Biometrics 000, 000–000 DOI: 000
000 0000

Supplementary Materials for "Weighted Q-learning for optimal dynamic treatment regimes with nonignorable missing covariates"

Jian Sun[1], Bo Fu[1,*], and Li Su[2]

[1]School of Data Science, Fudan University, Shanghai 200433, China

[2]MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, Cambridge, CB2 0SR, UK

**email:* fu@fudan.edu.cn

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

1. Web Appendix A: Calibrating the sensitivity parameter in the WQ-SA method

Although the importance of sensitivity analysis with respect to missing data has been recognized in research and regulatory agencies' guidelines for analysis and reporting of clinical trials (Little et al., 2012; European Medicines Agency, 2010, 2020), calibrating sensitivity parameters in nonignorable missingness models (and in related problems of sensitivity analysis for unmeasured confounding) is still a challenging task and an active area of research (Yin and Shi, 2019; Franks et al., 2020; Sjölander et al., 2022). We first discuss how to determine the sign of the sensitivity parameter using domain knowledge. Then we describe how to calibrate the magnitude of the sensitivity parameter in the WQ-SA methods.

1.1 *Determining the sign of the sensitivity parameter*

In practice, prior knowledge and domain expertise often can be used to reason about some characteristics of the missing data mechanism and determine the sign of sensitivity parameters. In the working model (4) of the main text for the WQ-SA method, the sign of the sensitivity parameter γ_t' has a relatively straightforward interpretation: for fixed $\boldsymbol{h}_t$ and a_t, if $\gamma_t' > 0$, then a smaller value of $y_{pse,t}$ leads to a smaller missingness probability, and vice versa. For example, in the MIMIC-III data, given the covariates and treatments up to stage t (i.e., $\boldsymbol{h}_t$ and a_t), patients with worse disease conditions (i.e., smaller values of $y_{pse,t}$) were more likely to be monitored frequently and less likely to have missing values of $y_{pse,t}$. Therefore, it is reasonable to assume that the sensitivity parameter γ_t' is positive in this scenario.

1.2 *The simulation-based sensitivity analysis approach by Yin and Shi (2019)*

We now turn to the more challenging task of specifying the plausible magnitude of the sensitivity parameters. The research on specifying the sensitivity parameters in nonignorable missingness models is relatively sparse in the literature. Yin and Shi (2019) developed a

simulation-based approach to searching for plausible values of the sensitivity parameters in logistic models for nonignorable missingness probabilities. This simulation-based method was initially introduced by Gelman et al. (1996) for assessing the goodness-of-fit with a Bayesian posterior predictive model. It was further employed for model checking with missing and latent data (Gelman et al., 2005; Copas and Shi, 2000; Copas and Eguchi, 2005; Daniels et al., 2012). Using K-nearest-neighbor (KNN) distance metrics, Yin and Shi (2019) compared the similarity between the *observed* outcome data with *replications* of *observed outcomes* simulated from the estimated full data distribution and the estimated missingness probabilities with fixed values of their sensitivity parameters. If the KNN distance metrics had small values, it indicated that the estimated full data distribution and the missingness probability model (including the specified values of the sensitivity parameters) were *compatible* with the observed outcome data, thus the specified sensitivity parameter values were considered to be plausible. Inspired by their methods, we propose a simulation-based approach to determining the plausible range of our sensitivity parameters in the WQ-SA method.

1.3 *Distinctions between the nonignorable missing outcome settings in Yin and Shi (2019) and the setting with nonignorable missing pseudo-outcomes*

Before we proceed to describe the proposed approach, we emphasize some distinctions between the nonignorable missing outcome settings examined by Yin and Shi (2019) and our DTR setting with nonignorable missing pseudo-outcomes. In typical nonignorable missing outcome settings in Yin and Shi (2019), the observed outcomes are directly available from the raw data. However, due to the potential outcome framework used in the DTR setting, the pseudo-outcomes before the final stage are conditional expectations of potential outcomes, thus not directly available from the raw data. Rather, we use estimated Q-functions to obtain estimates of the pseudo-outcomes in the backward-induction procedure. However, due to the nonignorable missing covariates in future stages, the estimates of pseudo-outcomes at the

current stage are partially missing as well and we need to apply missing data methods to address the selection bias from this missingness. In analogy to the typical non-ignorable outcome settings, we can treat the 'observed' pseudo-outcome *estimates* as the *'raw observed outcomes'* in the DTR setting and compare their distribution with the distribution of the *replications* of '*observed*' pseudo-outcomes in our simulated-based approach. This is reasonable because, as discussed in Section 2.6 of the main text, the pseudo-outcome estimates converge to the pseudo-outcomes when all the models for the Q functions and the missingness probabilities are correctly specified. When describing the implementation steps in the next section, we still employ the pseudo-outcome notation and their models as described in the main text.

1.4 *A simulation-based approach to calibrating the magnitude of the sensitivity parameter*

Let $Y_{pse,t}^{obs}$ and $Y_{pse,t}^{mis}$ denote the observed and unobserved values of $Y_{pse,t}$, respectively, in the sample with complete covariates until stage t (i.e., the patients with $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$). Similarly, let $\widehat{Y}_{pse,t}^{obs}$ denote the observed values of $\widehat{Y}_{pse,t}$ in the sample with $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$. We propose to evaluate the plausibility of each pre-specified value of the sensitivity parameter γ_t' by comparing the distribution of the *observed* pseudo-outcome estimates $\widehat{Y}_{pse,t}^{obs}$ with that of $Y_{pse,t}^{obs,\gamma_t'}$, which are the *replications* of *observed* pseudo-outcomes simulated based on the estimated Q-function at stage t and the estimated working missingness model (4) corresponding to a specified value of γ_t'.

Let $\text{pr}(\cdot)$ and $\text{pr}(\cdot \mid \cdot)$ denote the probability density function and the conditional probability density function, respectively. To simulate data with a specified value of γ_t', we first evaluate the density $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)$, which can be based on the following Lemma.

LEMMA 1: *In the sample with complete covariates until stage t (i.e.,* $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$*), the*

conditional density of the missing part of $Y_{pse,t}$ *(i.e., with* $R_{pse,t} = 0$*) can be expressed by*

$$\frac{\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)}{\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)} = \frac{\exp(\gamma_t' y_{pse,t})}{E\{\exp(\gamma_t' y_{pse,t}) \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\}}. \quad \text{(S1)}$$

In equation (S1), $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)$ and $E\{\exp(\gamma_t' y_{pse,t}) \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\}$ can be estimated by some nonparametric methods such as kernel regressions based on $\widehat{Y}_{pse,t}^{obs}$. Then for any given γ_t', we can generate samples from the estimated $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)$ and impute the missing values in $Y_{pse,t}$ for the patients with complete covariates until stage t (i.e., $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$). We denote these imputed pseudo-outcomes as $Y_{pse,t}^{\gamma_t',imp}$. Thus we have complete data $Y_{pse,t}^{\gamma_t',*} = (\widehat{Y}_{pse,t}^{obs}, Y_{pse,t}^{\gamma_t',imp})$ for the patients with complete covariates until stage t. Under the future-independence missingness assumption (Assumption 4), we have $E(Y_{pse,t} \mid \boldsymbol{h}_t, a_t) = E(Y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$, that is, the Q function model applies to the patients with $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$. Therefore, we fit the assumed Q-function model to $Y_{pse,t}^{\gamma_t',*}$ and estimate $E(Y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$. The density of the residuals of this conditional mean model can also be estimated by nonparametric methods. Combining the estimated conditional mean model and the estimated density of the residuals, we can obtain the conditional distribution $\text{pr}(y_{pse,t}^{\gamma_t',*} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$.

Next, we take a sample $Y_{pse,t}^{\gamma_t',**}$ from the estimated $\text{pr}(y_{pse,t}^{\gamma_t',*} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$ for the patients with complete covariates until stage t (i.e., $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$). The missingness indicator $R_{pse,t}^{\gamma_t',*}$ given $Y_{pse,t}^{\gamma_t',**}$ and γ_t' would be generated with the probability $\widehat{\pi}(\boldsymbol{h}_t, a_t, y_{pse,t}^{\gamma_t',**}; \gamma_t') = \left[1 + \exp\left\{\widehat{s}'_{\gamma_t',t}(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}^{\gamma_t',**}\right\}\right]^{-1}$, where $\widehat{s}'_{\gamma_t',t}(\boldsymbol{h}_t, a_t)$ is estimated by equation (5) in the main text. Drop the values of $Y_{pse,t}^{\gamma_t',**}$ for those with $R_{pse,t}^{\gamma_t',*} = 0$ and denote the remaining part of $Y_{pse,t}^{\gamma_t',**}$ as $Y_{pse,t}^{\gamma_t',obs}$, which we refer to as the replications of observed pseudo-outcomes. Finally, we compare the distribution of $Y_{pse,t}^{\gamma_t',obs}$ with that of $\widehat{Y}_{pse,t}^{obs}$ using the Wilcoxon rank-sum test. A large p-value of the Wilcoxon rank-sum test indicates that the replications of observed pseudo-outcomes are similar to the observed pseudo-outcome estimates $\widehat{Y}_{pse,t}^{obs}$. Therefore, the specified value of γ_t' is considered to be plausible.

For each candidate value of γ_t', we formalize the proposed simulation-based approach in the following steps:

(1) Obtain $\exp\{\widehat{s}'_{\gamma_t',t}(\boldsymbol{h}_t, a_t)\}$ using equation (5) in the main text.

(2) Based on the observed pseudo-outcome estimates $\widehat{Y}^{obs}_{pse,t}$, estimate $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)$ and $E\{\exp(\gamma_t' y_{pse,t}) \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\}$ nonparametrically. Afterwards, we can estimate $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)$ using equation (S1).

(3) Impute $Y^{\gamma_t',imp}_{pse,t}$ for the patients with $(\overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)$ by simulating from the estimated $\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)$. Denote $Y^{\gamma_t',*}_{pse,t} = (\widehat{Y}^{obs}_{pse,t}, Y^{\gamma_t',imp}_{pse,t})$.

(4) Based on $Y^{\gamma_t',*}_{pse,t}$, estimate $\text{pr}(y^{\gamma_t',*}_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$ with the assumed Q-function model and the nonparametric residual density estimator.

(5) Generate $Y^{\gamma_t',**}_{pse,t}$ from the estimated $\text{pr}(y^{\gamma_t',*}_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)$ for the patients with $\overline{\boldsymbol{R}}_t = \mathbf{1}_t$. Generate $R^{\gamma_t',*}_{pse,t}$ from $\widehat{\pi}(\boldsymbol{h}_t, a_t, y^{\gamma_t',**}_{pse,t}; \gamma_t')$. Let $Y^{\gamma_t',obs}_{pse,t}$ denote $Y^{\gamma_t',**}_{pse,t}$ with $R^{\gamma_t',*}_{pse,t} = 1$.

(6) Compare the distribution of $Y^{\gamma_t',obs}_{pse,t}$ and $\widehat{Y}^{obs}_{pse,t}$ using the Wilcoxon rank-sum test.

(7) Repeat Steps (3)-(6) for many Monte Carlo replicates, report the median of the p-values. Let MCR denote the number of Monte Carlo replicates.

Remark S1. In the above procedure, Steps (1)-(5) corresponds to Step (ii) in Section 2.2 of Yin and Shi (2019) but they have been tailored to our proposed WQ-SA method. It involves the specifications for the sensitivity parameter values, the parametric model for the Q-functions, and the exponential tilting form of the working missingness model. Hence, the similarity between $Y^{\gamma_t',obs}_{pse,t}$ and $\widehat{Y}^{obs}_{pse,t}$ can be used to assess both the plausibility of γ_t' and the assumed models in the WQ-SA method.

Remark S2. In Step (6) of the above procedure, we evaluated the similarity between $Y^{\gamma_t',obs}_{pse,t}$ and $\widehat{Y}^{obs}_{pse,t}$ using the Wilcoxon rank-sum test. The Wilcoxon rank-sum test, which is also known as the Mann–Whitney U test, is widely used to determine if a two group sample follows the same distribution. It is a nonparametric test that does not depend upon any

particular distributional form (or parameters). As suggested by Remark 3 in Yin and Shi (2019), various methods can be employed to measure the similarity between $Y_{pse,t}^{\gamma_t',obs}$ and $\widehat{Y}_{pse,t}^{obs}$. The Wilcoxon rank-sum test performed well in Simulation 3 in Web Appendix G and in the MIMIC-III data example, and it can offer an interpretable measure of the similarity between $Y_{pse,t}^{\gamma_t',obs}$ and $\widehat{Y}_{pse,t}^{obs}$. In contrast, the KNN distance was computationally expensive, and results relied on tuning parameters in our numerical studies. Therefore, we chose the Wilcoxon rank-sum test as the similarity measure.

Remark S3. To reduce the impact of random errors on the p-values caused by the generation of $Y_{pse,t}^{\gamma_t',imp}$ and $Y_{pse,t}^{\gamma_t',**}$, we repeated Steps (3)-(6) for MCR times. We reported the median of the resulting p-values since they had a skewed distribution.

Remark S4. In analogy to the conventional significance level 0.05 for hypothesis testing based on p-values, we considered a threshold of the median p-value 0.05 to determine a plausible range in Simulation 3 and MIMIC-III data analysis. Specifically, a value of γ_t' would be included in the plausible range if its corresponding median p-value was greater than 0.05.

Remark S5. The simulation-based approach is for determining a plausible range of the sensitivity parameter, rather than serving as a method for obtaining its point estimate. Following the discussion in Yin and Shi (2019), we recommend using a set of sensitivity parameters within a range determined by a tolerance level of p-value medians, instead of a single value of sensitivity parameter, in the WQ-SA method.

We evaluated the above approach in Simulation 3 in Web Appendix G and applied it to the MIMIC-III data analysis.

2. Web Appendix B: Asymptotic properties

The consistency and asymptotic normality of the estimators in Section 2.6 of the main text can be established under suitable regularity conditions for estimating equations and an

additional assumption that the optimal treatment is unique for all subjects at all stages. Note that the parameters are estimated separately using estimating equations recursively at each stage. Therefore, it is natural to establish the asymptotic properties of the stage-specific parameters recursively as well. For simplicity, we will focus on the two-stage setting, but extensions to the general T-stage setting follow directly.

THEOREM 1: *In the 2-stage scenario, if Assumptions 1-4 hold, the Q-functions at stage 2 are Lipschitz continuous and correctly specified, and* $\boldsymbol{M}_2(\boldsymbol{\theta}_2) = \mathbf{0}$ *has a unique solution,*

$$\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^* \xrightarrow{d} N(\mathbf{0}, \Sigma_{\boldsymbol{\theta}_2}),$$

$$\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) \xrightarrow{p} d_2^{\text{opt}}(\boldsymbol{h}_2) \text{ for all } \boldsymbol{h}_2 \in \mathcal{H}_2,$$

$$\widehat{Y}_{pse,1} \xrightarrow{p} Y_{pse,1},$$

where $\boldsymbol{\theta}_2^*$ *stands or the true values of* $\boldsymbol{\theta}_2$. $\widehat{\boldsymbol{\theta}}_2$, $\widehat{d}_2^{\text{opt}}(\cdot)$, *and* $\widehat{Y}_{pse,1}$ *denote the estimators of* $\boldsymbol{\theta}_2$, $d_2^{\text{opt}}(\cdot)$, *and* $Y_{pse,1}$, *respectively. Notations* $\xrightarrow{d}$ *and* $\xrightarrow{p}$ *stand for converge in distribution and converge in probability, respectively. The detailed form of* $\Sigma_{\boldsymbol{\theta}_2}$ *is given in Web Appendix C.*

THEOREM 2: *In the two-stage scenario, if Assumptions 1-5 hold, all models are Lipschitz continuous and correctly specified, and the estimating equations for the Q-functions and the estimating equation (3) in the main text each have a unique solution, then under certain additional regularity conditions,*

$$\widehat{\boldsymbol{\theta}}_{1,EE} - \boldsymbol{\theta}_1^* \xrightarrow{d} N(\mathbf{0}, \Sigma_{\boldsymbol{\theta}_1}),$$

$$\widehat{d}_{1,EE}^{\text{opt}}(\boldsymbol{h}_1) \xrightarrow{p} d_1^{\text{opt}}(\boldsymbol{h}_1) \text{ for all } \boldsymbol{h}_1 \in \mathcal{H}_1,$$

where $\boldsymbol{\theta}_1^*$ *stands or the true values of* $\boldsymbol{\theta}_1$. $\widehat{\boldsymbol{\theta}}_{1,EE}$ *and* $\widehat{d}_{1,EE}^{\text{opt}}(\cdot)$ *denote the estimators of* $\boldsymbol{\theta}_1$ *and* $d_1^{\text{opt}}(\cdot)$ *with the working model (1) in the main text estimated by the EE method, respectively. The regularity conditions and the detailed form of* $\Sigma_{\boldsymbol{\theta}_1}$ *are given in Web Appendix D.*

THEOREM 3: *In the two-stage scenario, if Assumptions 1-4 and 6 hold, all models are Lipschitz continuous and correctly specified, the sensitivity parameter* γ' *is correctly specified,*

and the estimating equations for the Q-functions each have a unique solution, then under certain additional regularity conditions,

$$\widehat{\boldsymbol{\theta}}_{1,SA} - \boldsymbol{\theta}_1^* \xrightarrow{d} N(\mathbf{0}, \Sigma'_{\boldsymbol{\theta}_1}),$$

$$\widehat{d}^{\text{opt}}_{1,SA}(\boldsymbol{h}_1) \xrightarrow{p} d_1^{\text{opt}}(\boldsymbol{h}_1) \text{ for all } \boldsymbol{h}_1 \in \mathcal{H}_1,$$

where $\boldsymbol{\theta}_1^*$ *stands or the true values of* $\boldsymbol{\theta}_1$. $\widehat{\boldsymbol{\theta}}_{1,SA}$ *and* $\widehat{d}^{\text{opt}}_{1,SA}(\cdot)$ *denote the estimators of* $\boldsymbol{\theta}_1$ *and* $d_1^{\text{opt}}(\cdot)$ *with the working model (4) in the main text estimated by the SA method, respectively. The regularity conditions and the detailed form of* $\Sigma'_{\boldsymbol{\theta}_1}$ *are given in Web Appendix E.*

3. Web Appendix C: Proof of Theorem 1

At stage 2, the parameters $\boldsymbol{\theta}_2$ are estimated by solving $\boldsymbol{M}_2(\boldsymbol{\theta}_2) = \mathbf{0}$, where $\boldsymbol{M}_2(\boldsymbol{\theta}_2) = \widehat{E}\left\{\mathbb{I}(\overline{\boldsymbol{R}}_2 = \mathbf{1}_2)\frac{\partial[\{Q_2(\boldsymbol{h}_2,a_2;\boldsymbol{\theta}_2)-y_{pse,2}\}^2]}{\partial\boldsymbol{\theta}_2}\right\}$. Under the conditions specified in Theorem 1, $\boldsymbol{M}_2(\boldsymbol{\theta}_2)$ is an unbiased estimating equation for $\boldsymbol{\theta}_2$. With the unbiasedness of these estimating equations, we can get the asymptotic property following Newey and McFadden (1994, Thm. 6.1). Using a first order Taylor expansion centered around $\boldsymbol{\theta}_2^*$, we find

$$\boldsymbol{M}_2(\widehat{\boldsymbol{\theta}}_2) = M(\boldsymbol{\theta}_2^*) + E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}(\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*) + o(1),$$

$$\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^* = -E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*) + o(1).$$

For simplicity, we use $\frac{\partial f(\alpha^*,\beta^*)}{\partial\alpha}$ to denote the value of $\frac{\partial f(\alpha,\beta)}{\partial\alpha}$ when $\alpha = \alpha^*, \beta = \beta^*$ for any function $f(\cdot)$ and parameter values (α^*, β^*) in the Web Appendix.

$$E\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)\right] = \mathbf{0},$$

$$Var\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)\right] = E\left(\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)\right]^{\otimes 2}\right).$$

Besides, $\boldsymbol{M}_2(\boldsymbol{\theta}_2)$ is defined as an empirical mean of a function of i.i.d. variable. According to the Central Limit Theorem, $(\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*) \xrightarrow{d} N\left(\mathbf{0}, E\left\{\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)\right]^{\otimes 2}\right\}\right)$.

The detailed form of $\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}$ can be derived by

$$\begin{aligned}
\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2} &= \widehat{E}\left[\mathbb{I}(\overline{\boldsymbol{R}}_2 = \mathbf{1}_2)\frac{\partial^2 \left\{Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*) - y_{pse,2}\right\}^2}{\partial \boldsymbol{\theta}_2^2}\right] \\
&= \widehat{E}\left(2\mathbb{I}(\overline{\boldsymbol{R}}_2 = \mathbf{1}_2)\frac{\partial\left[\left\{Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*) - y_{pse,2}\right\}\frac{\partial Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right]}{\partial \boldsymbol{\theta}_2}\right) \\
&= \widehat{E}\left(2\mathbb{I}(\overline{\boldsymbol{R}}_2 = \mathbf{1}_2)\left[\left\{\frac{\partial Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^2 + \left\{Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*) - y_{pse,2}\right\}\frac{\partial^2 Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2^2}\right]\right), \quad \text{(S2)}
\end{aligned}$$

where $\frac{\partial Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}$ and $\frac{\partial^2 Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2^2}$ depend on the specific form of $Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2)$.

Given that $E(\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*) = 0$, according to the Law of Large Numbers, $\widehat{\boldsymbol{\theta}}_2$ converges to $\boldsymbol{\theta}_2^*$ in probability. That is, for any $\epsilon > 0$, $P(|\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*| > \epsilon) = 0$ as $n \to \infty$. Furthermore, the Lipschitz continuity of $Q_2(\cdot)$ implies the existence of a Lipschitz constant $L_p > 0$ such that for all $(\boldsymbol{\theta}_2', \boldsymbol{\theta}_2)$, $|Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2') - Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2)| \le L_p|\boldsymbol{\theta}_2' - \boldsymbol{\theta}_2|$. Therefore, for all $\boldsymbol{h}_2 \in \mathcal{H}_2$ and $a_2 \in \{-1, 1\}$, we have

$$\begin{aligned}
\lim_{n\to\infty} P\{|Q_2(\boldsymbol{h}_2, a_2; \widehat{\boldsymbol{\theta}}_2) - Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)| \ge \epsilon\} &\le \lim_{n\to\infty} P(L_p|\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*| \ge \epsilon) \\
&= 0
\end{aligned}$$

That is, $Q_2(\boldsymbol{h}_2, a_2; \widehat{\boldsymbol{\theta}}_2)$ converges in probability to $Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2^*)$ for all fixed $(\boldsymbol{h}_2, a_2)$. Since $Q_2(\boldsymbol{h}_2, a_2; \boldsymbol{\theta}_2)$ can be decomposed as $q_{2,0}(\boldsymbol{h}_2; \boldsymbol{\beta}_2) + a_2 q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2)$, $q_{2,0}(\boldsymbol{h}_2; \widehat{\boldsymbol{\beta}}_2)$ and $q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2)$ converge in probability to $q_{2,0}(\boldsymbol{h}_2; \boldsymbol{\beta}_2^*)$ and $q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*)$ for all fixed $\boldsymbol{h}_2$, respectively.

Besides, $\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) = 2\mathbb{I}\left\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) > 0\right\} - 1$. Under the standard regularity assumption that $P\left\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) = 0\right\} = 0$, we have

$$\begin{aligned}
P\left\{\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) \ne d_2^{\text{opt}}(\boldsymbol{h}_2)\right\} &= P\left[\mathbb{I}\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) > 0\} \ne \mathbb{I}\{q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*) > 0\}\right] \\
&= P\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) > 0, q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*) < 0\} + P\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) < 0, q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*) > 0\}
\end{aligned}$$

Because $\lim_{n\to\infty} q_{2,1}(\boldsymbol{h}_t; \widehat{\boldsymbol{\psi}}_t) = q_{2,1}(\boldsymbol{h}_t; \boldsymbol{\psi}_t^*)$, we have $\lim_{n\to\infty} P\left\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) > 0, q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*) < 0\right\} = 0$ and $\lim_{n\to\infty} P\left\{q_{2,1}(\boldsymbol{h}_2; \widehat{\boldsymbol{\psi}}_2) < 0, q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2^*) > 0\right\} = 0$. Therefore, for all $\boldsymbol{h}_2 \in \mathcal{H}_2$,

$$\lim_{n\to\infty} P\left\{\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) \ne d_2^{\text{opt}}(\boldsymbol{h}_2)\right\} = 0,$$

which means that $\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) \xrightarrow{p} d_2^{\text{opt}}(\boldsymbol{h}_2)$ for any fixed $\boldsymbol{h}_2$.

Since $Y_{pse,1} = \max_{a_2\in\{-1,1\}} Q_2\left(\boldsymbol{H}_2, a_2; \boldsymbol{\theta}_2^*\right)$, we have $y_{pse,1} = Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \boldsymbol{\theta}_2^*\right\}$ and $\widehat{y}_{pse,1} = Q_2\left\{\boldsymbol{h}_2, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2); \widehat{\boldsymbol{\theta}}_2\right\}$ for a given $\boldsymbol{h}_2$. Based on the convergence of $\widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2)$ and $Q_2(\boldsymbol{h}_2, a_2; \widehat{\boldsymbol{\theta}}_2)$, we have

$$
\begin{aligned}
\lim_{n\to\infty} P(|\widehat{y}_{pse,1} - y_{pse,1}| \geq \epsilon) &= \lim_{n\to\infty} P\left[\left|Q_2\left\{\boldsymbol{h}_2, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2); \widehat{\boldsymbol{\theta}}_2\right\} - Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \boldsymbol{\theta}_2^*\right\}\right| \geq \epsilon\right] \\
&= \lim_{n\to\infty} P\left[\left|Q_2\left\{\boldsymbol{h}_2, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2); \widehat{\boldsymbol{\theta}}_2\right\} - Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \boldsymbol{\theta}_2^*\right\}\right| \geq \epsilon, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) \neq d_2^{\text{opt}}(\boldsymbol{h}_2)\right] \\
&\quad + \lim_{n\to\infty} P\left[\left|Q_2\left\{\boldsymbol{h}_2, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2); \widehat{\boldsymbol{\theta}}_2\right\} - Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \boldsymbol{\theta}_2^*\right\}\right| \geq \epsilon, \widehat{d}_2^{\text{opt}}(\boldsymbol{h}_2) = d_2^{\text{opt}}(\boldsymbol{h}_2)\right] \\
&= \lim_{n\to\infty} P\left[\left|Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \widehat{\boldsymbol{\theta}}_2\right\} - Q_2\left\{\boldsymbol{h}_2, d_2^{\text{opt}}(\boldsymbol{h}_2); \boldsymbol{\theta}_2^*\right\}\right| \geq \epsilon\right] \\
&= 0.
\end{aligned}
$$

Therefore, $\widehat{y}_{pse,1} \xrightarrow{p} y_{pse,1}$ for any fixed $\boldsymbol{h}_2$. Subsequently, $\widehat{Y}_{pse,1} \xrightarrow{p} Y_{pse,1}$.

4. Web Appendix D: Proof of Theorem 2

At stage 1, the parameters $\boldsymbol{\theta}_1$ are estimated by solving the estimating equation $\boldsymbol{M}_1(\boldsymbol{\theta}_1) = \boldsymbol{0}$, where $\boldsymbol{M}_1(\boldsymbol{\theta}_1) = \widehat{E}\left(\frac{\widehat{r}_{pse,1}}{\widehat{\pi}_1}\mathbb{I}(R_1 = 1)\frac{\partial\left[\{Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1) - \widehat{y}_{pse,1}\}^2\right]}{\partial\boldsymbol{\theta}_1}\right)$. Following Wallace and Moodie (2015) and Simoneau et al. (2020), we establish the asymptotic normality of $\widehat{\boldsymbol{\theta}}_1$ under the regularity condition that $P(\{q_{2,1}(\boldsymbol{H}_2; \widehat{\boldsymbol{\psi}}_2) = 0\}) = 0$. Since $(\widehat{Y}_{pse,t}, \widehat{R}_{pse,t}, \widehat{\pi}_t)$ are calculated with $(\widehat{\boldsymbol{\theta}}_2, \widehat{\gamma}_1)$, $\boldsymbol{M}_1$ depends $(\widehat{\boldsymbol{\theta}}_2, \widehat{\gamma}_1)$ and thus can also be denoted as $\boldsymbol{M}_1(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2, \widehat{\gamma}_1)$. The variance of $\boldsymbol{\theta}_1$ must adjust for the plug-in estimates of nuisance parameters $(\boldsymbol{\theta}_2, \gamma_1)$. By performing a first-order Taylor expression of $\boldsymbol{M}_1(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2, \widehat{\gamma}_1)$ around $(\boldsymbol{\theta}_2^*, \gamma_1^*)$, we have

$$
M_{\text{adj},1}(\boldsymbol{\theta}_1) = \boldsymbol{M}_1(\boldsymbol{\theta}_1, \widehat{\boldsymbol{\theta}}_2, \widehat{\gamma}_1) + E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*, \gamma_1^*)}{\partial(\boldsymbol{\theta}_2, \gamma_1)}\right\}\left\{\begin{pmatrix} \widehat{\boldsymbol{\theta}}_2 \\ \widehat{\gamma}_1 \end{pmatrix} - \begin{pmatrix} \boldsymbol{\theta}_2^* \\ \gamma_1^* \end{pmatrix}\right\} + o(1),
$$

where $E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*, \gamma_1^*)}{\partial(\boldsymbol{\theta}_2, \gamma_1)}\right\} = E\left\{\left(\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*, \gamma_1^*)}{\partial\boldsymbol{\theta}_2}, \frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*, \gamma_1^*)}{\partial\gamma_1}\right)\right\}$, and we assume that $\boldsymbol{M}_1$ is infinitely differentiable for $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \gamma_1)$. Therefore,

$$
\begin{aligned}
\boldsymbol{M}_{adj,1}(\widehat{\boldsymbol{\theta}}_1) &= \boldsymbol{M}_{adj,1}(\boldsymbol{\theta}_1^*) + E\left\{\frac{\partial \boldsymbol{M}_{adj,1}(\boldsymbol{\theta}_1^*)}{\partial\boldsymbol{\theta}_1}\right\}(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^*) + o(1), \\
\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^* &= -E\left\{\frac{\partial \boldsymbol{M}_{adj,1}(\boldsymbol{\theta}_1^*)}{\partial\boldsymbol{\theta}_1}\right\}^{-1} \boldsymbol{M}_{adj,1}(\boldsymbol{\theta}_1^*) + o(1).
\end{aligned}
$$

We need the following conditions to make sure that $\lim_{n\to\infty} \widehat{\gamma}_1 - \gamma_1^* = 0$ and subsequently $\lim_{n\to\infty} \boldsymbol{M}_{adj,1}(\boldsymbol{\theta}_1^*) = \mathbf{0}$.

Condition 1. The kernel $K(\boldsymbol{u}_1)$ has bounded derivatives of order δ, satisfies $\int K(\boldsymbol{u}_1)\mathrm{d}\boldsymbol{u}_1 = 1$, and has zero moments of order up to $m-1$ and nonzero m th-order moment.

Condition 2. The true function of $s_1(\boldsymbol{u}_1)$ is continuously differentiable and bounded on an open set containing the support of $\boldsymbol{u}_1$.

Condition 3. The moment $E\{\exp(4\gamma_1\widehat{y}_{pse,1})\}$ is finite and the function $E\{\exp(4\gamma_1\widehat{y}_{pse,1}) \mid \boldsymbol{u}_1\}\,\mathrm{pr}(\boldsymbol{u}_1)$ is bounded, where $\mathrm{pr}(\boldsymbol{u}_1)$ is the marginal density of $\boldsymbol{u}_1$.

Condition 4. The bandwidth $c_1 = c_{n_1}$ is such that $c_{n_1} \to 0, n_1 c_{n_1}^{p_1} \to \infty, n_1^{1/2} c_{n_1}^{p_1+2\delta} / \log n_1 \to \infty$, and $n_1 c_{n_1}^{2m} \to 0$ as $n_1 \to \infty$, where p_1 is the dimension of $\boldsymbol{u}_1$, n_1 is the number of units with fully-observed covariates at stage 1.

Let $\boldsymbol{B}_1'(\boldsymbol{\theta}_2, \gamma_1) = \frac{\partial \boldsymbol{B}_1(\boldsymbol{\theta}_2,\gamma_1)^{\mathrm{T}} W_1 \boldsymbol{B}_1(\boldsymbol{\theta}_2,\gamma_1)}{\partial \gamma_1}$, with conditions 1-4 and the convergence of $\boldsymbol{\theta}_2$, we have $\lim_{n\to\infty} E\{\boldsymbol{B}_1'(\widehat{\boldsymbol{\theta}}_2, \gamma_1^*)\} \xrightarrow{p} 0$ and $\lim_{n\to\infty} \boldsymbol{W}_1 \xrightarrow{p} \boldsymbol{W}_1^*$.

By Taylor expansion, we have:

$$
\begin{aligned}
&\begin{pmatrix} \widehat{\boldsymbol{\theta}}_2 \\ \widehat{\gamma}_1 \end{pmatrix} - \begin{pmatrix} \boldsymbol{\theta}_2^* \\ \gamma_1^* \end{pmatrix} = -E\left[\frac{\partial\{\boldsymbol{M}_2(\boldsymbol{\theta}_2^*), \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*, \gamma_1^*)\}}{\partial(\boldsymbol{\theta}_2, \gamma_1)}\right]^{-1} \begin{pmatrix} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\ \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*, \gamma_1^*) \end{pmatrix} + o(1) \\
&= -\left(E\begin{bmatrix} \frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2} & \frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \gamma_1} \\ \frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2} & \frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1} \end{bmatrix}\right)^{-1} \begin{pmatrix} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\ \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*, \gamma_1^*) \end{pmatrix} + o(1) \\
&= -\left(\begin{bmatrix} E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\} & \mathbf{0} \\ E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\} & E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\} \end{bmatrix}\right)^{-1} \begin{pmatrix} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\ \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*, \gamma_1^*) \end{pmatrix} + o(1) \\
&= \begin{bmatrix} -\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\right]^{-1} & \mathbf{0} \\ \left[E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\right]^{-1} E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\right]^{-1} & -\left[E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\right]^{-1} \end{bmatrix} \begin{pmatrix} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\ \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*, \gamma_1^*) \end{pmatrix} + o(1).
\end{aligned}
$$

Therefore, $\widehat{\gamma}_1 - \gamma_1^* = E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\right]^{-1}\boldsymbol{M}_2(\boldsymbol{\theta}_2^*) - \left[E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\right]^{-1}\boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)$,

$$
\begin{aligned}
M_{\text{adj},1}(\boldsymbol{\theta}_1) &= \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*) - E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\right]^{-1}\boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\
&+ E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\left[E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\right]^{-1}E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}\left[E\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\right]^{-1}\boldsymbol{M}_2(\boldsymbol{\theta}_2^*) \\
&- E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\left[E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}\right]^{-1}\boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*) + o(1).
\end{aligned}
$$

And we have $\widehat{\boldsymbol{\theta}}_{1,EE} - \boldsymbol{\theta}_1^* \to N(\mathbf{0},\Sigma_{\boldsymbol{\theta}_1})$, where $\Sigma_{\boldsymbol{\theta}_1} = E\left\{\left(\left[E\left\{\frac{\partial M_{adj,1}(\boldsymbol{\theta}_1^*)}{\partial \boldsymbol{\theta}_1}\right\}\right]^{-1}M_{adj,1}(\boldsymbol{\theta}_1^*)\right)^{\otimes 2}\right\}$.

The detailed form of $\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}$ is provided in equation (S2) in Web Appendix A. Assume that $\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}$ and $\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}$ exist. To obtain $M_{\text{adj},1}(\boldsymbol{\theta}_1)$ from observed data, we give the detailed form of $E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}$, $E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}$, $E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2}\right\}$, and $E\left\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}\right\}$ in the following text. Since the missing indicator of $Y_{pse,1}$ is mainly determined by the parametric form Q_2, under the assumption that the variables employed in the computation of $Y_{pse,1}$ remain unchanged, we have $\widehat{r}_{pse,1} = r_{pse,1}$.

Since $\widehat{y}_{pse,1}$ is calculated with $\boldsymbol{\theta}_2$, we denote it as $\widehat{y}_{pse,1}(\widehat{\boldsymbol{\theta}}_2)$. Note that $\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*) = y_{pse,1}$. Consequently,

$$
\begin{aligned}
&\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \boldsymbol{\theta}_2} = \frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*),\gamma_1^*)}{\partial\{\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*)\}}\frac{\partial\{\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*)\}}{\partial \boldsymbol{\theta}_2} \\
&= \frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,y_{pse,1},\gamma_1^*)}{\partial y_{pse,1}}\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2} \\
&= \frac{\partial \widehat{E}\left(\frac{r_{pse,1}}{\widehat{\pi}_1}r_1\frac{\partial[\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\}^2]}{\partial \boldsymbol{\theta}_1}\right)}{\partial y_{pse,1}}\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2} \\
&= \widehat{E}\left(2r_{pse,1}r_1\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\frac{\partial[\widehat{\pi}_1\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\}]}{\partial y_{pse,1}}\right)\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2} \\
&= \widehat{E}\left(2r_{pse,1}r_1\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\frac{\partial[1/\widehat{\pi}_1\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\}]}{\partial y_{pse,1}}\right)\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2} \\
&= \widehat{E}\left(2r_{pse,1}r_1\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\left[-\frac{1}{\widehat{\pi}_1}+\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\}\frac{\partial(1/\widehat{\pi}_1)}{\partial y_{pse,1}}\right]\right)\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2} \\
&= \widehat{E}\left(2r_{pse,1}r_1\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\left[\{\gamma_1^*Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-\gamma_1^*y_{pse,1}-1\}\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)+\gamma_1^*y_{pse,1}\}-1\right]\right)\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2},
\end{aligned}
$$

where $\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}$ and $\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2}$ depends on the specified form of the Q-functions.

Similarly, we obtain $\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1}$ using the following equations

$$\begin{aligned}
\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial \gamma_1} &= \frac{\partial \boldsymbol{M}_1\{\boldsymbol{\theta}_1,\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*),\gamma_1^*\}}{\partial\left[1/\widehat{\pi}_1\{\boldsymbol{u}_1,\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*),\gamma_1^*\}\right]}\frac{\partial\left[1/\widehat{\pi}_1\{\boldsymbol{u}_1,\widehat{y}_{pse,1}(\boldsymbol{\theta}_2^*),\gamma_1^*\}\right]}{\partial\gamma_1}\\
&= \frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,y_{pse,1},\gamma_1^*)}{\partial\left\{1/\widehat{\pi}_1(\boldsymbol{u}_1,y_{pse,1},\gamma_1^*)\right\}}\frac{\partial\left\{1/\widehat{\pi}_1(\boldsymbol{u}_1,y_{pse,1},\gamma_1^*)\right\}}{\partial\gamma_1}\\
&= \frac{\partial \widehat{E}\left(\frac{r_{pse,1}}{\widehat{\pi}_1}r_1\frac{\partial\left[\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\}^2\right]}{\partial\boldsymbol{\theta}_1}\right)}{\partial(1/\widehat{\pi}_1)}\frac{\partial\left\{1/\widehat{\pi}_1(\boldsymbol{u}_1,y_{pse,1},\gamma_1^*)\right\}}{\partial\gamma_1}\\
&= \widehat{E}\left[2r_{pse,1}r_1\left\{Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)-y_{pse,1}\right\}\frac{\partial Q_1(\boldsymbol{h}_1,a_1;\boldsymbol{\theta}_1)}{\partial\boldsymbol{\theta}_1}\right]\left[\frac{\partial\left\{1/\widehat{\pi}_1(\boldsymbol{u}_1,y_{pse,1},\gamma_1^*)\right\}}{\partial\gamma_1}\right].
\end{aligned} \tag{S3}$$

The explicit expression for $\frac{\partial(1/\widehat{\pi}_1)}{\partial\gamma_1}$ can be derived with the following equations

$$\begin{aligned}
\frac{\partial\{1/\widehat{\pi}_1(\boldsymbol{u}_1,y_{pse,1},\gamma_1^*)\}}{\partial\gamma_1} &= \frac{\partial\left[1+\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)\}\exp(\gamma_1^*y_{pse,1})\right]}{\partial\gamma_1}\\
&= \gamma_1^*\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)\}\exp(\gamma_1^*y_{pse,1})+\frac{\partial\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)\}}{\partial\gamma_1}\exp(\gamma_1^*y_{pse,1})\\
&= \gamma_1^*\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)\}\exp(\gamma_1^*y_{pse,1})+\frac{\partial\left\{\frac{\sum_{i=1}^n(1-r_{pse,1,i})K_{c_1}(\boldsymbol{u}_1-\boldsymbol{u}_{1,i})r_{1,i}}{\sum_{i=1}^n r_{pse,1,i}\exp(\gamma_1^*y_{pse,1,i})K_{c_1}(\boldsymbol{u}_1-\boldsymbol{u}_{1,i})r_{1,i}}\right\}}{\partial\gamma_1}\exp(\gamma_1^*y_{pse,1})\\
&= \exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{u}_1)\}\exp(\gamma_1^*y_{pse,1})\left\{\gamma_1^*+\frac{\sum_{i=1}^n r_{pse,1,i}y_{pse,1,i}\exp\left(\gamma_1^*y_{pse,1,i}\right)K_{c_1}\left(\boldsymbol{u}_1-\boldsymbol{u}_{1,i}\right)r_{1,i}}{\sum_{i=1}^n r_{pse,1,i}\exp\left(\gamma_1^*y_{pse,1,i}\right)K_{c_1}\left(\boldsymbol{u}_1-\boldsymbol{u}_{1,i}\right)r_{1,i}}\right\}.
\end{aligned} \tag{S4}$$

Combining equations (S3) and (S4), we can obtain the detailed form of $\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial\gamma_1}$.

The detailed form of $E\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial\gamma_1}\}$ and $E\{\frac{\partial \boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial\boldsymbol{\theta}_2}\}$ can be derived following Shao and Wang (2016) when the following regularity conditions hold.

Condition 5. There is a vector of the functional $G(r_{pse,1},\boldsymbol{u}_1,\boldsymbol{\omega})$ which is linear in $\boldsymbol{\omega} = (\omega_1,\omega_2)^{\mathrm{T}}$ and such that:

(i) for small enough $\|\boldsymbol{\omega}-\boldsymbol{\omega}^*\|$, $\|\widetilde{m}\left(r_{pse,1},\boldsymbol{u}_1,\widehat{y}_{pse,1},\boldsymbol{\omega},\gamma_1\right)-\widetilde{m}\left(r_{pse,1},\boldsymbol{u}_1,\widehat{y}_{pse,1},\boldsymbol{\omega}^*,\gamma_1\right)-G(r_{pse,1},\boldsymbol{u}_1,\boldsymbol{\omega}-\boldsymbol{\omega}^*)\| \leqslant c(r_{pse,1},\boldsymbol{u}_1)\left(\|\boldsymbol{\omega}-\boldsymbol{\omega}^*\|\right)^2$, where $\widetilde{m}(r_{pse,1},\boldsymbol{u}_1,\widehat{y}_{pse,1},\boldsymbol{\omega},\gamma_1)$ is the L-dimensional vector. And its l-th component $\widetilde{m}_l(r_{pse,1},\boldsymbol{u}_1,\widehat{y}_{pse,1},\boldsymbol{\omega},\gamma_1)$, is defined as $\mathbb{I}(\boldsymbol{z}_1=l)[r_{pse,1}\{1+\exp(\gamma_1\widehat{y}_{pse,1})\omega_1(\boldsymbol{u}_1)/\omega_2(\boldsymbol{u}_1)\}-1]$, where $\boldsymbol{\omega}^* = [E(1-r_{pse,1}\mid\boldsymbol{u}_1),E(r_{pse,1}\exp(\gamma_1^*y_{pse,1})\mid\boldsymbol{u}_1)]^{\mathrm{T}}$, and $E\{c(r_{pse,1},\boldsymbol{u}_1)\}<\infty$;

(ii) $\|G(r_{pse,1},\boldsymbol{u}_1,\boldsymbol{\omega})\| \leqslant e(r_{pse,1},\boldsymbol{u}_1)\|\boldsymbol{\omega}\|$ and $E\{e(r_{pse,1},\boldsymbol{u}_1)^2\}<\infty$;

(iii) there exists an almost everywhere continuous function $v_1(\boldsymbol{u}_1)$ with $\int \|v_1(\boldsymbol{u}_1)\|\mathrm{d}\boldsymbol{u}_1 < \infty$, $E\{G(r_{pse,1}, \boldsymbol{u}_1, \boldsymbol{\omega})\} = \int v_1(\boldsymbol{u}_1)\boldsymbol{\omega}(\boldsymbol{u}_1)\mathrm{d}\boldsymbol{u}_1$ for all $\|\boldsymbol{\omega}\| \leqslant \infty$, and $E\{\sup_{\|\zeta\|\leqslant\epsilon}\|v_1(\boldsymbol{u}_1+\zeta)\|^4\} < \infty$ for some $\epsilon > 0$.

Condition 6. For small enough $\|\boldsymbol{\omega}-\boldsymbol{\omega}_0\|, \widetilde{m}(r_{pse,1}, \boldsymbol{u}_1, \widehat{y}_{pse,1}, \boldsymbol{\omega}, \gamma_1)$ is continuously differentiable in γ_1 in a neighbourhood of $(y_{pse,1}, \gamma_1^*)$. And there is a function $k(r_{pse,1}, \boldsymbol{u}_1)$ satisfying $\left\|\frac{\partial\widetilde{m}(r_{pse,1},\boldsymbol{u}_1,\widehat{y}_{pse,1},\boldsymbol{\omega},\gamma_1)}{\partial\gamma_1} - \frac{\partial\widetilde{m}(r_{pse,1},\boldsymbol{u}_1,y_{pse,1},\boldsymbol{\omega}^*,\gamma_1^*)}{\partial\gamma_1}\right\| \leq k(r_{pse,1},\boldsymbol{u}_1)(\|\gamma_1^*-\gamma_1\|^\epsilon + \|\boldsymbol{\omega}-\boldsymbol{\omega}^*\|^\epsilon + \|\widehat{y}_{pse,1}-y_{pse,1}\|^\epsilon)$ for an $\epsilon > 0$, $E\{k(r_{pse,1},\boldsymbol{u}_1)\} < \infty$. Besides, $\boldsymbol{\Gamma}_{\gamma_1} = E\left\{\frac{\partial\widetilde{m}(r_{pse,1},\boldsymbol{u}_1,y_{pse,1},\boldsymbol{\omega}^*,\gamma_1^*)}{\partial\gamma_1}\right\}$ and $\boldsymbol{\Gamma}_{y_{pse,1}} = E\left\{\frac{\partial\widetilde{m}(r_{pse,1},\boldsymbol{u}_1,y_{pse,1},\boldsymbol{\omega}^*,\gamma_1^*)}{\partial\widehat{y}_{pse,1}}\right\}$ exist and are of full rank.

Assume that Conditions 1 – 6 hold. Then, as $n \to \infty$, $E\left\{\frac{\partial\boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial\gamma_1}\right\} \to 2\boldsymbol{\Gamma}_{\gamma_1}^{\mathrm{T}}\boldsymbol{W}_1^*\boldsymbol{\Gamma}_{\gamma_1}$, $E\left\{\frac{\partial\boldsymbol{B}_1'(\boldsymbol{\theta}_2^*,\gamma_1^*)}{\partial\boldsymbol{\theta}_2}\right\} \to 2\boldsymbol{\Gamma}_{y_{pse,1}}^{\mathrm{T}}\boldsymbol{W}_1^*\boldsymbol{\Gamma}_{y_{pse,1}}\frac{\partial y_{pse,1}}{\partial\boldsymbol{\theta}_2}$.

As mentioned by Shao and Wang (2016), the form of the asymptotic variance of this semiparameteric estimator is complicated. Consequently, we suggest using the bootstrap method to estimate it.

Besides, according to the Law of Large Numbers, $\widehat{\boldsymbol{\theta}}_{1,EE}$ converges to $\boldsymbol{\theta}_1$ in probability. Therefore, following the proof of Theorem 1, we can similarly prove the convergence in probability of $\widehat{d}_1^{\mathrm{opt}}$ by proving that $Q_1(\boldsymbol{h}_1, a_1; \widehat{\boldsymbol{\theta}}_{1,EE}) \xrightarrow{p} Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1)$ for any fixed $(\boldsymbol{h}_1, a_1)$.

5. Web Appendix E: Proof of Theorem 3

We need the following conditions to guarantee that $\lim_{n\to\infty}\exp\{\widehat{s}'_{\gamma_1',1}(\boldsymbol{g}_1)\} = \exp\{s'_{\gamma_1',1}(\boldsymbol{g}_1)\}$:

Condition 7. The kernel $K'(\boldsymbol{g}_1)$ has bounded derivatives of order δ', satisfies $\int K'(\boldsymbol{g}_1)\mathrm{d}\boldsymbol{g}_1 = 1$, and has zero moments of order up to $m-1$ and nonzero m th-order moment.

Condition 8. The true function of $s_1'(\boldsymbol{g}_1)$ is continuously differentiable and bounded on an open set containing the support of $\boldsymbol{g}_1$.

Condition 9. The moment $E[\exp(4\gamma_1\widehat{y}_{pse,1})]$ is finite and the function $E[\exp(4\gamma_1\widehat{y}_{pse,1}) \mid \boldsymbol{g}_1]\,\mathrm{pr}(\boldsymbol{g}_1)$ is bounded, where $\mathrm{pr}(\boldsymbol{g}_1)$ is the marginal density of $\boldsymbol{g}_1$.

Condition 10. The bandwidth $c_1' = c_{n_1}'$ is such that $c_{n_1}' \to 0, n_1{c_{n_1}'}^{p_1'} \to \infty, n_1^{1/2}{c_{n_1}'}^{p_1'+2\delta'}/\log n_1 \to$

∞, and $n_1 {c'_{n_1}}^{2m} \to 0$ as $n_1 \to \infty$, where p'_1 is the dimension of $\boldsymbol{g}_1$, n_1 is the number of units with fully-observed covariates at stage 1.

Under these conditions, $\exp\{\widehat{s}'_{\gamma'_1,1}(\boldsymbol{h}_1)\}$ is a consistent estimator of $\exp\{s'_{\gamma'_1,1}(\boldsymbol{h}_1)\}$. Therefore, the estimated missingness probability model is consistent. And $\boldsymbol{M}_1(\boldsymbol{\theta}_1)$ is an unbiased estimation equation for $\boldsymbol{\theta}_1$ (Newey and McFadden, 1994). Similar to Web Appendix D, where we provided the asymptotic properties of $\widehat{\boldsymbol{\theta}}_1$ for the WQ-EE method, we establish the asymptotic normality of $\widehat{\boldsymbol{\theta}}_1$ for the WQ-SA method under the regularity condition that $P(\{q_{2,1}(\boldsymbol{H}_2; \widehat{\boldsymbol{\psi}}_2) = 0\}) = 0$.

Let $M'_{\text{adj},1}(\boldsymbol{\theta}_1) = \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*) + E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}(\widehat{\boldsymbol{\theta}}_2 - \boldsymbol{\theta}_2^*) + o(1)$. Then we have

$$M'_{\text{adj},1}(\boldsymbol{\theta}_1) = \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*) - E\left\{\frac{\partial \boldsymbol{M}_1(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}\left\{\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}\right\}^{-1} \boldsymbol{M}_2(\boldsymbol{\theta}_2^*).$$

And $\widehat{\boldsymbol{\theta}}_{1,SA} - \boldsymbol{\theta}_1 \to N(\boldsymbol{0}, \Sigma'_{\boldsymbol{\theta}_1})$, where $\Sigma'_{\boldsymbol{\theta}_1} = E\left\{\left(\left[E\left\{\frac{\partial M'_{adj,1}(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\right\}\right]^{-1} M'_{adj,1}(\boldsymbol{\theta}_1)\right)^{\otimes 2}\right\}$. The explicit mathematical expression for $\frac{\partial \boldsymbol{M}_2(\boldsymbol{\theta}_2^*)}{\partial \boldsymbol{\theta}_2}$ can be found in equation (S2) in Web Appendix C. Employing a similar approach to that outlined in Web Appendix D, we obtain $\frac{\partial M'_{adj,1}(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1} = \widehat{E}\left(2r_{pse,1}r_1 \frac{\partial Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}\left[\{\gamma_1 Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1) - \gamma_1 y_{pse,1} - 1\}\exp\{\widehat{s}_{\gamma_1^*}(\boldsymbol{h}_1) + \gamma_1 y_{pse,1}\} - 1\right]\right)\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2}$, where $\frac{\partial Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}$ and $\frac{\partial y_{pse,1}}{\partial \boldsymbol{\theta}_2}$ depend on the specified form of the Q-functions. Because of the complexity of the asymptotic variance, we recommend utilizing the bootstrap method for its estimation. Following the derivation of Theorem 1, we can prove the convergence in probability of $\widehat{d}_1^{\text{opt}}$ by proving that $Q_1(\boldsymbol{h}_1, a_1; \widehat{\boldsymbol{\theta}}_{1,SA}) \xrightarrow{p} Q_1(\boldsymbol{h}_1, a_1; \boldsymbol{\theta}_1)$ for any fixed $(\boldsymbol{h}_1, a_1)$.

6. Web Appendix F: Proof of Lemma 1

Using Bayes' Theorem, we have

$$\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 0) = \frac{\text{pr}(y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 0)}{\text{pr}(\boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 0)},$$

$$\text{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 1) = \frac{\text{pr}(y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 1)}{\text{pr}(\boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \boldsymbol{1}_t, R_{pse,t} = 1)}.$$

This leads to the following equations:

$$\begin{aligned}
&\frac{\mathrm{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)}{\mathrm{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)} \\
&\quad = \frac{\mathrm{pr}(y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)}{\mathrm{pr}(\boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)} \frac{\mathrm{pr}(\boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)}{\mathrm{pr}(y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)} \\
&\quad = \frac{P(R_{pse,t} = 0 \mid y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)}{P(R_{pse,t} = 1 \mid y_{pse,t}, \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)} \frac{P(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)}{P(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)}.
\end{aligned}$$

Under Assumption 6, we have $\frac{P(R_{pse,t}=0|y_{pse,t},\boldsymbol{h}_t,a_t,\overline{\boldsymbol{R}}_t=\mathbf{1}_t)}{P(R_{pse,t}=1|y_{pse,t},\boldsymbol{h}_t,a_t,\overline{\boldsymbol{R}}_t=\mathbf{1}_t)} = \exp\{s_t'(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}\}$. Besides,

$$\begin{aligned}
&\frac{P(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)}{P(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)} \\
&\quad = \int \mathrm{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0) \frac{P(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)}{P(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t)} \mathrm{d}y_{pse,t} \\
&\quad = \int \frac{\mathrm{pr}\left(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t\right)}{P\left(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t\right)} P\left(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, y_{pse,t}\right) \mathrm{d}y_{pse,t} \\
&\quad = \int \frac{\mathrm{pr}\left(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\right)}{P\left(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, y_{pse,t}\right)} P\left(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, y_{pse,t}\right) \mathrm{d}y_{pse,t} \\
&\quad = \int \frac{P\left(R_{pse,t} = 0 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, y_{pse,t}\right)}{P\left(R_{pse,t} = 1 \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, y_{pse,t}\right)} \mathrm{pr}\left(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\right) \mathrm{d}y_{pse,t} \\
&\quad = \int \exp\{s_t'(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}\} \mathrm{pr}\left(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\right) \mathrm{d}y_{pse,t} \\
&\quad = E[\exp\{s_t'(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}\} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\mathrm{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 0)}{\mathrm{pr}(y_{pse,t} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1)} &= \frac{\exp\{s_t'(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}\}}{E[\exp\{s_t'(\boldsymbol{h}_t, a_t) + \gamma_t' y_{pse,t}\} \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1]} \\
&= \frac{\exp(\gamma_t' y_{pse,t})}{E\{\exp(\gamma_t' y_{pse,t}) \mid \boldsymbol{h}_t, a_t, \overline{\boldsymbol{R}}_t = \mathbf{1}_t, R_{pse,t} = 1\}}.
\end{aligned}$$

This proves the equation (S1) in Lemma 1.

7. Web Appendix G: The m-out-of-n bootstrap method for parameter inference in non-regularity scenarios

The asymptotic properties for stage-1 parameters in Theorem 2 and Theorem 3 in Web Appendix B are based on regularity conditions (i.e., $p_{n-regu} = P\{q_{2,1}(\boldsymbol{h}_2; \boldsymbol{\psi}_2) = 0\} = 0$).

However, the inference for the parameters in Q-learning for optimal DTRs may face the challenge of non-regularity (Chakraborty et al., 2010), which can undermine the validity of the asymptotic distribution provided in Web Appendix B and the ordinary n-out-of-n bootstrap method (Andrews, 2000). Specifically, in the two-stage scenario, the stage-1 pseudo-outcome and estimators are non-smooth functions of $\widehat{\boldsymbol{\psi}}_2$ when $P\{q_{2,1}(\boldsymbol{h}_2;\widehat{\boldsymbol{\psi}}_2)=0\}>0$. As a result, the asymptotic distribution of $(\widehat{\boldsymbol{\theta}}_1-\boldsymbol{\theta}_1)$ would be non-normal. Small perturbations in the marginal distribution of $\boldsymbol{H}_2$, the parameter $\boldsymbol{\psi}_2$, or both, can lead to abrupt differences in the limiting distribution of $\widehat{\boldsymbol{\theta}}_1$, and thus make them non-regular. In this case, we followed Chakraborty et al. (2013) and employed the m-out-of-n bootstrap method, which is consistent in non-smooth scenarios, to construct confidence intervals.

The m-out-of-n bootstrap uses a resample size m smaller than the original sample size n to asymptotically avoid the non-negligible probability concentrated at the discontinuous points of $\widehat{Y}_{pse,1}$ and $\widehat{\boldsymbol{\theta}}_1$ as functions of $\widehat{\boldsymbol{\psi}}_2$ (Xu et al., 2015). Specifically, we would draw B bootstrap samples, each with a sample size m, and obtain B parameter estimates. The 95% confidence interval is then constructed using the 2.5th and 97.5th percentiles of the B parameter estimates.

For the choice of resample size m, Chakraborty et al. (2013) proposed a data adaptive method that is directly connected with a measure of non-regularity. Specifically, they estimated p_{n-regu} by $\widehat{p}_{n-regu}=\widehat{E}(\mathbb{I}[n\{q_{2,1}(\boldsymbol{h}_2;\widehat{\boldsymbol{\psi}}_2)\}^2\le \boldsymbol{h}_2\widehat{\Sigma}_{21}\boldsymbol{h}_2^{\mathrm{T}}\chi^2_{1,1-\nu}])$, where $\widehat{\Sigma}_{21}$ was the plug-in estimator of $n\mathrm{Cov}(\widehat{\boldsymbol{\psi}}_2,\widehat{\boldsymbol{\psi}}_2)$, $\chi^2_{1,1-\nu}$ was the $(1-\nu)\times 100$ percentile of the χ^2 distribution with 1 degree of freedom, and ν was set as 0.001 in their simulation. We used the same value of ν in our simulations and MIMIC-III data analysis. The resample size m was set as $m=\left\lceil n^{\frac{1+\alpha(1-\widehat{p}_{n-regu})}{1+\alpha}}\right\rceil$, where $\lceil x\rceil$ denotes the smallest integer $\ge x$ and α was selected from a grid of candidate values from 0 to 1 using a double bootstrap method. Following Chakraborty

et al. (2013), we applied the following double bootstrap method for choosing α from a grid of candidate values in our simulation and the MIMIC-III data example.

(1) Denote the estimate of the parameter of interest from original data as $\boldsymbol{c}^{\mathrm{T}}\widehat{\boldsymbol{\psi}}_1$, where $\boldsymbol{c}$ is an arbitrary vector of the same dimension as $\boldsymbol{\psi}_1$.

(2) Draw B_1 n-out-of-n first-level bootstrap samples from original data and calculate the $\widehat{p}_{n-regu}^{(b_1)}$ with $\nu = 0.001$ using the sample where $\overline{\boldsymbol{R}}_2^{(b_1)} = \mathbf{1}_2^{(b_1)}$ (i.e., the sample with complete covariates in the b_1th first-level bootstrap resampling) for $b_1 = 1, \ldots, B_1$. Fix α at the smallest value in the grid.

(3) Compute the corresponding values of $\widehat{m}^{(b_1)}$ by

$$\widehat{m}^{(b_1)} = \left\lceil n \frac{1+\alpha\left(1-\widehat{p}_{n-regu}^{(b_1)}\right)}{1+\alpha} \right\rceil,$$

$b_1 = 1, \ldots, B_1$.

(4) Conditional on each first-level bootstrap sample, draw B_2 $m^{(b_1)}$-out-of-n second-level (nested) bootstrap samples and estimate the parameter of interest $\boldsymbol{c}^{\mathrm{T}}\widehat{\boldsymbol{\psi}}_{1,\widehat{m}^{(b_1)}}^{(b_1 b_2)}$, $b_1 = 1, \ldots, B_1$, $b_2 = 1, \ldots, B_2$.

(5) For $b_1 = 1, ..., B_1$, compute the 2.5th and 97.5th percentiles of $\boldsymbol{c}^{\mathrm{T}}\widehat{\boldsymbol{\psi}}_{1,\widehat{m}^{(b_1)}}^{(b_1 b_2)}$; say $\widehat{L}_{db}^{(b_1)}$ and $\widehat{U}_{db}^{(b_1)}$, respectively.

(6) Estimate the coverage rate of the bootstrap confidence interval from all the first-level bootstrap datasets as

$$\frac{1}{B_1}\sum_{b_1=1}^{B_1} \mathbb{I}\left\{\widehat{L}_{db}^{(b_1)} \le \boldsymbol{c}^{\mathrm{T}}\widehat{\boldsymbol{\psi}}_1 \le \widehat{U}_{db}^{(b_1)}\right\}.$$

(7) If the above coverage rate is at or above the nominal level, then pick the current value of α as the final value. Otherwise, increment α to the next highest value in the grid.

(8) Repeat steps (3)–(7), until the coverage rate of the double bootstrap confidence interval attains the nominal coverage rate or the grid is exhausted.

As discussed in Chakraborty et al. (2013), the number of bootstrap replications in the above double bootstrap procedure to select α (i.e., B_1 and B_2) does not need to be equal

to the number of bootstrap replications used for inference (i.e., B) after α is selected. Since the kernel regression for estimating the missingness probability models in the WQ-EE and WQ-SA methods is computationally expensive, we set $B_1 = B_2 = 200$ in our simulation and the MIMIC-III data example.

8. Web Appendix H: Simulation

We conducted four simulation studies to evaluate the finite-sample performance of the WQ-EE and the WQ-SA methods. In Simulation 1, we focused on the scenario where Assumptions 1-5 were satisfied and demonstrated the performance of the WQ-EE method. In Simulation 2, we evaluated the robustness of the WQ-EE method against the violations of Assumption 5 when (1) the chosen instrumental variable was weakly associated with $R_{pse,t}$ and (2) there were interactions between $\boldsymbol{u}_t$ and $y_{pse,t}$ in the true missingness probability model. In Simulation 3, we evaluated the proposed simulation-based approach to specifying a plausible range of the sensitivity parameter and the performance of the WQ-SA method based on the specified range. In Simulation 4, we investigated the influence of an increased number of total stages when covariates are nonignorably missing.

For comparison, we also assessed three other methods for handling partially missing covariate data in Q-learning: a naive method which ignored partially observed covariates ('naive'), complete case analysis ('CC'), and multiple imputation ('MI') (Little and Rubin, 2014). Q-learning when all the covariates were fully observed ('All') was used as the benchmark. In the naive method, we omitted all terms containing the partially observed covariates in the specified Q-functions. Consequently, the Q-functions were misspecified in this scenario as Assumption 2 was violated. In the CC method, we excluded all patients whose covariates were subject to missingness in estimation. In the MI method, we first imputed missing covariates using predictive mean matching multiple times, which was implemented through the R package `MICE`. Subsequently, we applied the standard Q-learning algorithm to the

imputed data sets. Finally, we calculated the average of the estimated parameters from the imputed datasets to obtain the final estimates of the Q-function parameters.

All methods were evaluated in two aspects. Firstly, we compared the performances of the methods for estimating the blip function parameters $\boldsymbol{\psi}_t$. Secondly, we compared their abilities to identify the true optimal dynamic treatment regime (DTR) and the values of the estimated optimal DTRs. Besides, we also examined the coverage rates of the bootstrap confidence intervals for blip function parameters when the models in the WQ-EE and the WQ-SA methods were correctly specified in Simulation 1 and Simulation 3, respectively.

8.1 *Simulation 1: comparative performance of the WQ-EE method when Assumptions 1-5 were satisfied*

We simulated data from an observational study with two stages of intervention. There were two covariates $X_{t,1}$ and $X_{t,2}$ at stage $t = 1, 2$. Among these covariates, only $X_{1,2}$ and $X_{2,2}$ were partially observed, and their missingness indicators were R_1 and R_2, respectively. Denote $\text{expit}(b) = \exp(b)/\{1 + \exp(b)\}$ for b in $\mathcal{R}$. The variables were generated as follows:

$$\begin{pmatrix} X_{1,1} \\ X_{2,1} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right);$$

$X_{1,2} \sim \text{Uniform}(0, 2); R_1 \sim \text{Bernoulli}(\{1 + \exp(-3 + X_{1,2})\}^{-1});$

$A_1 \sim 2 \times \text{Bernoulli}(\text{expit}(-1 + X_{1,1} + X_{1,2} - R_1)) - 1;$

$Y_1 = A_1(\psi_{10} + 2 + \psi_{11}X_{1,2}) + \beta_{10} + \beta_{11}X_{1,1} + \beta_{12}X_{1,2} + \epsilon_1; \epsilon_1 \sim N(0, 3);$

$X_{2,2} \sim \text{Uniform}(0, 2); R_2 \sim \text{Bernoulli}(\{1 + \exp(-1 + 0.5X_{1,2} - Y_1 - X_{2,1} - 0.5X_{2,2})\}^{-1});$

$A_2 \sim 2 \times \text{Bernoulli}(\text{expit}(-1 - X_{1,1} - X_{1,2} + Y_1 + X_{2,1} - R_2)) - 1;$

$Y_2 = A_2(\psi_{20} + \psi_{2A}A_1 + \psi_{22}X_{2,2}) + \beta_{2A}A_1 + \beta_{22}X_{2,1} + \beta_{23}X_{2,2} + \epsilon_2; \epsilon_2 \sim N(0, 1).$

The true values of the parameters $\boldsymbol{\theta}_1 = (\psi_{10}, \psi_{11}, \beta_{10}, \beta_{11}, \beta_{12})$ and $\boldsymbol{\theta}_2 = (\psi_{20}, \psi_{2A}, \psi_{22}, \beta_{2A}, \beta_{22}, \beta_{23})$ were $(1, -1, 1.5, 0.5, -0.5)$ and $(1, -1, 1, -1, 1, -0.5)$, respectively. We generated 1000 data sets with sample sizes of 500 and 2000.

We considered $Y = Y_1 + Y_2$ as the final outcome. Under this setting, the true optimal treatment rule at stage 2 was $2\mathbb{I}(\psi_{20}+\psi_{2A}a_1+\psi_{22}x_{2,2} \geq 0)-1$. Because $\boldsymbol{\theta}_2 = (1, -1, 1, -1, 1, -0.5)$ and $1 - a_1 + x_{2,2} \geq 0$, the optimal choice of A_2 was always 1. Then, the pseudo-outcome at stage 1, $Y_{pse,1}$, was equal to $1 - 2A_1 + Y_1 + X_{2,1} + 0.5X_{2,2}$. Therefore, $Q_1(x_{1,1}, x_{1,2}, a_1) = 1.5 + \beta_{10} + a_1(\psi_{10} + \psi_{11}x_{1,2}) + (0.5 + \beta_{11})x_{1,1} - \beta_{12}x_{1,2}$. And the optimal treatment rule at stage 1 was $2\mathbb{I}(\psi_{10} + \psi_{11}X_{1,2} \geq 0) - 1$. Note that the missingness indicators of $X_{1,2}$ and $X_{2,2}$ were directly associated with their own values, thus they were nonignorable missing. R_1 only depended on $X_{1,2}$, and R_2 only depended on $(X_{1,2}, Y_1, X_{2,1}, X_{2,2})$, so the future-independent missingness assumption held. The missingness indicator for $Y_{pse,1}$ was the same as that for $X_{2,2}$, since $Y_{pse,1} = 1 - 2A_1 + Y_1 + X_{2,1} + 0.5X_{2,2}$ and $(A_1, Y_1, X_{2,1})$ were fully observed. Besides, since $P(R_2 = 1 \mid \boldsymbol{H}_2) = \{1 + \exp(-1 + 0.5X_{1,2} - Y_1 - X_{2,1} - 0.5X_{2,2})\}^{-1}$, it follows that $P(R_{pse,1} = 1 \mid \boldsymbol{H}_1, A_1, Y_{pse,1}) = \{1 + \exp(0.5X_{1,2} - 2A_1 - Y_{pse,1})\}^{-1}$, which indicates that $Y_{pse,1}$ was nonignorable missing and $X_{1,1}$ was conditional independent of $R_{pse,1}$ given $(X_{1,2}, A_1, Y_{pse,1})$. Since $X_{1,1}$ was also associated with $Y_{pse,1}$ after conditioning on $(X_{1,2}, A_1)$, we employed $X_{1,1}$ as a nonresponse instrumental variable.

[Figure 1 about here.]

Web Figure 1 shows the empirical distributions of the estimates of $\boldsymbol{\psi}_t = (\psi_{10}, \psi_{11}, \psi_{20}, \psi_{2A}, \psi_{22})$, which directly influence the estimated optimal DTR. At stage 2, the proposed WQ-EE method and the CC method yielded unbiased estimates of the blip function parameters since the future-independent missingness assumption held and the outcome at stage 2 was fully observed. In contrast, the naive and MI methods led to non-negligible biases of the blip function parameters. At stage 1, the estimators in the WQ-EE method remained unbiased

while the estimators in the naive, CC, and MI methods displayed relatively large biases, which were not mitigated as the sample size increased. Note that the naive estimates of ψ_{11} and ψ_{22} were unavailable, as we excluded the terms $A_1X_{1,2}$ and $A_2X_{2,2}$ from the specified Q-functions when implementing this method. And the misspecified Q-functions led to the relatively poor performance of the naive estimates. As for the bias of the CC estimates, it can be attributed to the violation of the ignorable missingness assumption of the pseudo-outcomes, which is necessary for the equation $E(Y_{pse,1} \mid \boldsymbol{h}_1, a_1) = E(Y_{pse,1} \mid \boldsymbol{h}_1, a_1, \overline{\boldsymbol{R}}_2 = \mathbf{1}_2)$ to hold. The bias of the MI estimates stemmed from the nonignorable missingness. It has been discussed that the standard MI method would be biased when data are nonignorably missing (Schafer and Graham, 2002).

The proportion of patients whose optimal treatments in both stages were correctly identified was high when employing the WQ-EE method, as demonstrated in Web Table 1. With the smaller sample size ($n = 500$), the WQ-EE method achieved an average correct classification rate of 89.9% across the 1000 simulated datasets, with a mean final outcome of 2.98. In general, larger sample sizes led to higher proportions of patients with their true optimal DTR correctly identified. In contrast, the DTRs obtained from the CC and MI methods had much lower correct classification rates and the mean final outcomes under the estimated DTRs were lower than that from the WQ-EE method.

[Table 1 about here.]

We further examined the coverage rates of the constructed confidence intervals. Following the recommendations and numerical studies of Shao and Wang (2016), we constructed the confidence intervals with the bootstrap method due to the complicated form of the asymptotic covariance matrix. For stage-2 parameters, non-regularity is not an issue, allowing the use of the ordinary n-out-of-n bootstrap method to construct confidence intervals.

For stage-1 parameters, we employed the m-out-of-n bootstrap method described in Web Appendix G to construct confidence intervals.

Because the kernel regression for estimating missingness probabilities in the WQ-EE method slows down the computation, the adaptive selection of α for determining the bootstrap resample size m in each simulated dataset is computationally prohibitive. To reduce the computation burden in our simulations, we set $B_1 = B_2 = 200$ in the double bootstrap method and set α at the fixed value that was determined by calculating the average of the α values adaptively selected from 10 simulated datasets, each with a sample size 500. The parameter of interest in the double bootstrap procedure was set as ψ_{10}.

With the determined α, we conducted 500 bootstrap resampling for each of the 1000 simulated datasets to calculate the n-out-of-n and m-out-of-n percentile bootstrap confidence intervals for the stage-2 and stage-1 blip function parameters, respectively. And the coverage rate of the bootstrap confidence intervals was calculated by the percentage of the confidence intervals that covered the true value of corresponding parameter.

Web Table 2 shows the results for the coverage rates of the bootstrap confidence intervals when employing the WQ-EE method in Simulation 1. We see that these coverage rates were close to the nominal level 95%.

[Table 2 about here.]

8.2 *Simulation 2: comparative performance of the WQ-EE method when Assumption 5 is violated*

To investigate the robustness of the proposed WQ-EE method when Assumption 5 was violated, we kept all other data generating mechanisms unchanged, while altering the true model for $R_{pse,1}$ to $P(R_{pse,1} = 1) = \{1 + \exp(\gamma_z X_{1,1} + 0.5X_{1,2} - 2A_1 - Y_{pse,1})\}^{-1}$. When $\gamma_z = 0$, the data generating mechanism was the same as that in Simulation 1 and Assumption 5 was satisfied. However, when $\gamma_z \neq 0$, $X_{1,1}$ was associated with $R_{pse,1}$ conditional on $(X_{1,2}, A_1, Y_{pse,1})$.

Thus there were no nonresponse instrumental variables. We set $\gamma_z \in \{-0.4, -0.2, 0.2, 0.4\}$ such that $X_{1,1}$ was weakly associated with $P(R_{pse,1} = 1)$ conditional on $X_{1,2}$ and $Y_{pse,1}$. We generated 1000 Monte Carlo samples with sample size $n = 500$.

Since the future-independent missingness assumption still held in Simulation 2, the estimators for $(\psi_{20}, \psi_{2A}, \psi_{22})$ in the WQ-EE method were still consistent. We report the performance of the estimators for (ψ_{10}, ψ_{11}) and the value and correct classification rate of the estimated optimal DTRs. The results were calculated by averaging over 1000 simulated data sets and are shown in Web Table 3. We see that the proposed WQ-EE method was robust when X_{11} was weakly related to $R_{pse,1}$. Despite the declining performance of the corresponding optimal DTR as the absolute value of γ_z increased, the WQ-EE method consistently outperformed the other three methods with higher mean final outcomes and correct classification rates.

[Table 3 about here.]

We also considered scenarios where there were interactions between $y_{pse,t}$ and $\boldsymbol{u}_t$ in the true missingness probability model. The true model for $R_{pse,1}$ was modified as $P(R_{pse,1} = 1) = [1 + \exp\{0.5X_{1,2} - 2A_1 - Y_{pse,1} + \gamma_{uy}(X_{1,2} - 1)Y_{pse,1}\}]^{-1}$, while the data generating mechanism for other covariates was the same as that in Simulation 1. The specific form for the interaction, $(X_{1,2} - 1)Y_{pse,1}$, was chosen to ensure that the missing data proportions were largely similar between Simulations 1 and 2. We set $\gamma_{uy} \in \{-0.2, -0.1, 0.1, 0.2\}$ by allowing weak interactions between $X_{1,2}$ and $Y_{pse,1}$. We again generated 1000 simulated data sets with sample size $n = 500$. The results of the compared methods are shown in Web Table 4. The WQ-EE method exhibited superior performance compared to the other three methods despite the introduction of interaction terms involving $Y_{pse,1}$, showcasing its robustness in the presence of mild model misspecification.

[Table 4 about here.]

8.3 *Simulation 3: The performance of the WQ-SA method and the proposed simulation-based approach to calibrating the magnitude of the sensitivity parameters*

To evaluate the performance of the WQ-SA method and the proposed simulation-based approach to calibrating the magnitude of the sensitivity parameters, we simulated data from a 2-stage scenario with two covariates at each stage. At stage $t = 1, 2$, there existed two covariates, $X_{t,1}$ and $X_{t,2}$. Among these covariates, only $X_{1,2}$ and $X_{2,2}$ were partially observed, with their corresponding missingness indicators denoted as R_1 and R_2, respectively. The data-generation mechanism was as follows:

$$\begin{pmatrix} X_{1,1} \\ X_{2,1} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right);$$

$X_{1,2} \sim \text{Uniform}(0, 2)$;

$R_1 \sim \text{Bernoulli}(\{1 + \exp(-1 - X_{1,2})\}^{-1})$;

$A_1 \sim 2 \times \text{Bernoulli}(\text{expit}(1 - X_{1,1} - X_{1,2} + R_1)) - 1$;

$Y_1 = A_1(\psi'_{10} + 2 + \psi'_{11}X_{1,2}) + \beta'_{10} + \beta'_{11}X_{1,1} + \beta'_{12}X_{1,2} + \epsilon'_1; \epsilon'_1 \sim N(0, 3)$;

$X_{2,2} \sim \text{Uniform}(0, 2)$;

$R_2 \sim \text{Bernoulli}([1 + \exp\{2.5 - \exp(2X_{1,1}) - 0.5X_{1,2} + A_1 + Y_1 + X_{2,1} + 0.5X_{2,2}\}]^{-1})$;

$A_2 \sim 2 \times \text{Bernoulli}(\text{expit}(-1 + X_{1,1} - 2A_1 + X_{2,1})) - 1$;

$Y_2 = A_2(\psi'_{20} + \psi'_{2A}A_1 + \psi'_{22}X_{2,2}) + \beta'_{2A}A_1 + \beta'_{22}X_{2,1} + \beta'_{23}X_{2,2} + \epsilon'_2; \epsilon'_2 \sim N(0, 1)$.

The true values of the parameters $\boldsymbol{\theta}'_1 = (\psi'_{10}, \psi'_{11}, \beta'_{10}, \beta'_{11}, \beta'_{12})$ and $\boldsymbol{\theta}'_2 = (\psi'_{20}, \psi'_{2A}, \beta'_{22}, \beta'_{2A}, \beta'_{22}, \beta'_{23})$ were $(-1, 1, -2, -0.5, -1)$ and $(1, -1, 1, -1, 1, -0.5)$, respectively. The final outcome was set to be $Y = Y_1 + Y_2$.

The true optimal treatment rule at stage 2 was $2\mathbb{I}(1 - a_1 + x_{2,2} \geq 0) - 1$. Since $a_1 \in \{-1, 1\}$ and $x_{2,2} \geq 0$, the optimal choice of A_2 was always 1. With this optimal treatment at stage 2, $Y_{pse,1} =$

$1-2A_1+Y_1+X_{2,1}+0.5X_{2,2}$. Then, the optimal treatment rule at stage 1 was $2\mathbb{I}(\psi_{10}+\psi_{11}x_{1,2} \geq 0) - 1$. Since the missingness indicators of $X_{1,2}$ and $X_{2,2}$ were directly related to their own values, they were nonignorably missing. The future-independent missingness assumption held because R_1 only depended on $X_{1,2}$, and R_2 only depended on $(X_{1,1}, X_{1,2}, A_1, Y_1, X_{2,1}, X_{2,2})$. Because $Y_{pse,1} = 1 - 2A_1 + Y_1 + X_{2,1} + 0.5X_{2,2}$, $R_{pse,1} = R_2$, $P(R_{pse,1} = 1) = \{1.5 - \exp(2X_{1,1}) - 0.5X_{1,2} + 3A_1 + Y_{pse,1}\}^{-1}$, $Y_{pse,1}$ was nonignorably missing and there was no nonresponse instrumental variable. And the true value of the sensitivity parameter was $\gamma_1' = 1$ in this scenario.

We first evaluated the performance of the simulation-based approach to calibrating the magnitude of the sensitivity parameter.

As discussed in Web Appendix A, we assumed the sign of the sensitivity parameter γ_1' was known and chose $[0, 7]$ as an initial range for γ_1'. To zoom in to a more plausible range, we employed the simulation-based approach proposed in Web Appendix A. Specifically, we first generated 1000 data sets with sample size $n = 500$ and 2000, respectively. For each data set, to evaluate the candidate values of γ_1', we implemented the steps described in the end of Web Appendix A. Particularly, in Step (2), $\text{pr}(y_{pse,1} \mid x_{1,1}, x_{1,2}, a_1, R_1 = 1, R_2 = 1)$ was obtained by combing the kernel regression estimates of $E(y_{pse,1} \mid x_{1,1}, x_{1,2}, a_1, R_1 = 1, R_2 = 1)$ and its residual density estimates, which were obtained by kernel density estimation method with the R function `density`. And $E\{\exp(\gamma_1' y_{pse,1}) \mid x_{1,1}, x_{1,2}, a_1, R_1 = 1, R_2 = 1\}$ was also estimated by kernel regression methods. In Step (3), based on the estimated $\text{pr}(y_{pse,1} \mid x_{1,1}, x_{1,2}, a_1, R_1 = 0, R_2 = 1)$, we imputed $Y_{pse,1}^{\gamma_1', imp}$ for 1000 Monte Carlo replication times. And in Step (4), we estimated $\text{pr}(y_{pse,1}^{\gamma_1',*} \mid x_{1,1}, x_{1,2}, a_1, R_1 = 1)$ with the correctly specified Q-function for the conditional mean and R function `density` for the residuals. Further, we generated the replications of observed pseudo-outcomes and compared them with the observed pseudo-outcome estimates using the Wilcoxon ranked sum test in Steps (5) and

(6), respectively. Finally, we calculated the medians of the Wilcoxon ranked sum test p-values from the 1000 Monte Carlo replicates for each simulated data set, and we reported the summary statistics of these p-value medians from the 1000 simulated data sets and the probability to be included in the range of plausible values determined by the threshold of 0.05 in Web Table 5.

[Table 5 about here.]

We see that for both sample sizes, the summary statistics corresponding to the true value $\gamma_1' = 1$ were largest among candidate values of γ_1', and they decreased as the specified sensitivity parameter moving away from its true value. In analogy to the conventional significance level 0.05 for hypothesis testing based on p-values, we considered a threshold of the p-value medians 0.05 to determine a plausible range of γ_1' for each simulated data set, where a value of γ_1' was included in the determined range if its corresponding p-value median was greater than 0.05. The true value for γ_1' was 100% included in the determined ranges when applying this threshold to the 1000 data sets. This indicated that applying the 0.05 threshold for the p-value medians can ensure that the determined plausible ranges would cover the true value of γ_1'.

With larger sample sizes, the summary statistics of the p-value medians for the true value of γ_1' was stable while those for other values of γ_1' decreased. With a sample size of 2000, $\gamma_1' = 3$ was included in the determined plausible range for sensitivity parameter in 54 out of 1000 data sets, and larger values of γ_1' were selected even less frequently.

We compared the performance of the WQ-SA method with $\gamma_1' \in [0, 7]$ with other methods. First we assessed their performance in estimating blip function parameters and optimal DTRs. Similar to Simulations 1 and 2, the future-independent missingness assumption still held in Simulation 3, and we reported the performance of the estimators for (ψ_{10}', ψ_{11}') and

the values and correct classification rates of the estimated optimal DTRs. The results were calculated by averaging over 1000 simulated datasets and are shown in Web Table 6.

[Table 6 about here.]

We see that when the specified sensitivity parameter was equal to its true value, the WQ-SA method had the best performance. Notably, when $\gamma_t' = 0$ was specified, it was equivalent to assuming ignorable missingness, that is, the missingness of $Y_{pse,1}$ was conditionally independent with its own values. Except for this specification, the WQ-SA method exhibited relative robustness to the sensitivity parameter specification and consistently outperformed the naive, CC, and MI methods with higher mean final outcomes and correct classification rates under the estimated optimal DTRs.

Finally, we examined the coverage rates of the confidence intervals for blip function parameters when employing the WQ-SA method with the correctly specified sensitivity parameter value $\gamma_t' = 1$. Similar to Simulation 1, the confidence intervals for stage-2 parameters were constructed by ordinary n-out-of-n bootstrap method, while those for stage-1 parameters were constructed by m-out-of-n bootstrap method described in Web Appendix G. And the bootstrap resample sizes were selected in the same way as in Simulation 1. Web Table 7 shows the coverage rates of the bootstrap confidence intervals, which were close to the nominal level 95%.

[Table 7 about here.]

In summary, Simulation 3 demonstrated that the proposed simulation-based approach could identify a plausible range for the sensitivity parameter. Additionally, it showed that the WQ-SA method was robust and outperformed the naive, CC, and MI methods within a wide range of the sensitivity parameter. Furthermore, it confirmed that the 95% confidence intervals constructed by the bootstrap percentiles performed well when all the models were correctly specified for the WQ-SA method.

8.4 *Simulation 4: the impact of an increased number of total stages with nonignorable missing covariates*

To investigate the influence of the nonignorable missing covariates when the number of total stages increases, we designed a simulation study with three stages of intervention. To facilitate comparisons with the 2-stage scenario in Simulation 1, we used the same data generation mechanism for stage 1 and stage 2 as that in Simulation 1, and generated data for stage 3 as follows:

$$
\begin{aligned}
&X_{3,1} \sim N(0,1); X_{3,2} \sim \text{Uniform}(0,2);\\
&R_3 \sim \text{Bernoulli}(\{1+\exp(-Y_1-Y_2-0.5X_{3,1}-|X_{3,2}-1|)\}^{-1});\\
&A_3 \sim 2\times \text{Bernoulli}(\text{expit}(-1+X_{3,1}+X_{3,2}))-1;\\
&Y_3 = \beta_{30}+A_3(\psi_{30}+\psi_{3,1}X_{3,2})+\beta_{31}X_{31}+\epsilon_3; \epsilon_3 \sim N(0,1).
\end{aligned}
$$

The true values of the parameters $\boldsymbol{\theta}_3 = (\psi_{30}, \psi_{31}, \beta_{30}, \beta_{31})$ were $(-1, 1, -0.5, 0.5)$. R_3 was the missingness indicator for the partially observed variable $X_{3,2}$, while $X_{3,1}$ was fully observed. Similar to Simulation 1, we generated 1000 data sets with sample sizes 500 and 2000.

The final outcome was assumed to be $Y = Y_1 + Y_2 + Y_3$. The true optimal treatment rule at stage 3 was $2\mathbb{I}(\psi_{30}+\psi_{3,1}X_{3,2} \geq 0) - 1$. With the optimal treatment at stage 3, the pseudo-outcome at stage 2 was $Y_{pse,2} = -0.5+|-1+X_{3,2}|+0.5X_{3,1}+Y_1+Y_2$. Because the data generation mechanism at stages 1 and 2 was unchanged from Simulation 1 and $E(Y_{pse,2} \mid \boldsymbol{H}_2, A_2, Y_2) = Y_1 + Y_2$, the Q-functions, the optimal treatment rules at stages 1 and 2, and $(Y_{pse,1}, R_{pse,1})$ were the same as those in Simulation 1. Since R_3 only depended on $(X_{3,1}, X_{3,2})$, the future-independent missingness assumption held. Among the partially observed variables, only $X_{3,2}$ was employed in the calculation of $Y_{pse,2}$. Therefore, $R_{pse,2} = R_3$. Combining the form of the pseudo-outcomes and the missingness mechanism, we have $P(R_{pse,2} = 1 \mid \boldsymbol{H}_2, A_2, Y_{pse,2}) = \{1+\exp(-Y_{pse,2})\}^{-1}$. Since $(X_{2,1}, X_{2,2})$ were associated with $Y_{pse,2}$ and independent of $R_{pse,2}$

given $Y_{pse,2}$, we employed $(X_{2,1}, X_{2,2})$ as nonresponse instrumental variables when estimating the missingness model for $R_{pse,2}$.

[Table 8 about here.]

Web Table 8 presents the performance of the methods in comparison in the 3-stage scenario. Compared to Web Table 1 in the 2-stage scenario, the means of the values under the estimated DTRs for all methods decreased. For the All and the WQ-EE method, this decrease was small and it became smaller as the sample size increased. It can be explained by the finite-sample bias accumulated in the backward-induction procedure. In contrast, for the naive, CC, and MI methods, the deterioration of performance was more pronounced, indicating that, if it is not handled properly, the consequence of the nonignorable missing covariates problem would be more severe when the number of total stages increases. To better understand the reasons behind the performance deterioration, we examined the correct classification rates at different stages. Since we have presented the correct classification rates at stages 1 and 2 in Simulation 1, and the true Q-functions and optimal treatments rules at stages 1 and 2 in Simulation 1 and Simulation 4 were the same, we examined the correct classification rates at stage 3 and at stages 1 and 2, respectively.

[Table 9 about here.]

We used the results in Web Table 9 to investigate why the increased number of total stages led to lower overall correct classification rate if nonignorable missing covariate were not handled properly. For the naive method, its correct classification rate at stages 1 and 2 in Simulation 4 was almost the same as that in Simulation 1 (Web Table 1), and the decrease in its overall correct classification rate in Web Table 8 was mainly explained by its poor average correct classification rate at stage 3, which was near 0.5. In contrast, for the CC and the MI method, the average correct classification rates at stage 3 was higher than 0.93, but their average correct classification rates at stages 1 and 2 dropped a lot compared to those in

Simulation 1 (Web Table 1). And this decrease was due to the bias from nonignorable missing pseudo-outcomes accumulated over more stages in the backward induction procedure.

In summary, an increased number of total stages would generally lead to lower overall correct classification rates of the estimated optimal DTRs when covariates had nonignorable missing values. If the nonignorable missingness was properly addressed, the decrease in the overall correct classification rates would be small and could be mitigated by increasing the sample size. However, if the nonignorable missingness was not properly handled, the decrease in overall correct classification rates could be substantial due to the bias accumulated in the backward-induction procedure.

9. Web Appendix I: Additional details of the MIMIC-III data and the estimation procedure

The proposed methods were applied to the electronic medical records and administrative data from the MIMIC-III database (Johnson et al., 2016). MIMIC-III is a retrospectively collected and freely available database accessible through PhysioNet (Goldberger et al., 2000). The MIMIC-III database contains de-identified and anonymized health information from over 45,000 patients who were treated in the Intensive Care Units (ICUs) of Beth Israel Deaconess Medical Center in Boston, Massachusetts. We utilized similar patient selection criteria as in Speth et al. (2022). The detailed patient selection criteria are presented in Web Figure 2.

[Figure 2 about here.]

As depicted in Web Figure 2, This study included adult patients (age $\geq$ 18) diagnosed with acute emergent sepsis who were transferred from the Emergency Department (ED) and admitted to the medical ICU. Patients must be sent to the medical ICU within 1 hour after hospital admission and had documented fluid intake within the first hour after hospital admission. Furthermore, following Speth et al. (2022), to ensure a relatively homogeneous

patient cohort, we only included patients who survived for a minimum of 48 hours after ICU admission.

Since there were very few patients with hemodynamic variables recorded precisely at 180 minutes after ICU admission, we calculated the mean of the hemodynamic variables recorded within a half-hour window around the third hour. That is, we used the data collected between 2.5 and 3.5 hours after ICU admission as surrogates for the values of the hemodynamic variables three hours after ICU admission. If a hemodynamic variable had no record during this specific time interval, we considered the corresponding record three hours after admission to be missing. Under this criterion, the missing data proportion of hemodynamic variables is provided in Web Table 10.

[Table 10 about here.]

Following Xu et al. (2015), we performed cross-validation and m-out-of-n bootstrap to assess the performance of the estimated optimal DTRs. Specifically, to estimate the expected out-of-sample improvement of SOFA score at 24 hours post-admission, we randomly split the MIMIC-III data for 1000 times. For each split, we used 4/5 of the data (rounded to the nearest integer) as the training set and the remaining data as the testing set. The semiparametric missingness probability model and the Q-functions were re-estimated based on the training set of that split. The expected outcome in the testing set, under the estimated DTR, was then computed using the estimated missingness probability models and Q-functions. It should be noted that the optimal DTR may differ across different splits. Nevertheless, the average of the expected optimal outcomes in the testing sets from repeated splitting can be interpreted as the expected SOFA score at 24 hours after ICU admission that would have been obtained with out-of-sample data under the estimated regimes.

For the inference of the improvement of SOFA score at 24 hours post-admission, we further calculated the bootstrap standard deviation and the 95% confidence interval using the m-out-

of-n bootstrap method based on the original data (i.e., without cross-validation). We used the double bootstrap method described in Web Appendix G to determine the bootstrap resample size m, with the number of bootstrap replications set as $B_1 = B_2 = 200$. Specifically, when selecting m and conducting m-out-of-n bootstrap for the MI method, we followed the suggestion from Shen et al. (2023) and used 25 imputations.

References

Andrews, D. W. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica* pages 399–405.

Chakraborty, B., Laber, E. B., and Zhao, Y. (2013). Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme. *Biometrics* **69,** 714–723.

Chakraborty, B., Murphy, S., and Strecher, V. (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical methods in medical research* **19,** 317–343.

Copas, J. and Eguchi, S. (2005). Local model uncertainty and incomplete-data bias (with discussion). *Journal of the Royal Statistical Society Series B: Statistical Methodology* **67,** 459–513.

Copas, J. and Shi, J. Q. (2000). Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics* **1,** 247–262.

Daniels, M. J., Chatterjee, A. S., and Wang, C. (2012). Bayesian model selection for incomplete data using the posterior predictive distribution. *Biometrics* **68,** 1055–1063.

European Medicines Agency (2010). Guideline on missing data in confirmatory clinical trials. `https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-missing-data-confirmatory-clinical-trials_en.pdf`.

European Medicines Agency (2020). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for

clinical trials. https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-e9-r1-addendum-estimands-and-sensitivity-analysis-clinical-trials-guide en.pdf.

Franks, A., D'Amour, A., and Feller, A. (2020). Flexible sensitivity analysis for observational studies without observable implications. *Journal of the American Statistical Association* **115,** 1730–1746.

Gelman, A., Meng, X.-L., and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica sinica* pages 733–760.

Gelman, A., Van Mechelen, I., Verbeke, G., Heitjan, D. F., and Meulders, M. (2005). Multiple imputation for model checking: Completed-data plots with missing and latent data. *Biometrics* **61,** 74–85.

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation* **101,** e215–e220.

Johnson, A. E., Pollard, T. J., Shen, L., Lehman, L.-w. H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Anthony Celi, L., and Mark, R. G. (2016). Mimic-iii, a freely accessible critical care database. *Scientific data* **3,** 1–9.

Little, R. J., D'agostino, R., Cohen, M. L., Dickersin, K., Emerson, S. S., Farrar, J. T., Frangakis, C., Hogan, J. W., Molenberghs, G., Murphy, S. A., et al. (2012). The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine* **367,** 1355–1360.

Little, R. J. and Rubin, D. B. (2014). *Statistical Analysis with Missing Data*, volume 333. John Wiley & Sons.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing.

Handbook of econometrics **4,** 2111–2245.

Schafer, J. L. and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods* **7,** 147.

Shao, J. and Wang, L. (2016). Semiparametric inverse propensity weighting for nonignorable missing data. *Biometrika* **103,** 175–187.

Shen, J., Hubbard, R. A., and Linn, K. A. (2023). Estimation and evaluation of individualized treatment rules following multiple imputation. *Statistics in Medicine* **42,** 4236–4256.

Simoneau, G., Moodie, E. E., Nijjar, J. S., Platt, R. W., Investigators, S. E. R. A. I. C., et al. (2020). Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association* **115,** 1531–1539.

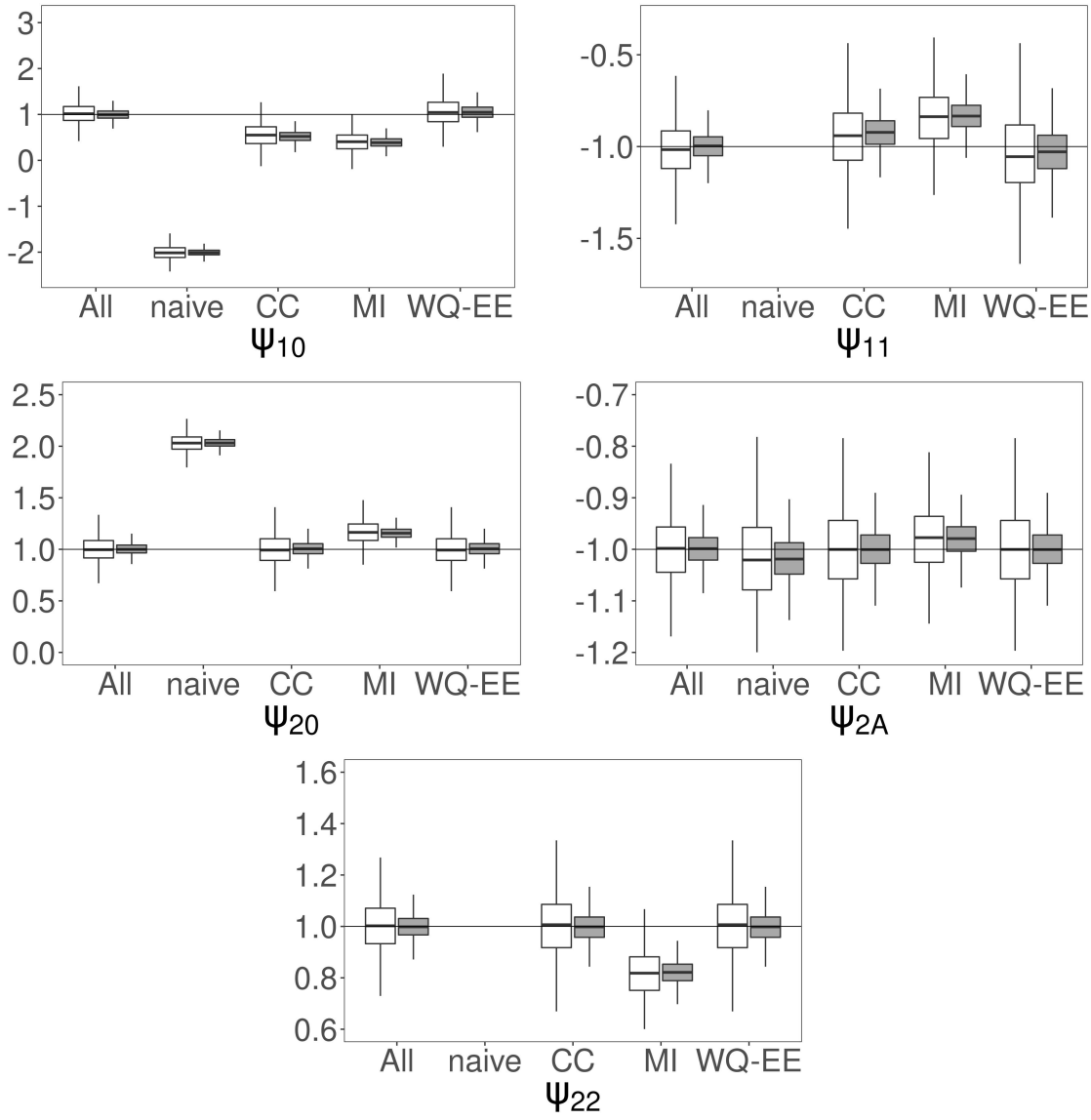
Sjölander, A., Gabriel, E. E., and Ciocănea-Teodorescu, I. (2022). Sensitivity analysis for causal effects with generalized linear models. *Journal of Causal Inference* **10,** 441–479.

Speth, K. A., Elliott, M. R., Marquez, J. L., and Wang, L. (2022). Penalized spline-involved tree-based (pensit) learning for estimating an optimal dynamic treatment regime using observational data. *Statistical Methods in Medical Research* **31,** 2338–2351.

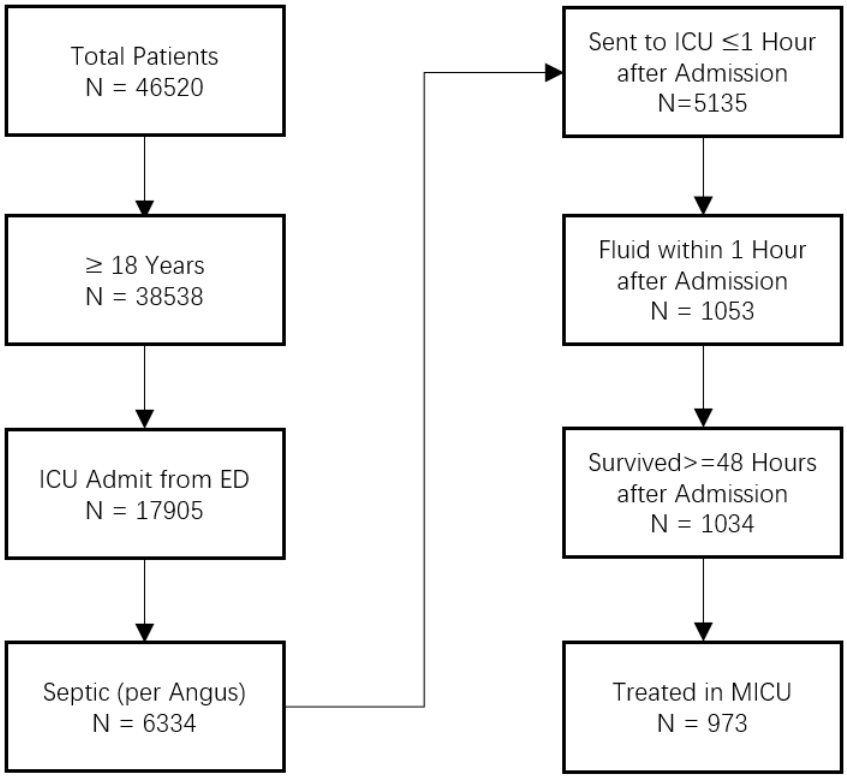
Wallace, M. P. and Moodie, E. E. (2015). Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics* **71,** 636–644.

Xu, Y., Yu, M., Zhao, Y.-Q., Li, Q., Wang, S., and Shao, J. (2015). Regularized outcome weighted subgroup identification for differential treatment effects. *Biometrics* **71,** 645–653.

Yin, P. and Shi, J. Q. (2019). Simulation-based sensitivity analysis for non-ignorably missing data. *Statistical methods in medical research* **28,** 289–308.



Web Figure 1: Boxplots of blip function parameter estimates from the compared methods in Simulation 1. In each boxplot, white and gray boxes are the results for the sample sizes 500 and 2,000, respectively. The horizontal line marks the true values of the blip function parameter estimates. Note that the naive method did not generate estimates of ψ_{11} and ψ_{22} as ψ_{11} and ψ_{22} are the coefficients of partially missing $X_{1,2}$ and $X_{2,2}$ in the true treatment rules.



Web Figure 2: Eligibility criteria and patient population for the selected cohort from the MIMIC-III database.

Web Table 1: Performance of estimated DTRs in Simulation 1

	n=500		n=2000	
	Value	Opt%	Value	Opt%
All	2.988 (0.057)	0.928 (0.052)	2.997 (0.028)	0.964 (0.024)
naive	2.501 (0.053)	0.500 (0.022)	2.500 (0.027)	0.499 (0.011)
CC	2.879 (0.128)	0.773 (0.111)	2.898 (0.059)	0.777 (0.057)
MI	2.832 (0.139)	0.729 (0.108)	2.849 (0.071)	0.730 (0.058)
WQ-EE	2.974 (0.069)	0.899 (0.073)	2.994 (0.029)	0.949 (0.036)

Note: Value, the value of the estimated regime; Opt%, the percentage of patients correctly classified to their true optimal treatments in both stages; The numbers outside parentheses represent the sample means, while the numbers inside parentheses represent the sample standard errors.

Web Table 2: Coverage rates of the 95% bootstrap confidence intervals when employing the WQ-EE method in Simulation 1

	m	ψ_{10}	ψ_{11}	ψ_{20}	ψ_{2A}	ψ_{22}
Sample size 500	473	0.947	0.937	0.951	0.941	0.946
Sample size 2000	1956	0.947	0.945	0.953	0.950	0.945

Note: m, the average resample size in the m-out-of-n bootstrap method in 1000 simulated datasets, rounded to integer. For stage-1 parameters (ψ_{10}, ψ_{11}), the 95% confidence intervals were constructed by m-out-of-n bootstrap percentiles. For stage-2 parameters $(\psi_{20}, \psi_{2A}, \psi_{22})$, the 95% confidence intervals were constructed by standard (n-out-of-n) bootstrap percentiles.

Web Table 3: Performance of the WQ-EE method without valid nonresponse instrumental variable in comparison with alternative methods in Simulation 2

	ψ_{10}		ψ_{11}		Value		Opt%	
Method	Bias	Std	Bias	Std	Mean	Std	Mean	Std
			$\gamma_z = -0.4$					
naive	-1.005	0.168	—	—	2.498	0.054	0.499	0.022
CC	0.068	0.188	-0.447	0.277	2.881	0.131	0.773	0.113
MI	0.166	0.169	-0.588	0.242	2.838	0.144	0.738	0.116
WQ-EE	-0.042	0.237	0.113	0.318	2.971	0.078	0.896	0.080
			$\gamma_z = -0.2$					
naive	-1.008	0.166	—	—	2.498	0.054	0.500	0.023
CC	0.062	0.184	-0.443	0.266	2.885	0.126	0.781	0.113
MI	0.161	0.165	-0.591	0.231	2.838	0.142	0.733	0.113
WQ-EE	-0.043	0.243	0.090	0.315	2.972	0.081	0.898	0.076
			$\gamma_z = 0.2$					
naive	-1.000	0.155	—	—	2.502	0.052	0.499	0.022
CC	0.073	0.185	-0.482	0.253	2.871	0.125	0.766	0.110
MI	0.174	0.164	-0.627	0.227	2.816	0.144	0.718	0.113
WQ-EE	-0.040	0.239	0.031	0.310	2.974	0.070	0.899	0.072
			$\gamma_z = 0.4$					
naive	-1.007	0.151	—	—	2.501	0.053	0.499	0.023
CC	0.088	0.185	-0.495	0.274	2.862	0.132	0.755	0.114
MI	0.183	0.163	-0.639	0.238	2.808	0.146	0.705	0.114
WQ-EE	-0.020	0.233	0.053	0.321	2.968	0.074	0.895	0.078

Note: Value, the value of the estimated regime; Opt%, the percentage of patients correctly classified to their true optimal treatments in both stages; Bias, the bias of the sample mean; Std, the sample standard error; Mean, the sample mean.

Web Table 4: Performance of the WQ-EE method when interaction terms between $y_{pse,t}$ and $\boldsymbol{u}_t$ were incorrectly omitted, in comparison with alternative methods in Simulation 2

	ψ_{10}		ψ_{11}		Value		Opt%	
Method	Bias	Std	Bias	Std	Mean	Std	Mean	Std
			$\gamma_{uy} = -0.2$					
naive	-1.011	0.160	—	—	2.500	0.053	0.500	0.022
CC	0.124	0.184	-0.514	0.269	2.863	0.139	0.760	0.115
MI	0.215	0.162	-0.656	0.233	2.808	0.152	0.712	0.115
WQ-EE	-0.067	0.239	0.075	0.324	2.972	0.072	0.898	0.073
			$\gamma_{uy} = -0.1$					
naive	-1.002	0.153	—	—	2.500	0.053	0.499	0.022
CC	0.084	0.182	-0.474	0.270	2.871	0.132	0.764	0.116
MI	0.185	0.159	-0.624	0.230	2.820	0.141	0.715	0.115
WQ-EE	-0.062	0.228	0.065	0.318	2.971	0.072	0.897	0.075
			$\gamma_{uy} = 0.1$					
naive	-1.002	0.157	—	—	2.499	0.053	0.500	0.022
CC	0.046	0.182	-0.441	0.258	2.886	0.119	0.780	0.105
MI	0.147	0.163	-0.595	0.231	2.832	0.143	0.734	0.107
WQ-EE	-0.021	0.233	0.022	0.307	2.974	0.074	0.898	0.077
			$\gamma_{uy} = 0.2$					
naive	-1.010	0.160	—	—	2.497	0.051	0.499	0.023
CC	0.046	0.188	-0.444	0.263	2.883	0.124	0.772	0.112
MI	0.146	0.168	-0.590	0.234	2.835	0.140	0.732	0.111
WQ-EE	0.023	0.242	-0.011	0.308	2.972	0.079	0.895	0.082

Note: Value, the value of the estimated regime; Opt%, the percentage of patients correctly classified to their true optimal treatments in both stages; Bias, the bias of the sample mean; Std, the sample standard error; Mean, the sample mean.

Web Table 5: Performance of the proposed simulation-based approach to determining a plausible range of the sensitivity parameter in Simulation 3

γ_t'	0	1	2	3	4	5	6	7
				Sample size = 500				
minimum	0.211	0.216	0.049	0.006	0.002	0.000	0.000	0.000
median	0.591	0.615	0.452	0.273	0.201	0.143	0.101	0.073
mean	0.578	0.601	0.440	0.297	0.232	0.185	0.150	0.122
maximum	0.754	0.756	0.744	0.762	0.695	0.665	0.671	0.662
inclusion probability	1.000	1.000	0.997	0.966	0.895	0.814	0.722	0.605
				Sample size = 2000				
minimum	0.105	0.338	0.001	0.000	0.000	0.000	0.000	0.000
median	0.377	0.624	0.045	0.010	0.003	0.001	0.000	0.000
mean	0.381	0.618	0.061	0.015	0.008	0.005	0.003	0.002
maximum	0.681	0.757	0.342	0.225	0.089	0.108	0.094	0.072
inclusion probability	1.000	1.000	0.448	0.055	0.026	0.013	0.012	0.007

Note: minimum, minimum of p-value medians in the 1000 data sets; median, median of p-value medians in the 1000 data sets; mean, mean of p-value medians in the 1000 data sets; maximum, maximum of p-value medians in the 1000 data sets; inclusion probability, the probability to be included in the selected range based on threshold 0.05.

Web Table 6: Performance of the WQ-SA method in Simulation 3

	ψ'_{10}		ψ'_{11}		Value		Opt%	
Method	Bias	Std	Bias	Std	Mean	Std	Mean	Std
Sample size = 500								
All	0.016	0.234	-0.011	0.158	-1.018	0.045	0.923	0.055
naive	1.171	0.162	—	—	-1.498	0.035	0.503	0.023
CC	-0.319	0.298	-0.078	0.209	-1.131	0.121	0.767	0.114
MI	-0.269	0.264	-0.178	0.186	-1.184	0.138	0.721	0.113
WQ-SA ($\gamma'_1 = 0$)	-0.037	0.213	-0.395	0.298	-1.134	0.124	0.760	0.111
WQ-SA ($\gamma'_1 = 1$)	-0.056	0.241	0.034	0.335	-1.033	0.071	0.893	0.084
WQ-SA ($\gamma'_1 = 2$)	-0.049	0.251	0.083	0.347	-1.034	0.071	0.891	0.084
WQ-SA ($\gamma'_1 = 3$)	-0.056	0.252	0.062	0.339	-1.036	0.070	0.889	0.086
WQ-SA ($\gamma'_1 = 4$)	-0.072	0.256	0.058	0.344	-1.037	0.077	0.887	0.088
WQ-SA ($\gamma'_1 = 5$)	-0.070	0.264	0.025	0.349	-1.041	0.077	0.883	0.091
WQ-SA ($\gamma'_1 = 6$)	-0.053	0.263	-0.015	0.361	-1.042	0.079	0.880	0.093
WQ-SA ($\gamma'_1 = 7$)	-0.060	0.271	-0.023	0.359	-1.044	0.080	0.879	0.095
Sample size = 2000								
All	0.005	0.118	-0.002	0.080	-1.003	0.020	0.965	0.025
naive	1.173	0.077	—	—	-1.500	0.014	0.500	0.011
CC	-0.324	0.153	-0.075	0.106	-1.101	0.053	0.774	0.057
MI	-0.296	0.134	-0.152	0.096	-1.148	0.065	0.731	0.056
WQ-SA ($\gamma'_1 = 0$)	-0.036	0.102	-0.421	0.143	-1.120	0.056	0.756	0.055
WQ-SA ($\gamma'_1 = 1$)	-0.026	0.140	0.035	0.182	-1.007	0.022	0.948	0.037
WQ-SA ($\gamma'_1 = 2$)	-0.019	0.132	-0.036	0.177	-1.008	0.023	0.943	0.043
WQ-SA ($\gamma'_1 = 3$)	-0.021	0.158	-0.019	0.205	-1.009	0.024	0.939	0.043
WQ-SA ($\gamma'_1 = 4$)	-0.020	0.163	-0.066	0.200	-1.012	0.026	0.931	0.051
WQ-SA ($\gamma'_1 = 5$)	-0.033	0.165	-0.098	0.206	-1.019	0.029	0.915	0.056
WQ-SA ($\gamma'_1 = 6$)	-0.032	0.164	-0.137	0.203	-1.027	0.035	0.897	0.062
WQ-SA ($\gamma'_1 = 7$)	-0.051	0.168	-0.139	0.209	-1.031	0.037	0.887	0.063

Note: WQ-SA (·), the WQ-SA method based on corresponding sensitivity parameters; Value, the value of the estimated regime; Opt%, the percentage of patients correctly classified to their true optimal treatments in both stages; Bias, the bias of the sample mean; Std, the sample standard error; Mean, the sample mean.

Web Table 7: Coverage rates of the bootstrap 95% confidence intervals when employing the WQ-SA method in Simulation 3

	m	ψ_{10}	ψ_{11}	ψ_{20}	ψ_{2A}	ψ_{22}
Sample size 500	413	0.961	0.958	0.950	0.960	0.961
Sample size 2000	1864	0.946	0.944	0.947	0.948	0.939

Note: m, the average resample size in the m-out-of-n bootstrap method in 1000 simulated datasets, rounded to integer. For stage-1 parameters (ψ_{10}, ψ_{11}), the 95% confidence intervals were constructed by m-out-of-n bootstrap percentiles. For stage-2 parameters $(\psi_{20}, \psi_{2A}, \psi_{22})$, the 95% confidence intervals were constructed by standard (n-out-of-n) bootstrap percentiles.

Web Table 8: Performance of estimated DTRs in Simulation 4

	n=500		n=2000	
	Value	Opt%	Value	Opt%
All	2.983 (0.064)	0.904 (0.059)	2.995 (0.031)	0.950 (0.028)
naive	2.012 (0.064)	0.255 (0.019)	2.006 (0.031)	0.252 (0.009)
CC	2.650 (0.174)	0.576 (0.114)	2.622 (0.122)	0.560 (0.070)
MI	2.348 (0.192)	0.450 (0.046)	2.466 (0.103)	0.479 (0.020)
WQ-EE	2.947 (0.085)	0.837 (0.097)	2.983 (0.034)	0.910 (0.049)

Note: Value, the value of the estimated regime; Opt%, the percentage of patients correctly classified to their true optimal treatments in all the stages; The numbers outside parentheses represent the sample means, while the numbers inside parentheses represent the sample standard errors.

Web Table 9: Correct classification rates in Simulation 4

	n=500		n=2000	
	Stages 1&2	Stage 3	Stages 1&2	Stage 3
All	0.919 (0.060)	0.980 (0.017)	0.958 (0.027)	0.990 (0.008)
naive	0.501 (0.021)	0.506 (0.021)	0.501 (0.011)	0.504 (0.010)
CC	0.594 (0.117)	0.958 (0.035)	0.568 (0.067)	0.981 (0.016)
MI	0.485 (0.040)	0.931 (0.058)	0.506 (0.016)	0.936 (0.026)
WQ-EE	0.873 (0.098)	0.958 (0.035)	0.928 (0.054)	0.981 (0.016)

Note: Stages 1&2, the correct classification rate at stages 1 and 2; Stage 3, the correct classification rate at stage 3; The numbers outside parentheses represent the sample means, while the numbers inside parentheses represent the sample standard errors.

Web Table 10: The missing proportion of hemodynamic variables 3 hours after ICU admission in the MIMIC-III data

covariates	missing proportion
heartrate at 3 hours	7.50 %
mean blood pressure at 3 hours	10.2 %
SpO2 at 3 hours	10.4 %
respiratory rate at 3 hours	10.2 %
temperature at 3 hours	19.8 %
urine output in 0-3 hours	0.00 %