Appendix

A Proof of Proposition 1

Proof. First, we show that an optimal λ^* for problem (7) exits. By the definition of the Lagrangian dual function, we have that $\mathcal{L}(\lambda)$ in (7) is an infimum of a collection of linear functions. Thus, it holds that $\mathcal{L}(\lambda)$ is a convex function. Also, it is not difficult to see that as $\lambda \to +\infty$, we have $\mathcal{L}(\lambda) \to +\infty$. Thus, together with the convexity of $\mathcal{L}(\lambda)$, we have that $\mathcal{L}(\lambda)$ has compact level sets. That is, for any $\alpha \in \mathbb{R}$, the set $\{\lambda : \mathcal{L}(\lambda) \leqslant \alpha\}$ is compact. By the Bolzano-Weistrass Theorem, there exists an optimal Lagrangian multiplier λ^* that minimizes $\mathcal{L}(\lambda)$.

Then, given the optimal Lagrangian multiplier λ^*. Since, by assumption, the primal solution $\widehat{\boldsymbol{\beta}}$ exists, we have that the function $\widehat{\mathcal{M}}(\boldsymbol{\beta})$ is bounded above. We have that a dual optimal solution $\widetilde{\boldsymbol{\beta}}$ exists. ■

B Proof of Theorem 2

Proof. In the proof, for ease of presentation, we let the constraint for the primal problem be $\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}) \geqslant q$.

We consider the case where the optimal Lagrangian multiplier $\lambda^* = 0$ or $\lambda^* > 0$. We first have that, if $\lambda^* = 0$, we have that $\boldsymbol{\beta}^* = \operatorname{argmax} \widehat{\mathcal{M}}(\boldsymbol{\beta})$ by (8). Since $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) \geqslant q$ by the feasibility of $\widetilde{\boldsymbol{\beta}}$, we have $\widetilde{\boldsymbol{\beta}}$ is also a primal optimal solution, and our clam holds.

If $\lambda^* > 0$, we show in Lemma 12 that one of the two cases hold

i. There exists a dual optimal solution such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) = q$.

ii. There exist at least two solutions achieve the dual optimal objective, denoted as $\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\beta}}'$, such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) < q$ and $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}') > q$.

Considering the two cases separately, for case (i), there exists a dual optimal solution $\widetilde{\boldsymbol{\beta}}$ such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) = q$. By the weak duality, we have

$$\widehat{\mathcal{M}}(\widehat{\boldsymbol{\beta}}) \geqslant \widehat{\mathcal{M}}(\widetilde{\boldsymbol{\beta}}) + \lambda^*\{q - \widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}})\} = \widehat{\mathcal{M}}(\widetilde{\boldsymbol{\beta}}),$$

and our claim holds as desired. Note that in this case, the dual optimal solution actually also achieves the primal optimality.

We then focus on case (ii). Given the multiplier λ^*, there exist multiple solutions achieve the dual optimality. Suppose that there are m of them. Let these solutions be $\boldsymbol{\beta}_{(1)},\ldots,\boldsymbol{\beta}_{(m)}$ be the sequence of solutions ranked by their corresponding primal objective values that

$$\widehat{\mathcal{M}}(\boldsymbol{\beta}_{(1)}) \leqslant \widehat{\mathcal{M}}(\boldsymbol{\beta}_{(2)}) \leqslant \cdots \leqslant \widehat{\mathcal{M}}(\boldsymbol{\beta}_{(m)}).$$

Meanwhile, by the dual optimality, we have that

$$\lambda^*\widehat{\mathcal{M}}(\boldsymbol{\beta}_{(1)}) + \lambda^*\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(1)}) = \widehat{\mathcal{M}}(\boldsymbol{\beta}_{(2)}) + \lambda^*\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(2)}) = \cdots = \widehat{\mathcal{M}}(\boldsymbol{\beta}_{(m)}) + \lambda^*\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(m)}).$$

Since $\lambda^* > 0$, we have

$$\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(1)}) \geqslant \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(2)}) \geqslant \cdots \geqslant \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(m)}).$$

Meanwhile, by our assumption, we have that there exists some $k \in [m]$ such that

$$\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(k)}) \geqslant q \geqslant \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(k+1)}).$$

This shows that there exists a dual solution, $\boldsymbol{\beta}_{(k+1)}$ in this case, that satisfies the primal constraint, and the duality gap is upper bounded by $\widetilde{\lambda}(\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(k+1)} - q)$. Note that by the discrete nature of the sample quantile function $\widehat{\mathcal{Q}}_\tau(\cdot)$, the primal solution's corresponding sample quantile value is $\widehat{\mathcal{Q}}_\tau(\widehat{\boldsymbol{\beta}})$, which might be different from q. We thus have, the duality bound can be bounded by $\widetilde{\lambda}\big\{\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}_{(k+1)}) - \widehat{\mathcal{Q}}_\tau(\widehat{\boldsymbol{\beta}})\big\}$, which concludes our proof. ∎

Lemma 12. *For the dual problem* (8)*, suppose that the optimal Lagrangian multiplier* $\lambda^* > 0$. *One of the following two cases must hold that*

i. There exists a dual optimal solution such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) = q$.

ii. There exist at least two solutions achieve the dual optimal objective, denoted as $\widetilde{\boldsymbol{\beta}}$ and $\widetilde{\boldsymbol{\beta}}'$, such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) < q$ and $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}') > q$.

Proof. We prove the lemma by contradiction. We assume the contrary that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) < q$ for all dual optimal solutions $\widetilde{\boldsymbol{\beta}}$ that achieve the dual optimal objective. (Note that the other case $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) < q$ follows by similar arguments.) We have that

$$\begin{aligned}\mathcal{L}(\lambda^*) &= \underset{\boldsymbol{\beta}}{\text{maximize}}\ \widehat{\mathcal{M}}(\boldsymbol{\beta}) + \lambda^*\{q - \widehat{\mathcal{Q}}_\lambda(\boldsymbol{\beta})\} \\ &= \underset{\ell\in[n]}{\text{maximize}}\ \widehat{\mathcal{M}}(\boldsymbol{\beta}^{(\ell)}) + \lambda^*\{q - \widehat{\mathcal{Q}}_\lambda(\boldsymbol{\beta}^{(\ell)})\},\end{aligned}$$

where $\boldsymbol{\beta}^{(\ell)} = \text{argmax}_{\boldsymbol{\beta}:\mathcal{Q}_\tau(\boldsymbol{\beta})=y_\ell} \widehat{\mathcal{M}}(\boldsymbol{\beta})(\boldsymbol{\beta})$, by the fact that $\mathcal{Q}_\tau(\boldsymbol{\beta}) = y_i$ for some $i \in [n]$.

By our assumption that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}})$ is strictly less than q. As shown in Lemma 13, we have that for small $\varepsilon > 0$, we have

$$\begin{aligned}\mathcal{L}(\lambda^* + \varepsilon) &= \underset{\ell\in[n]}{\text{maximize}}\ \widehat{\mathcal{M}}(\boldsymbol{\beta}) + (\lambda^* + \varepsilon)\{q - \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta})\} \\ &< \underset{\ell\in[n]}{\text{maximize}}\ \widehat{\mathcal{M}}(\boldsymbol{\beta}) + \lambda^*\{q - \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta})\} \\ &= \mathcal{L}(\lambda^*).\end{aligned}$$

However, since λ^* is the optimal Lagrangian multiplier by our assumption, it minimizes the function $\mathcal{L}(\lambda^*)$. The above result gives a contradiction, and our result holds as desired. ■

Lemma 13. *Suppose that the dual optimal Lagrangian multiplier $\lambda^* > 0$. Let $\boldsymbol{\beta}^{(\ell)} = \text{argmax}\{\widehat{\mathcal{M}}(\boldsymbol{\beta}) : \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}) = y_\ell\}$ for $\ell = 1, ..., n$. With loss of generality, assume $y_1 < y_2 < \cdots < y_n$. It holds that if $0 < \varepsilon < \min_{\ell\in\{2,...,n\}} \{\mathcal{M}(\boldsymbol{\beta}^{(\ell-1)}) - \mathcal{M}(\boldsymbol{\beta}^{(\ell)}\}$,*

$$\mathcal{L}(\lambda^* + \varepsilon) = \underset{\ell\in[n]}{\textit{minimize}} \widehat{\mathcal{M}}(\boldsymbol{\beta}) + (\lambda^* + \varepsilon)\{q - \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta})\}.$$

Proof. When we perturb the optimal λ^* to $\lambda^* + \varepsilon$, the corresponding dual solution becomes

$$\widetilde{\boldsymbol{\beta}}' = \underset{\boldsymbol{\beta}}{\text{argmax}}\, \widehat{\mathcal{M}}(\boldsymbol{\beta}) + (\lambda^* + \varepsilon)\{q - \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta})\}.$$

By our choice of ε, it is not difficult to see that our claim holds as desired. ■

C Proof of Proposition 4

Proof. For ease of presentation, we denote by $\breve{M}(\breve{\boldsymbol{\beta}})$ and $\breve{Q}_\tau(\breve{\boldsymbol{\beta}})$ the sample mean and τ-th quantile of the treatment effects following deterministic decision rule $f_i = \mathbb{1}(\mathbf{x}_i^\top \breve{\boldsymbol{\beta}} > 0)$.

We first prove that there exists a $\widetilde{\boldsymbol{\beta}}$ that if we follow the stochastic ITR $f(\mathbf{x}_i, \widetilde{\boldsymbol{\beta}}) = \mathbb{P}(f_i = 1) = \{1 + \exp(-\mathbf{x}_i^\top \widetilde{\boldsymbol{\beta}})\}^{-1}$, the corresponding objective can be arbitrarily close to the objective achieved by the deterministic ITR $f_i = \mathbb{1}(\mathbf{x}_i^\top \breve{\boldsymbol{\beta}} > 0)$, and the quantile constraint is approximately satisfied by the stochastic ITR. We have that for any $\breve{\boldsymbol{\beta}}$ and $\delta > 0$, there exists some $\widetilde{\boldsymbol{\beta}}$ such that $\left|\mathbb{1}(\mathbf{x}_i^\top \breve{\boldsymbol{\beta}} > 0) - \left\{1 + \exp(-\mathbf{x}_i^\top \widetilde{\boldsymbol{\beta}})\right\}^{-1}\right| \leqslant \delta$ for all $\mathbf{x}_i$. (Note that here we implicitly assume that $\mathbf{x}_i \neq 0$. If we indeed have some $\mathbf{x}_i = \mathbf{0}$, we may perturb the data by letting all $\mathbf{x}_i' = \mathbf{x}_i + \delta$ for some δ such that all $\mathbf{x}_i' \neq 0$.) This implies that by considering stochastic ITRs that $f(\mathbf{x}_i, \boldsymbol{\beta}) = \mathbb{P}(f_i = 1) = \{1 + \exp(-\mathbf{x}_i^\top \boldsymbol{\beta})\}^{-1}$, we have for any given $\varepsilon_1 > 0$, there exists some $\widetilde{\boldsymbol{\beta}}$, such that the corresponding objective satisfies $\widehat{\mathcal{M}}(\widetilde{\boldsymbol{\beta}}) > \breve{\mathcal{M}}(\breve{\boldsymbol{\beta}}) - \varepsilon_1$ and the corresponding quantile constraint satisfies $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) > \breve{\mathcal{Q}}_\tau(\breve{\boldsymbol{\beta}}) - \varepsilon_1$.

In addition, we have that by our assumptions that all outcomes are bounded, and $\breve{\boldsymbol{\beta}}$ achieves the quantile constraint in population. Also, as shown above, for any $\varepsilon_1 > 0$, there exists some $\widetilde{\boldsymbol{\beta}}$ such that $\widehat{\mathcal{Q}}_\tau(\widetilde{\boldsymbol{\beta}}) > \breve{\mathcal{Q}}_\tau(\breve{\boldsymbol{\beta}}) - \varepsilon_1$. We thus have that if n is large enough, problem (6) is feasible. Note that as $n \to \infty$ both $\widehat{\mathcal{M}}(\boldsymbol{\beta})$ and $\breve{\mathcal{M}}(\boldsymbol{\beta})$ converge to $\mathcal{M}(\boldsymbol{\beta}) = E(Y^*(\boldsymbol{\beta}))$. Meanwhile, we have $\widehat{\mathcal{M}}(\widetilde{\boldsymbol{\beta}}) > \breve{\mathcal{M}}(\breve{\boldsymbol{\beta}}) - \varepsilon_1$. We then have for any $\varepsilon_2 > 0$, the solution to our problem (6), $\widehat{\boldsymbol{\beta}}$, satisfies that $\mathcal{M}(\widehat{\boldsymbol{\beta}}) \geqslant \mathcal{M}(\breve{\boldsymbol{\beta}}) - \varepsilon_1 - \varepsilon_2$ with probability approaching one, and satisfies the quantile constraint in (6). Since ε_1 and ε_2 are arbitrary, our claim follows as desired. ■

D Proof of Theorem 5

Proof. Denote by $\mathbb{P}_n$ the empirical measure of the observed samples. Let $\boldsymbol{\beta}^*$ be a minimizer to the loss function under the quantile constraint in expectation that

$$\boldsymbol{\beta}^* = \underset{\boldsymbol{\beta} \in \mathcal{B}}{\operatorname{argmax}}\, \mathcal{M}(\boldsymbol{\beta}), \text{ subject to } \mathcal{Q}_\tau(\boldsymbol{\beta}) \geqslant q.$$

First, we have that as $\mathcal{Q}_\tau(\boldsymbol{\beta}^*) \geqslant q$, by Theorem 1 of Wang et al. (2018), it is not difficult to see that, as n increases, $\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}^*) \geqslant q - C \cdot n^{-1/2}$ for some constant C with probability goes to 1. Thus, we have that as n increases, $\boldsymbol{\beta}^*$ is a feasible point for problem (6) with probability goes to 1.

Meanwhile, by the definition that $\widehat{\boldsymbol{\beta}}$ is the maximizer for the empirical mean function under the constraint, we have that for any n large enough, $\widehat{\mathcal{M}}(\widehat{\boldsymbol{\beta}}) \geqslant \widehat{\mathcal{M}}(\boldsymbol{\beta}^*)$ for all $\boldsymbol{\beta}^* \in \mathcal{B}$ and satisfies $\widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}^*) \geqslant q - C \cdot n^{-1/2}$ with high probability. Thus, we only need to prove that $\widehat{\mathcal{M}}(\widehat{\boldsymbol{\beta}}) \to \mathcal{M}(\widehat{\boldsymbol{\beta}})$ in probability.

By our assumption that $\boldsymbol{\beta} \in \mathcal{B}$, and $\mathcal{B}$ is compact, we have that $\widehat{\boldsymbol{\beta}}$ is bounded. This implies that $\{\widehat{\mathcal{M}}(\boldsymbol{\beta}) : \boldsymbol{\beta} \in \mathcal{B}\}$ belongs to a Donsker class because it is not difficult to see $\widehat{\mathcal{M}}(\boldsymbol{\beta})$ is Lipschitz continuous with respect to $\boldsymbol{\beta}$. Consequently, we have

$$\sqrt{n}\big\{\widehat{\mathcal{M}}(\widehat{\boldsymbol{\beta}}) - \mathcal{M}(\widehat{\boldsymbol{\beta}})\big\} = \mathcal{O}_P(1).$$

Our claim holds as desired. ■

E Proof of Theorem 6

Proof. The proof is based on an application of Theorem 5.6 of Steinwart et al. (2007). Specifically, let $\mathcal{G}$ be the function class

$$\mathcal{G} = \big\{\widehat{\mathcal{M}}(\boldsymbol{\beta}) - \widehat{\mathcal{M}}(\boldsymbol{\beta}^*) : \boldsymbol{\beta} \in \mathcal{Q}_\tau(q)\big\},$$

where $\boldsymbol{\beta}^* \in \operatorname{argmax}_{\boldsymbol{\beta} \in \mathcal{Q}_\tau(q)} \mathcal{M}(\boldsymbol{\beta})$, and $\mathcal{Q}_\tau(q) = \{\boldsymbol{\beta} : \mathcal{Q}_\tau(\boldsymbol{\beta}^*) \geqslant q\}$. We first have that $\mathbb{E}(g) \leqslant 0$ for any $g \in \mathcal{G}$ as $\boldsymbol{\beta}^*$ is a maximizer in expectation. Note that our loss function is Lipschitz conitnuous with respect to $\boldsymbol{\beta}$. Denote that Lipschitz constant as C_L, we have $|g| \leqslant C_L \|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|$. As we assume that $\boldsymbol{\beta} \in \mathcal{B}(M)$, we have $|g| \leqslant B = 2MC_L$. Consequently, squaring both sides and taking expectations, we have $\mathbb{E}(g^2) \leqslant \mathbb{E}(g) + 4B^2$.

Next, for the covering number $N\big(B^{-1}\mathcal{G}, \varepsilon, L_2(\mathbb{P}_n)\big)$, we have

$$\begin{aligned}\log N\big(B^{-1}\mathcal{G}, \varepsilon, L_2(\mathbb{P}_n)\big) &\leqslant \log N\big(B^{-1}\{\widehat{\mathcal{M}}(\boldsymbol{\beta}) : \boldsymbol{\beta} \in \mathcal{B}(M)\}, \varepsilon, L_2(\mathbb{P}_n)\big) \\ &\leqslant \log N\big(\mathcal{B}(M), B\varepsilon/C_L, , L_2(\mathbb{P}_n)\big) \\ &\leqslant \log N\big(\mathcal{B}(1), 2\varepsilon, L_2(\mathbb{P}_n)\big).\end{aligned}$$

Thus, by Theorem 2.1 of Steinwart et al. (2007), we have that for some constant C,

$$\sup_{\mathbb{P}_n} \log N\big(B^{-1}\mathcal{G}, \varepsilon, L_2(\mathbb{P}_n)\big) \leqslant C\varepsilon^{-2}.$$

Consequently, by Theorem 5.6 of Steinwart et al. (2007), there exists a constant C_S such that for all $n \geqslant 1$ and $\tau \geqslant 1$, we have that

$$\mathbb{P}^*\big(\mathcal{M}(\widehat{\boldsymbol{\beta}}) < \mathcal{M}(\boldsymbol{\beta}^*) - C_S\varepsilon(n, C_1, B, \tau)\big) \leqslant e^{-\tau},$$

where

$$\varepsilon(n, C_1, B, \tau) = B \cdot \left(\frac{1}{n} + \frac{4}{\sqrt{n}}\right) + (B + C_1)\frac{\tau}{n}.$$

Our claim holds as desired. ■

F Dynamic Treatment Regime with Intermediate Outcome

In this section, we extend the dynamic treatment regime discussed in Section 5 to the more general case where we observe intermediate outcome at each stage. Similar to Section 5, we consider 2-stage dynamic treatment regime for ease of presentation, and the methods and results for the general T-stage case can be easily generalized. We also assume that the data are from some SMART trial.

The main difference between the setup with intermediate outcome is that after the first stage, we observe an intermediate outcome $Y_i^{(1)}$ for sample i, and after the second stage, we observe an outcome $Y_i^{(2)}$. Let $\boldsymbol{H}_i^{(1)} = \boldsymbol{X}_i^{(1)}$ and $\boldsymbol{H}_i^{(2)} = (\boldsymbol{X}_i^{(1)\top}, A_i^{(1)}, Y_i^{(1)}, \boldsymbol{X}_i^{(2)\top})^\top$. We consider candidate stochastic F-ITR indexed by $\boldsymbol{\beta} = \{\boldsymbol{\beta}^{(1)}, \boldsymbol{\beta}^{(2)}\}$ such that $f_j(\boldsymbol{H}_i^{(j)}, \boldsymbol{\beta}^{(j)}) = \mathbb{P}(A_i^{(j)} = 1|\boldsymbol{H}_i^{(j)}) = \{1 + \exp(-\boldsymbol{H}_i^{(j)\top}\boldsymbol{\beta}^{(j)})\}$ for $j = 1, 2$.

Suppose we have random samples $\{\mathbf{x}_i^{(1)}, a_i^{(1)}, y_i^{(1)}, \mathbf{x}_i^{(2)}, a_i^{(2)}, y_i^{(2)}\}_{i\in[n]}$, and we let $\mathbf{h}_i^{(1)} = \mathbf{x}_i^{(1)}$, and $\mathbf{h}_i^{(2)} = (\mathbf{x}_i^{(1)\top}, a_i^{(1)}, y_i^{(1)}, \mathbf{x}_i^{(2)\top})^\top$. We consider a backward fitting approach to estimating the optimal F-ITR. Specifically, letting $c_i^{(2)}(\boldsymbol{\beta}^{(2)}) = a_i^{(2)} f^{(2)}(\mathbf{h}_i^{(2)}, \boldsymbol{\beta}^{(2)}) + (1 - a_i^{(2)})\{1 - f^{(2)}(\mathbf{h}_i^{(2)}, \boldsymbol{\beta}^{(2)})\}$, we estimate the regime for stage 2 by

$$\widehat{\boldsymbol{\beta}}^{(2)} \in \operatorname{argmax} \widehat{\mathcal{M}}^{(2)}(\boldsymbol{\beta}^{(2)}), \text{ subject to } \widehat{\mathcal{Q}}_{\tau_2}^{(2)}(\boldsymbol{\beta}^{(2)}) \geqslant q - C_2/\sqrt{n}, \tag{13}$$

where $\widehat{\mathcal{M}}^{(2)}(\boldsymbol{\beta}^{(2)})$ and $\widehat{\mathcal{Q}}_{\tau_2}^{(2)}(\boldsymbol{\beta}^{(2)})$ are the estimators for the mean and τ_2-th quantile of outcome in stage 2 that

$$\widehat{\mathcal{M}}^{(2)}(\boldsymbol{\beta}^{(2)}) = \operatorname{argmin}_\mu n^{-1} \sum_{i=1}^n c_i^{(2)}(\boldsymbol{\beta}^{(2)})(y_i^{(2)} - \mu)^2,$$

and

$$\widehat{\mathcal{Q}}^{(2)}(\boldsymbol{\beta}^{(2)}) = \operatorname{argmin}_q n^{-1} \sum_{i=1}^n c_i^{(2)}(\boldsymbol{\beta}^{(2)})\rho_{\tau_2}(y_i^{(2)} - q),$$

and C_2 is a constant.

After getting $\widehat{\boldsymbol{\beta}}^{(2)}$, we estimate the regime for stage 1. First, similar to (11), we let

$$\begin{aligned} c_i^{(1)}(\boldsymbol{\beta}^{(1)}) = & \frac{a_i^{(1)} a_i^{(2)}}{\pi_1 \pi_2} \cdot f_i^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)}) f_i^{(2)}(\mathbf{h}_i^{(2)}, \widehat{\boldsymbol{\beta}}^{(2)}) \\ & + \frac{a_i^{(1)}(1 - a_i^{(2)})}{\pi_1(1 - \pi_2)} \cdot f_i^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)})\big(1 - f_i^{(2)}(\mathbf{h}_i^{(2)}, \widehat{\boldsymbol{\beta}}^{(2)})\big) \\ & + \frac{(1 - a_i^{(1)}) a_i^{(2)}}{(1 - \pi_1)\pi_2} \cdot \big(1 - f_i^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)})\big) f_i^{(2)}(\mathbf{h}_i^{(2)}, \widehat{\boldsymbol{\beta}}^{(2)}) \\ & + \frac{(1 - a_i^{(1)})(1 - a_i^{(2)})}{(1 - \pi_1)(1 - \pi_2)} \cdot \big(1 - f_i^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)})\big)\big(1 - f_i^{(2)}(\mathbf{h}_i^{(2)}, \widehat{\boldsymbol{\beta}}^{(2)})\big). \end{aligned}$$

Then, we estimate the F-ITR at stage 1 by

$$\widehat{\boldsymbol{\beta}}^{(1)} = \operatorname*{argmax}_{\boldsymbol{\beta}^{(1)}} \widehat{\mathcal{M}}(\boldsymbol{\beta}^{(1)}), \text{ subject to } \widehat{\mathcal{Q}}_{\tau_1}^{(1)}(\boldsymbol{\beta}^{(1)}) \geqslant q_1 - C_1/\sqrt{n}, \text{ and } \widehat{\mathcal{Q}}_\tau(\boldsymbol{\beta}^{(1)}) \geqslant q - C/\sqrt{n}, \tag{14}$$

where

$$\widehat{\mathcal{M}}(\boldsymbol{\beta}^{(1)}) = \text{argmin}_{\mu} n^{-1} \sum_{i=1}^{n} c_i^{(1)}(\boldsymbol{\beta}^{(1)})(y_i^{(1)} + y_i^{(2)} - \mu)^2$$

is the estimator of the mean of total outcome, and

$$\widehat{\mathcal{Q}}_{\tau}(\boldsymbol{\beta}^{(1)}) = \text{argmin}_{q} n^{-1} \sum_{i=1}^{n} c_i^{(1)}(\boldsymbol{\beta}^{(1)})\rho_{\tau}(y_i^{(1)} + y_i^{(2)} - q)$$

is the estimator for the τ-th quantile of the total outcome, and

$$\widehat{\mathcal{Q}}_{\tau_1}^{(1)}(\boldsymbol{\beta}^{(1)}) = \text{argmin}_{q} n^{-1} \sum_{i=1}^{n} c_i(\boldsymbol{\beta}^{(1)})\rho_{\tau_1}(y_i^{(1)} - q),$$

where $c_i(\boldsymbol{\beta}^{(1)}) = a_i^{(1)} f^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)}) + (1 - a_i^{(1)})\{1 - f^{(1)}(\mathbf{h}_i^{(1)}, \boldsymbol{\beta}^{(1)})\}$, is the estimator for the τ_1-th quantile of the stage 1 intermediate outcome.

For the estimator $\widehat{\boldsymbol{\beta}} = \{\widehat{\boldsymbol{\beta}}^{(1)}, \widehat{\boldsymbol{\beta}}^{(2)}\}$ derived above, we can get similar $\mathcal{O}_P(n^{-1/2})$ rate of convergence to the optimal risk $\mathcal{M}(\boldsymbol{\beta}^*)$, while satisfying the quantile constraints by backward induction and similar arguments in the proof of Theorem 6.

Theorem 14. *Suppose that $\boldsymbol{\beta}^* = \{\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*\}$ belongs to a compact set $\mathcal{B}(M)$, where $M > 0$ is a constant. Then we have that for all $\tau \geqslant 1$ we have*

$$\mathbb{P}^*\big(\mathcal{M}(\widehat{\boldsymbol{\beta}}) \geqslant \mathcal{M}(\boldsymbol{\beta}^*) - \varepsilon\big) \geqslant 1 - e^{-\tau},$$

where $\mathbb{P}^*$ denotes the outer probability for possibly nonmeasureable sets, and $\varepsilon = \mathcal{O}(n^{-1/2})$. Or, equivalently,

$$|\mathcal{M}(\widehat{\boldsymbol{\beta}}) - \mathcal{M}(\boldsymbol{\beta}^*)| = \mathcal{O}_P(n^{-1/2}).$$

In addition, we have that, with probability goes to 1,

$$\mathcal{Q}_{\tau}^{(1)}(\widehat{\boldsymbol{\beta}}^{(1)}) \geqslant q_1, \;\; \mathcal{Q}_{\tau}^{(2)}(\widehat{\boldsymbol{\beta}}^{(2)}) \geqslant q_2, \;\; \textit{and } \mathcal{Q}_{\tau}(\widehat{\boldsymbol{\beta}}) \geqslant q,$$

where $\mathcal{Q}_{\tau_1}^{(1)}(\boldsymbol{\beta}^{(1)})$ denotes the τ_1-th quantile of the stage 1 intermediate outcome, $\mathcal{Q}_{\tau_2}^{(2)}(\boldsymbol{\beta}^{(2)})$ denotes the τ_2-th quantile of the stage 2 intermediate outcome, and $\mathcal{Q}_{\tau}(\widehat{\boldsymbol{\beta}})$ denotes the τ-th quantile of the total outcome.