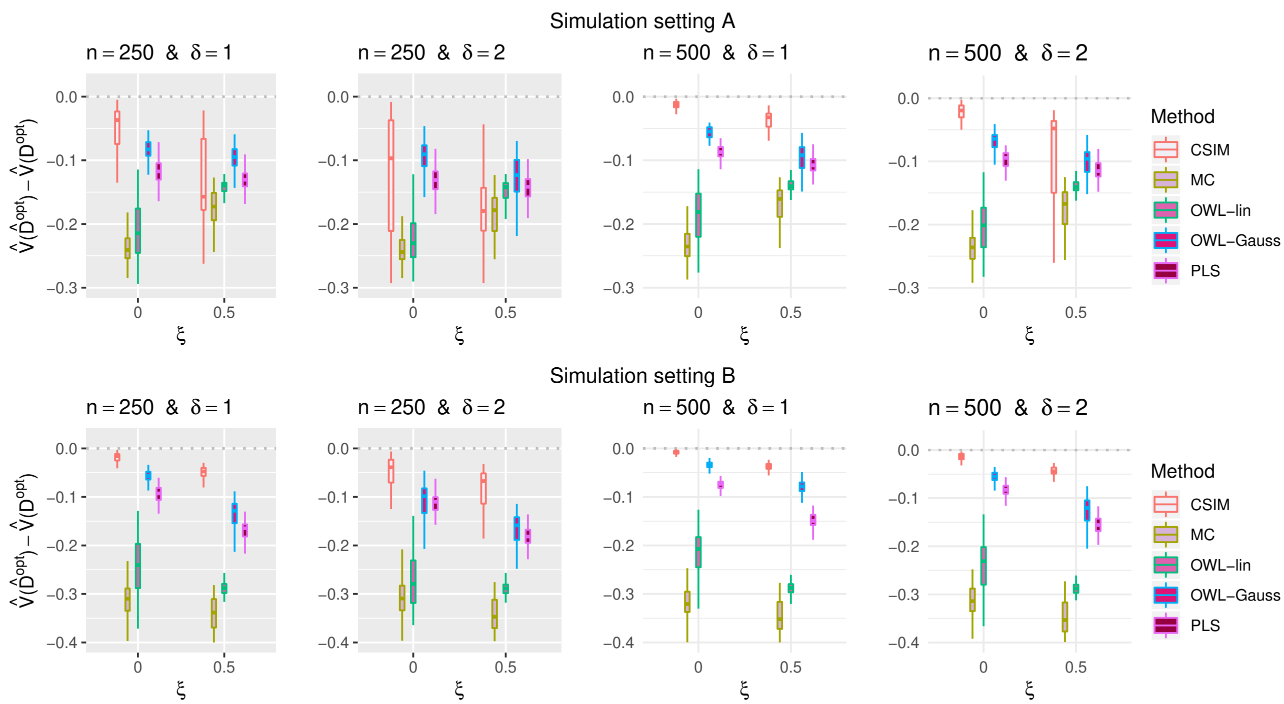

Figure S.3: Boxplots comparing 5 approaches to estimating $\mathcal{D}^{opt}$, given each scenario indexed by $\delta \in \{1, 2\}$ and $\xi \in \{0, 0.5\}$, for the simulation setting "A" in the top panels and the setting "B" in the bottom panels, for the correlated covariate cases. For each scenario, 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 $\mathcal{D}^{opt}$.

the main effect dampens the performance of all approaches to estimating optimal individualized treatment rules, the proposed approach consistently targets the interaction effect, and its performance is near optimal when $n = 500$.

B.6. Coverage probability of normal-approximated bootstrap confidence intervals

In Section 5 of the main manuscript, we compute a normal-theory based 95% bootstrap confidence interval $(\widehat{\alpha}_{0j} - 1.96\sqrt{\widehat{\text{var}}[\widehat{\alpha}_{0j}]},\ \widehat{\alpha}_{0j} + 1.96\sqrt{\widehat{\text{var}}[\widehat{\alpha}_{0j}]})$ for α_{0j} $(j = 1, \ldots, p)$ of the single-index coefficient vector $\boldsymbol{\alpha}_0 = (\alpha_{01}, \ldots, \alpha_{0p})^\top$ of the underlying model (S.1), where $\widehat{\text{var}}[\widehat{\alpha}_{0j}]$ denotes the sampling variance estimate of $\widehat{\alpha}_{0j}$ $(j = 1, \ldots, p)$, obtained based on 500 bootstrap replicates of $\widehat{\boldsymbol{\alpha}}_0$. This normal-approximated confidence interval is based on the result of Theorem 2 of the main manuscript that the estimator $\widehat{\alpha}_{0j}$ for α_{0j} is asymptotically normally distributed. However, a proof that this is a valid approach to construct confidence intervals is not currently available; thus this should be regarded as heuristics. In particular, this approach may not yield accurate 95% confidence intervals if some bootstrap replicates of $\widehat{\boldsymbol{\alpha}}_0$ fall into local optimums, inflating the estimate of sampling variance.

To evaluate the coverage probability of these normal-theory bootstrap confidence intervals, we conduct a set of simulation experiments. We consider the setting of Simulation set "A" in Section 4 of the main manuscript, with $\delta = 1$ and $\xi = 0$, i.e., the outcome Y is generated from:

$$Y_i = 0.8\cos(\boldsymbol{\eta}^\top \boldsymbol{X}_i) + (-1)^{T_i}(e^{-(\boldsymbol{\alpha}_0^\top \boldsymbol{X}_i - 0.5)^2} - 0.5) + \epsilon_i \quad (i = 1, \ldots, n), \tag{S.37}$$

in which $\boldsymbol{\eta} = (-1, 1, -1, 1, -1, 1, 0, 0, 0, 0)^\top$ and $\boldsymbol{\alpha}_0 = (1, 0.5, 0.25, 0.125, 0, 0, 0, 0, 0, 0)^\top$, each normalized to have unit norm. We vary the sample size $n \in \{50, 100, 200, 400, 800, 1600\}$. In (S.37), the term $\mu(\boldsymbol{X}) := 0.8\cos(\boldsymbol{\eta}^\top \boldsymbol{X})$ represents the $\boldsymbol{X}$ main effect whereas the term $f_T(\boldsymbol{\alpha}_0^\top \boldsymbol{X}) := (-1)^T(e^{-(\boldsymbol{\alpha}_0^\top \boldsymbol{X} - 0.5)^2} - 0.5)$ represents the T-by-$\boldsymbol{X}$ interaction effect, and the aforementioned normal-theory bootstrap confidence intervals are constructed for the single-index coefficient $\boldsymbol{\alpha}_0 = (\alpha_{01}, \ldots, \alpha_{0p})^\top$ associated with the interaction effect term in (S.37).

In Table S.3, although the "actual" coverage probabilities appear to get relatively closer to the "nominal" coverage probability of 0.95 as the sample size n increases, they are often larger than their nominal level, i.e., the bootstrap confidence intervals are often too wide. This is because some of the bootstrap replicates of the estimator $\widehat{\boldsymbol{\alpha}}_0$ fall into their local optima; ideally, all of the 500 bootstrap replicates of $\widehat{\boldsymbol{\alpha}}_0$ should be from the global minima of the corresponding bootstrapped criterion functions. We note that the convergence of the iterative algorithm to global minimums depends on initialization. For binary treatment cases, throughout the paper (and in obtaining the results given in Table S.3), we have used the least square estimate of $\boldsymbol{\alpha}$ (i.e., the minimizer of (15)) of the modified covariates linear model (2) of the main manuscript as an initial value of the iterative procedure. However, this initialization sometimes yields bootstrap replications of $\widehat{\boldsymbol{\alpha}}_0$ that are only local (i.e., not global) minima of the associated bootstrapped criterion functions, that are often "outliers" in comparison to the rest of the bootstrap replications, resulting in an inflated bootstrap sampling variance estimate of the estimator $\widehat{\boldsymbol{\alpha}}_0$. In obtaining the results of Table S.3, there appeared to be a handful of such outliers out of 500 bootstrap replications of $\widehat{\boldsymbol{\alpha}}_0$ (note, 500 such replications are computed for each of the 200 simulation runs in each scenario) that impact the coverage probability.

In contrast to the simulations that gave Table S.3, for the simulation results in Table S.4, we have used as an initial value for $\widehat{\boldsymbol{\alpha}}_0$, a value that is close to the true coefficient vector $\boldsymbol{\alpha}_0 = (1, 0.5, 0.25, 0.125, 0, 0, 0, 0, 0, 0)^\top \in \mathbb{R}^{10}$ (normalized to have unit norm), rather than using the MC linear model estimate as an initial value. Specifically, we have used the vector $(1, 0.7, 0.4, 0.1, 0.05, 0.05, 0.05, 0.05, 0.05)^\top \in \mathbb{R}^{10}$ (normalized to have unit norm) as an initial estimate of the iterative procedure to estimate the true coefficient $\boldsymbol{\alpha}_0$. The idea is that, by starting from a value that is close to $\boldsymbol{\alpha}_0$ (which is feasible only in a simulation setting), one could minimize the chance of $\widehat{\boldsymbol{\alpha}}_0$ missing the global optimum in this simulation experiment. In practice, one way to mitigate this problem of getting trapped into local optima (as mentioned in the Discussion section of the main manuscript) is to consider multiple initial values in estimation. We did not pursue this approach in this simulation experiment, as it becomes computationally expensive: multiple starting points (say, 100) $\times$ 500 bootstrap replications for estimating the sampling variance × 200 such simulation runs for computing the coverage proportion, for each of the 6 scenarios.

In comparison to the results in Table S.3, the results in Table S.4 appear to provide better coverage proportions that are close to the nominal level of 0.95. This suggests that, provided that the iterative procedure for obtaining $\widehat{\boldsymbol{\alpha}}_0$ avoids getting trapped in a local optimum, the normal-theory based bootstrap confidence intervals can be reasonably accurate.

Table S.3: The proportion of time (out of 200 simulation runs) that the normal-theory based bootstrap 95% confidence interval contains the true value of α_j $(j = 1, \ldots, 10)$, with varying $n \in \{50, 100, 200, 400, 800, 1600\}$, when the modified covariate model estimate is used as an initial value for the iterative procedure for obtaining $\widehat{\boldsymbol{\alpha}}_0$.

n	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
50	0.545	0.735	0.910	0.965	0.990	0.980	0.975	0.975	0.940	0.950
100	0.865	0.840	0.915	0.930	0.960	0.955	0.975	0.940	0.955	0.960
200	0.990	0.935	0.965	0.970	0.960	0.960	0.960	0.970	0.985	0.980
400	0.990	0.990	0.995	0.980	0.985	0.990	0.990	0.995	0.990	0.985
800	0.950	0.970	0.950	0.975	0.970	0.975	0.960	0.985	0.955	0.980
1600	0.975	0.960	0.960	0.965	0.970	0.960	0.960	0.960	0.980	0.960

Table S.4: The proportion of time (out of 200 simulation runs) that the normal-theory based bootstrap 95% confidence interval contains the true value of α_j $(j = 1, \ldots, 10)$, with varying $n \in \{50, 100, 200, 400, 800, 1600\}$, when a value close to the true coefficient $\boldsymbol{\alpha}_0$ is used as an initial value for the iterative procedure for obtaining $\widehat{\boldsymbol{\alpha}}_0$.

n	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
50	0.765	0.870	0.950	0.965	0.970	0.985	0.975	0.960	0.980	0.945
100	0.950	0.945	0.975	0.940	0.985	0.960	0.980	0.955	0.975	0.960
200	0.995	0.975	0.975	0.960	0.960	0.975	0.970	0.965	0.980	0.990
400	0.955	0.970	0.970	0.955	0.960	0.960	0.970	0.955	0.950	0.965
800	0.940	0.940	0.920	0.970	0.945	0.945	0.935	0.965	0.930	0.950
1600	0.940	0.950	0.955	0.950	0.950	0.965	0.955	0.960	0.950	0.975

References

Rubin, D. (1978). Bayesian inference for causal effects: the role of randomization. *The Annals of Statistics* **6,** 34–58.

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.

Wang, L. and Yang, L. (2007). Spline single-index prediction model. *Technical Report. https://arxiv.org/abs/0704.0302* .

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

Wood, S. N. (2019). mgcv: Mixed GAM computation vehicle with automatic smoothness estimation. *R package version 1.8.28* .

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