A Proofs

A.1 Proof of Theorem 1

1. **Consistency of $\widehat{\boldsymbol{\lambda}}$.**

   Let us define

$$\boldsymbol{\psi}_n(\boldsymbol{\lambda}) = \frac{1}{N}\left\{\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k)\boldsymbol{X}_k - \sum_{k\in P}\boldsymbol{X}_k\right\},$$
$$\boldsymbol{\psi}(\boldsymbol{\lambda}) = \mathbb{E}\left\{\left(R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) - 1\right)\boldsymbol{X}_k\right\}.$$

   Let $\Psi_n(\boldsymbol{\lambda}) = -||\boldsymbol{\psi}_n(\boldsymbol{\lambda})||$ and $\Psi(\boldsymbol{\lambda}) = -||\boldsymbol{\psi}(\boldsymbol{\lambda})||$. Hence, $\widehat{\boldsymbol{\lambda}} \in \arg\max_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \Psi_n(\boldsymbol{\lambda})$. We check the conditions of Theorem 5.9 in van der Vaart (2000). First,

$$\mathbb{E}(\boldsymbol{\psi}_n(\boldsymbol{\lambda})|\mathcal{U}) = \frac{1}{N}\sum_{k\in P}\left[R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) - 1\right]\boldsymbol{X}_k,$$

   which implies that $\mathbb{E}\{\boldsymbol{\psi}_n(\boldsymbol{\lambda})\} = \boldsymbol{\psi}(\boldsymbol{\lambda})$. Since $R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) - 1$ is bounded, it follows that $\left\{R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) - 1\right\}\boldsymbol{X}_k$ admits second-order moments. Then, $\mathbb{V}[\mathbb{E}(\boldsymbol{\psi}_n(\boldsymbol{\lambda})|\mathcal{U})]$ converges to 0. Also,

$$\mathbb{E}\left[\mathbb{V}\{\boldsymbol{\psi}_n(\boldsymbol{\lambda})|\mathcal{U}\}\right] = \mathbb{E}\left[\mathbb{V}\left\{\frac{1}{N}\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k)\boldsymbol{X}_k\middle|\mathcal{U}\right\}\right]. \tag{25}$$

   By Assumption 3-(i), the right-hand side of (25) tends to 0. Then $\boldsymbol{\psi}_n(\boldsymbol{\lambda}) \to \boldsymbol{\psi}(\boldsymbol{\lambda})$ in L^2, and thus in probability.

   Because $\boldsymbol{Z}_k$ has compact support and $\boldsymbol{\Lambda}$ is compact, there exists a compact interval I including with probability one the interval $[\min_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \boldsymbol{\lambda}^\top \boldsymbol{Z}_k, \max_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \boldsymbol{\lambda}^\top \boldsymbol{Z}_k]$. Moreover, $F(.)$ is uniformly continuous on I. Now, fix ε and let δ be such that for any $(a,b)\in I^2$, $|a-b|<\delta$ implies $|F(a)-F(b)|<\varepsilon$. Let C be such that $\|\boldsymbol{Z}_k\| \le C$ with probability one. Consider balls of center $\boldsymbol{\lambda}_b$ and of radius δ/C for $\boldsymbol{\lambda}$. Then, for $\boldsymbol{\lambda}$ within such a ball, we get, by the triangular and Cauchy-Schwarz inequalities,

$$\begin{aligned}\|\boldsymbol{\psi}_n(\boldsymbol{\lambda}) - \boldsymbol{\psi}_n(\boldsymbol{\lambda}_b)\| &\le \frac{1}{N}\sum_{k\in S} d_k R_k \left|F(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_b^\top \boldsymbol{Z}_k)\right| \|\boldsymbol{X}_k\| \\ &\le \frac{\varepsilon}{N}\sum_{k\in S} d_k R_k \|\boldsymbol{X}_k\|. \end{aligned} \tag{26}$$

Similarly, $\|\boldsymbol{\psi}(\boldsymbol{\lambda}) - \boldsymbol{\psi}(\boldsymbol{\lambda}_b)\| \leq \varepsilon \mathbb{E}(\|\boldsymbol{X}_k\|)$. Now, by assumption, $\boldsymbol{\Lambda}$ is compact. It can then be recovered by B balls of centers $\boldsymbol{\lambda}_b$ ($b = 1...B$) and of radius δ/C. Then, using $\big|||a|| - ||b||\big| \leq ||a-b||$, we get

$$\sup_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} |\Psi_n(\boldsymbol{\lambda}) - \Psi(\boldsymbol{\lambda})| \leq \sup_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} ||\boldsymbol{\psi}_n(\boldsymbol{\lambda}) - \boldsymbol{\psi}(\boldsymbol{\lambda})||$$
$$\leq \max_{b=1...B} \|\boldsymbol{\psi}_n(\boldsymbol{\lambda}_b) - \boldsymbol{\psi}(\boldsymbol{\lambda}_b)\| + \varepsilon \left\{ \mathbb{E}(\|\boldsymbol{X}_k\|) + \frac{1}{N}\sum_{k\in S} d_k R_k \|\boldsymbol{X}_k\| \right\}.$$

The first term on the right-hand side converges in probability to zero by pointwise consistency of $\boldsymbol{\psi}_n(\boldsymbol{\lambda})$. By Assumption 3-(i) and reasoning as above, the term within brackets converges in probability to $\mathbb{E}\{(1+R_k)\|\boldsymbol{X}_k\|\}$. Therefore, with probability tending to one, the left-hand side is smaller than $\varepsilon[1 + 2\mathbb{E}\{(1+R_k)\|\boldsymbol{X}_k\|\}]$. Because ε was arbitrary, we have proved

$$\sup_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} |\Psi_n(\boldsymbol{\lambda}) - \Psi(\boldsymbol{\lambda})| \xrightarrow{P} 0.$$

Hence, condition (i) in Theorem 5.9 of van der Vaart (2000) holds.

We now check condition (ii). First, by Assumptions 1 and 2,

$$\mathbb{E}\left[\left(R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1\right) \boldsymbol{X}_k | \boldsymbol{Z}_k\right] = \left[F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\mathbb{E}\left(R_k|\boldsymbol{Z}_k\right) - 1\right] \mathbb{E}\left(\boldsymbol{X}_k|\boldsymbol{Z}_k\right) = 0.$$

Thus, $\boldsymbol{\psi}(\boldsymbol{\lambda}_0) = 0$ and $\Psi(\boldsymbol{\lambda}_0) = 0$. Suppose that there exists $\boldsymbol{\lambda}_1$ such that $\Psi(\boldsymbol{\lambda}_1) = 0$. Then $\boldsymbol{\psi}(\boldsymbol{\lambda}_1) = 0$ and because $\mathbb{E}(\boldsymbol{X}_k|R_k = 1, \boldsymbol{Z}_k) = \boldsymbol{\Gamma}\boldsymbol{Z}_k$, we obtain, by the law of iterated expectation,

$$\boldsymbol{\Gamma}\, \mathbb{E}\left[R_k \left\{F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k)\right\} \boldsymbol{Z}_k\right] = 0.$$

Because the rank of $\boldsymbol{\Gamma}$ is equal to $\dim(\boldsymbol{Z}_k)$,

$$\mathbb{E}\left[R_k \left\{F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k)\right\} \boldsymbol{Z}_k\right] = 0.$$

This, in turn, implies that

$$\mathbb{E}\left[R_k \left(F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k)\right) \left(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k - \boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k\right)\right] = 0. \tag{27}$$

Now, because $F(.)$ is strictly increasing, we have

$$\left\{F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k)\right\} \left(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k - \boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k\right) \geq 0 \text{ with equality iff } \boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k - \boldsymbol{\lambda}_1^\top \boldsymbol{Z}_k = 0.$$

Hence, (27) implies that $(\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_1)^\top \boldsymbol{Z}_k = 0$ almost surely. This and the fact that $\mathbb{E}(\boldsymbol{Z}_k \boldsymbol{Z}_k^\top)$ is nonsingular imply that $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_0$.

Thus, $\Psi(\boldsymbol{\lambda}) = 0$ implies that $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$. Second, by the same argument following (26), we have, for any $\boldsymbol{\lambda}, \boldsymbol{\lambda}'$ such that $\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\| < \delta/C$,

$$\|\Psi(\boldsymbol{\lambda}) - \Psi(\boldsymbol{\lambda}')\| \leq \|\boldsymbol{\psi}(\boldsymbol{\lambda}) - \boldsymbol{\psi}(\boldsymbol{\lambda}')\| \leq \varepsilon \mathbb{E}\left(\|\boldsymbol{X}_k\|\right).$$

Hence, Ψ is continuous. Thus, for any $\varepsilon' > 0$,

$$\inf_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}: \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_0\| \geq \varepsilon} \|\Psi(\boldsymbol{\lambda})\| = \min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}: \|\boldsymbol{\lambda} - \boldsymbol{\lambda}_0\| \geq \varepsilon} \|\Psi(\boldsymbol{\lambda})\| > 0 = \|\Psi(\boldsymbol{\lambda}_0)\|$$

and condition (ii) in Theorem 5.9 of van der Vaart (2000) holds.

As a result, both conditions of this theorem are satisfied, and $\widehat{\boldsymbol{\lambda}}$ is consistent.

Consistency of $\widehat{t}_C$.

First,

$$\begin{aligned}\left(\widehat{t}_C - t_y\right)/N = & \frac{1}{N}\sum_{k \in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) Y_k - \frac{1}{N}\sum_{k \in P} Y_k \\ & + \frac{1}{N}\sum_{k \in S} d_k R_k \left\{F(\widehat{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\right\} Y_k. \end{aligned} \tag{28}$$

We now show that both terms on the right-hand side of (28), denoted by A_1 and A_2 hereafter, tend to zero in probability. For A_1, we use arguments similar to those used for the pointwise consistency of $\boldsymbol{\psi}_n(\lambda)$. We have

$$\mathbb{E}(A_1 | \mathcal{U}) = \frac{1}{N}\sum_{k \in P} \left\{R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1\right\} Y_k.$$

Moreover, by the law of iterated expectation and (10),

$$\mathbb{E}\left\{\left[R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1\right] Y_k\right\} = \mathbb{E}\left\{\left[\mathbb{E}(R_k \mid \boldsymbol{Z}_k) F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1\right] \mathbb{E}\left[Y_k | \boldsymbol{Z}_k\right]\right\} = 0.$$

Hence, $\mathbb{E}(A_1) = 0$ and $\mathbb{V}\{\mathbb{E}(A_1|\mathcal{U})\}$ converges to 0. Moreover,

$$\mathbb{E}\left\{\mathbb{V}(A_1 | \mathcal{U})\right\} = \mathbb{E}\left\{\mathbb{V}\left(\frac{1}{N}\sum_{k \in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) Y_k \middle| \mathcal{U}\right)\right\},$$

and the right-hand side converges to 0 by Assumption 3-(i). Thus, $\mathbb{V}(A_1)$ tends to 0 and A_1 converges to 0 in probability.

Turning to A_2, note that $F(.)$ is uniformly continuous on the compact set. Set $\varepsilon > 0$ and δ as above. By consistency of $\widehat{\boldsymbol{\lambda}}$, $||\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0|| < \delta/C$ with probability close to one. Then, by the Cauchy-Schwarz inequality, $\max_k ||(\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0)^\top \boldsymbol{Z}_k|| < \delta$, implying that

$$\max_k \left| F(\widehat{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) \right| < \varepsilon,$$

with a probability close to one. Hence, with such a probability,

$$|A_2| < \varepsilon \left(\frac{1}{N} \sum_{k \in S} d_k |Y_k| \right).$$

By Assumption 3-(i), the term into parentheses converges to $\mathbb{E}(|Y_k|)$. Therefore, A_2 converges to zero, and the result follows. $\square$

2. **Linearization of $\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0$.**

Let us define $\widehat{\mathbf{G}}(\boldsymbol{\lambda}) = \frac{1}{N} \sum_{k \in S} d_k R_k F'(\boldsymbol{\lambda}^\top \boldsymbol{Z}_k) \boldsymbol{X}_k \boldsymbol{Z}_k^\top$. Then, by the first-order condition of (5) and the mean value theorem,

$$\begin{aligned} 0 &= \frac{\widehat{\mathbf{G}}(\widehat{\boldsymbol{\lambda}})^\top}{N} \left\{ \sum_{k \in S} d_k R_k F(\widehat{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k) \boldsymbol{X}_k - \sum_{k \in P} \boldsymbol{X}_k \right\} \\ &= \frac{\widehat{\mathbf{G}}(\widehat{\boldsymbol{\lambda}})^\top}{N} \left\{ \sum_{k \in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) \boldsymbol{X}_k - \sum_{k \in P} \boldsymbol{X}_k \right\} + \widehat{\mathbf{G}}(\widehat{\boldsymbol{\lambda}})^\top \widehat{\mathbf{G}}(\widetilde{\boldsymbol{\lambda}})(\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0), \end{aligned} \tag{29}$$

where $\widetilde{\boldsymbol{\lambda}} = \widetilde{t}\boldsymbol{\lambda}_0 + (1 - \widetilde{t})\widehat{\boldsymbol{\lambda}}$ for some $\widetilde{t} \in [0, 1]$.

Because $F'(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)$ is bounded, we have, by the same arguments as when showing $\boldsymbol{\psi}_n(\boldsymbol{\lambda}) \xrightarrow{P} \boldsymbol{\psi}(\boldsymbol{\lambda})$,

$$\widehat{\mathbf{G}}(\boldsymbol{\lambda}_0) \xrightarrow{P} \mathbf{G}.$$

Now fix $\varepsilon > 0$. F' is continuous, and therefore uniformly continuous on the interval I defined above. Thus, there exists $\delta_2 > 0$ such that for any $(a, b) \in I^2$, $|a - b| < \delta_2$ implies $|F'(a) - F'(b)| < \varepsilon$. By Step 1, with a large probability, $\|\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0\| < \delta_2/C$

and $\|\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0\| < \delta_2/C$. Then, by the triangular and Cauchy-Schwarz inequality, with a large probability,

$$\left\|\widehat{\mathbf{G}}(\widehat{\boldsymbol{\lambda}}) - \widehat{\mathbf{G}}(\boldsymbol{\lambda}_0)\right\| \leq \frac{\varepsilon}{N}\sum_{k\in S} d_k R_k \|\boldsymbol{X}_k \boldsymbol{Z}_k^\top\|.$$

The same holds if $\widehat{\boldsymbol{\lambda}}$ is replaced with $\widetilde{\boldsymbol{\lambda}}$. Because $\boldsymbol{Z}_k$ is bounded, the right-hand side converges to 0 by Assumption 3-(i). Hence, both $\widehat{\mathbf{G}}(\widehat{\boldsymbol{\lambda}})$ and $\widehat{\mathbf{G}}(\widetilde{\boldsymbol{\lambda}})$ converge in probability to $\mathbf{G}$. This and (29) imply that

$$\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0 = -\frac{(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top}{N}\left\{\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\boldsymbol{X}_k - \sum_{k\in P}\boldsymbol{X}_k\right\}\{1 + o_P(1)\}.$$

Asymptotic normality of $\widehat{t}_C$.

First, we have

$$\begin{aligned}
\frac{\widehat{t}_C}{N} &= \frac{1}{N}\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) Y_k + \frac{1}{N}\sum_{k\in S} d_k R_k \left\{F(\widehat{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\right\} Y_k \\
&= \frac{1}{N}\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) Y_k - \left\{\frac{1}{N}\sum_{k\in S} d_k R_k Y_k F'(\overline{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k)\boldsymbol{Z}_k^\top\right\} \\
&\quad \times \frac{(\mathbf{G}^\top\mathbf{G})^{-1}\mathbf{G}^\top}{N}\left\{\sum_{k\in S} d_k R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\boldsymbol{X}_k - \sum_{k\in P}\boldsymbol{X}_k\right\}\{1 + o_P(1)\},
\end{aligned}$$

where $\overline{\boldsymbol{\lambda}} = \bar{t}\boldsymbol{\lambda}_0 + (1-\bar{t})\widehat{\boldsymbol{\lambda}}$ for some $\bar{t} \in [0,1]$. By the same argument as above,

$$\frac{1}{N}\sum_{k\in S} d_k R_k Y_k F'(\overline{\boldsymbol{\lambda}}^\top \boldsymbol{Z}_k)\boldsymbol{Z}_k^\top \xrightarrow{P} \mathbb{E}(\rho_k Y_k \boldsymbol{Z}_k^\top).$$

Hence, since $\boldsymbol{\gamma} = \mathbf{G}(\mathbf{G}^\top\mathbf{G})^{-1}\mathbb{E}(\rho_k Y_k \boldsymbol{Z}_k^\top)$ and $W_k = (R_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1)(Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k)$, we get

$$\widehat{t}_C - t_y = \left\{\sum_{k\in S} d_k(W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k) - \sum_{k\in P} W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k + \sum_{k\in P} W_k\right\}\{1 + o_P(1)\}. \tag{30}$$

To prove the result, we now check the conditions of Theorem 2 in Chen and Rao (2007). Note first that $\mathcal{U}$ plays the role of $\mathcal{B}_n$ in their theorem, $\sum_{k\in S} d_k(W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k) - \sum_{k\in P} W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k$ corresponds to U_n, $\sum_{k\in P} W_k$ corresponds to V_n, $\sigma_{2N} \equiv$

$\mathbb{V}\left\{\sum_{k\in S} d_k(W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k)\middle|\mathcal{U}\right\}^{1/2}$ corresponds to σ_{2n} and $\sigma_{1N} \equiv \mathbb{V}\left(\sum_{k\in P} W_k\right)^{1/2}$ corresponds to σ_{1n}.

Then, by Assumptions 1 and 2,

$$\mathbb{E}(W_k|\boldsymbol{Z}_k) = \left[\mathbb{E}(R_k|\boldsymbol{Z}_k)F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k) - 1\right]\mathbb{E}\left(Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k|\boldsymbol{Z}_k\right) = 0.$$

Hence, $\mathbb{E}(W_k) = 0$ and by the central limit theorem,

$$\sigma_{1N}^{-1}\left(\sum_{k\in P} W_k\right) \xrightarrow{d} \mathcal{N}(0,1).$$

Also, $\sum_{k\in P} W_k$ is $\mathcal{U}$-measurable. Hence, their condition 1 holds.

Next,

$$\mathbb{E}\left\{\sum_{k\in S} d_k(W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k) - \sum_{k\in P} W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k \middle| \mathcal{U}\right\} = 0$$

and Condition (1) in Chen and Rao (2007) holds by Assumption 3-(iii) and Polya's theorem (see the remark below Theorem 2 in Chen and Rao, 2007). Thus, their condition 2 holds. Finally their condition 3 holds by Assumption 3-(ii).

Therefore,

$$\frac{\sum_{k\in S} d_k(W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k) - \sum_{k\in P} W_k + Y_k - \boldsymbol{\gamma}^\top \boldsymbol{X}_k + \sum_{k\in P} W_k}{\sqrt{\sigma_{1N}^2 + \sigma_{2N}^2}} \xrightarrow{d} \mathcal{N}(0,1).$$

The result follows by (30) and Slutsky's lemma. □

A.2 Proof of Theorem 2

1. By assumption, a solution to (14) exists. By the same reasoning as the one used to show $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}_1$ in the proof of Theorem 1, the solution is unique. Moreover, still reasoning as in the first step of the proof of Theorem 1, we have $\widehat{\boldsymbol{\lambda}} \xrightarrow{P} \boldsymbol{\lambda}_\infty$. □

2. We can decompose the total error of estimation of $\widehat{t_C}$ as in (13). Using Assumption 3 and the same arguments as in the second step of the proof of Theorem 1, the first three terms of the right-hand side converge to 0 in probability. On the other hand, the

fourth term on the right-hand side does not tend to zero in probability. By the law of large number and Assumption 3-(i), it converges towards

$$\mathbb{E}\left[R_k\ \left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\right\}\ Y_k\right].$$

This in turn can be rewritten as:

$$\mathbb{E}\left\{R_k\ \left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\right\}\ Y_k\right\} \;=\; \mathbb{E}\left(f_k\ R_k\ Y_k\ \boldsymbol{Z}_k^\top\right)\ (\boldsymbol{\lambda}_\infty - \boldsymbol{\lambda}_0)\,,$$

where f_k is defined in (16). Next, we prove (15). We have

$$\begin{aligned}
-\mathbb{E}\left[f_k F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)^{-1} X_k \boldsymbol{Z}_k^\top(\boldsymbol{\lambda}_\infty - \boldsymbol{\lambda}_0)\right] &= \mathbb{E}\left\{\left(1 - \frac{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k)}{F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)}\right) X_k\right\} \\
&= \mathbb{E}\left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k)\,(R_k X_k - \mathbb{E}(R_k|Z_k) X_k)\right\} \\
&= \mathbb{E}\left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k)\ \mathbb{C}ov\,(\boldsymbol{X}_k,\ R_k \mid \boldsymbol{Z}_k)\right\},
\end{aligned}$$

where the second equality comes from the nonresponse model and (14), and the third equality from the law of iterated expectation. This shows (15), which in turn implies that

$$\begin{aligned}
&\mathbb{E}\left\{R_k\ \left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k) - F(\boldsymbol{\lambda}_0^\top \boldsymbol{Z}_k)\right\}\ Y_k\right\} \\
= &\; -\ \left\{\left\{\mathbb{E}(f_k R_k \boldsymbol{Z}_k \boldsymbol{X}_k^\top)\right\}^{-1}\ \mathbb{E}(f_k R_k \boldsymbol{Z}_k Y_k)\right\}^\top\ \mathbb{E}\left\{F(\boldsymbol{\lambda}_\infty^\top \boldsymbol{Z}_k)\ \mathbb{C}ov\,(\boldsymbol{X}_k,\ R_k \mid \boldsymbol{Z}_k)\right\}. \square
\end{aligned}$$