BIOMETRICS 000, 000–000
000 XXXXX

DOI: 000

Web-based Supplementary Materials for "Evaluating Marker-Guided Treatment Selection Strategies"

Roland A. Matsouaka[1†,*], Junlong Li[2†,], and Tianxi Cai[2,***]**

[1]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA
[2]Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA
† Equal Contributors
**email:* rmatsoua@hsph.harvard.edu
***email:* juli@hsph.harvard.edu
****email:* tcai@hsph.harvard.edu

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

APPENDIX

Web Appendix A: Asymptotic Properties of $\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}}$ and $\widehat{\mathcal{I}}_{\text{calib}}$

Although the working models could be semi-parametric and $\widehat{\boldsymbol{\beta}}$ could be infinite dimensional, for the derivation of the asymptotic theory, we focus on the case when $\widehat{\boldsymbol{\beta}}$ is finite dimensional. Throughout we use notation $g'(x) = dg(x)/dx$ and $\mathbf{g}'_k(\mathbf{x}_1, ..., \mathbf{x}_K) = \partial g(\mathbf{x}_1, ..., \mathbf{x}_K)/\partial \mathbf{x}_k$. We assume that the covariates are bounded, $\bar{\boldsymbol{\beta}}$ is an interior point of a compact parameter space, $D(\bar{\boldsymbol{\beta}}, \mathbf{X})$ is a continuous random variable which has a continuously differentiable density function $f_D(\cdot)$ bounded away from 0 with finite support $[\tau_l, \tau_r]$, and $Y^{(k)} \mid D(\bar{\boldsymbol{\beta}}, \mathbf{X})$ has a finite second moment for any $D(\bar{\boldsymbol{\beta}}, \mathbf{X}) \in [\tau_l, \tau_r]$. We require the regularity condition that $\widehat{\boldsymbol{\beta}} \to \bar{\boldsymbol{\beta}} = (\bar{\boldsymbol{\beta}}_0^\intercal, \bar{\boldsymbol{\beta}}_1^\intercal)^\intercal$ for $\bar{\boldsymbol{\beta}}$ and $\widehat{\mathcal{W}}_{\boldsymbol{\beta}} = n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}}) = n^{-\frac{1}{2}} \sum_{i=1}^n \Psi_{\boldsymbol{\beta} i} + o_p(1)$ which converges weakly in distribution to a multivariate normal. We consider that $\bar{\Delta}(\cdot)$ is continuously differentiable with $|\bar{\Delta}'(s)| \leqslant b_{\bar{\Delta}'} < \infty$. We also assume that there are at most $K_0 < \infty$ number of solutions to $\bar{\Delta}'(s) = 0$, there exists intervals $\tau_l = s_1 < \cdots < s_{K_0+1} = \tau_r$ such that $\bar{\Delta}(s)$ is monotone in $[s_k, s_{k+1}]$, for $k = 1, ..., K_0$; and hence $\bar{\Delta}(s) = \xi$ has at most K_0 roots. Furthermore, when at least one solution to $\bar{\Delta}(s) = \xi$ exists, we denote by $\{\mathfrak{s}_{k_r} \in (s_{k_r}, s_{k_r+1}) : r = 1, ..., R_0\}$ all the solutions and assume that $|\bar{\Delta}'(\mathfrak{s}_{k_r})| \geqslant L_{\bar{\Delta}'} > 0$. This regularity condition ensures that the total variation of $\mathfrak{I}(s) = I\{\bar{\Delta}(s) \geqslant \xi\}$ is bounded by K_0.

We next derive the convergence properties of $\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}} = \widehat{\mathbb{V}}_a(\widehat{\Delta}, \widehat{\boldsymbol{\beta}}) + \bar{Y}_0$ to $\mathbb{V}_{\mathcal{I}_{\text{calib}}} = \mathbb{V}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + \mu_0$, where $\bar{Y}_0 = n_0^{-1} \sum_{i=1}^n Y_i I(G_i = 0)$, $\mu_0 = E(Y^{(0)})$,

$$\widehat{\mathbb{V}}_a(\Delta, \boldsymbol{\beta}) = n^{-1} \sum_{i=1}^n I\left[\Delta\{D(\boldsymbol{\beta}, \mathbf{X}_i)\} \geqslant \xi\right] \widehat{\mathcal{Y}}_{\xi_i}, \quad \mathbb{V}_a(\Delta, \boldsymbol{\beta}) = E\left(I\left[\Delta\{D(\boldsymbol{\beta}, \mathbf{X}_i)\} \geqslant \xi\right] \mathcal{Y}_{\xi_i}\right)$$

$\widehat{\mathcal{Y}}_{\xi_i} = (Y_i - \xi) I(G_i = 1)/\widehat{\pi}_1 - Y_i I(G_i = 0)/\widehat{\pi}_0$ and $\mathcal{Y}_{\xi_i} = (Y_i - \xi) I(G_i = 1)/\pi_1 - Y_i I(G_i = 0)/\pi_0$.

To decompose various sources of variation, we define $\widehat{W}_{\mathbb{V}}$ as

$$\widehat{W}_{\mathbb{V}} \equiv \widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}} - \mathbb{V}_{\mathcal{I}_{\text{calib}}} = \widehat{W}_{\Delta} + \widehat{\mathbb{V}}_a(\bar{\Delta}, \widehat{\boldsymbol{\beta}}) - \widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + \widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + \bar{Y}_0 - \mathbb{V}_{\mathcal{I}_{\text{calib}}}$$

where $\widehat{W}_\Delta \equiv \widehat{\mathbb{V}}_a(\widehat{\Delta}, \widehat{\boldsymbol\beta}) - \widehat{\mathbb{V}}_a(\bar{\Delta}, \widehat{\boldsymbol\beta}) = \widehat{W}_{\Delta 1} + \widehat{W}_{\Delta 2}$ with

$$\widehat{W}_{\Delta 1} = \int [I\{\widehat{\Delta}_\xi(s) \geqslant 0\} - I\{\Delta_\xi(s) \geqslant 0\}]\widehat{H}(\widehat{\boldsymbol\beta}; ds)$$

$$\widehat{W}_{\Delta 2} = \int [I\{\widehat{\Delta}_\xi(s) \geqslant 0\} - I\{\Delta_\xi(s) \geqslant 0\}]\Delta_\xi(s)\widehat{F}(\widehat{\boldsymbol\beta}; ds),$$

and $\widehat{H}(\boldsymbol\beta; s) = n^{-1}\sum_{i=1}^n I\{D(\boldsymbol\beta, \mathbf{X}_i) \leqslant s\}[\widehat{\mathcal{Y}}_{\xi_i} - \Delta_\xi\{D(\boldsymbol\beta, \mathbf{X}_i)\}]$, $\widehat{\Delta}_\xi(d) = \widehat{\Delta}(d) - \xi$, $\Delta_\xi(d) = E(\mathcal{Y}_{\xi_i} \mid D(\bar{\boldsymbol\beta}, \mathbf{X})_i = d) = \bar{\Delta}(d) - \xi$, and $\widehat{F}(\boldsymbol\beta; s) = n^{-1}\sum_{i=1}^n I\{D(\boldsymbol\beta, \mathbf{X}_i) \leqslant s\}$.

From $\sup_{\mathbf{x}} |D(\widehat{\boldsymbol\beta}, \mathbf{x}) - D(\bar{\boldsymbol\beta}, \mathbf{x})| = O_p(n^{-\frac{1}{2}})$ and similar arguments in Cai et al. (2011),

$$\widehat{\epsilon}_\Delta = \sup_s |\widehat{\Delta}(s) - \bar{\Delta}(s)| = o_p\{(nh)^{-1/2}\log(n)\}.$$

Now consider $\widehat{W}_{\Delta 1}$. Note that $n^{\frac{1}{2}}\left\{\widehat{H}(\boldsymbol\beta; s) - H(\boldsymbol\beta; s)\right\} = n^{-\frac{1}{2}}\sum_{i=1}^n H_i(\boldsymbol\beta; s) + o_p(1)$, where $H_i(\boldsymbol\beta; s) = I\{D(\boldsymbol\beta, \mathbf{X}_i) \leqslant s\}[\mathcal{Y}_{\xi_i} - \Delta_\xi\{D(\boldsymbol\beta, \mathbf{X}_i)\}] - m_1(\boldsymbol\beta, s)\{I(G_i = 1)/\pi_1 - 1\} + m_0(\boldsymbol\beta, s)\{I(G_i = 0)/\pi_0 - 1\}$ and $H(\boldsymbol\beta; s) = E[H_i(\boldsymbol\beta; s)]$, $m_j(\boldsymbol\beta; s) = E[I\{D(\boldsymbol\beta, \mathbf{X}_i) \leqslant s\}\mathcal{Y}_{\xi_i} \mid G_i = j]$. It follows from a functional central limit theorem that $n^{\frac{1}{2}}\left\{\widehat{H}(\boldsymbol\beta; s) - H(\boldsymbol\beta; s)\right\}$ converges weakly to a gaussian process and hence is equicontinuous in $\boldsymbol\beta$. Hence, $n^{\frac{1}{2}}\widehat{H}(\widehat{\boldsymbol\beta}; s) = n^{\frac{1}{2}}\{\widehat{H}(\widehat{\boldsymbol\beta}; s) - H(\bar{\boldsymbol\beta}; s)\} = n^{\frac{1}{2}}\left\{\widehat{H}(\bar{\boldsymbol\beta}; s) - H(\bar{\boldsymbol\beta}; s)\right\} + \dot{\mathbf{H}}_1(\bar{\boldsymbol\beta}; s)^\intercal n^{\frac{1}{2}}(\widehat{\boldsymbol\beta} - \bar{\boldsymbol\beta})$ which converges weakly to a zero-mean gaussian process. Therefore, for any $\delta = o_p(1)$,

$$\widehat{\epsilon}_H(\delta) = \sup_s \left|\widehat{H}(\widehat{\boldsymbol\beta}; s) - \widehat{H}(\widehat{\boldsymbol\beta}; s + \delta)\right| = o_p(n^{-\frac{1}{2}}),$$

and consequently $|\widehat{W}_{\Delta 1}| \leqslant \int I\{\bar{\Delta}(s) - \xi \in [-\widehat{\epsilon}_\Delta, \widehat{\epsilon}_\Delta]\}|\widehat{H}(\widehat{\boldsymbol\beta}; ds)|$.If $\inf_s |\bar{\Delta}(s) - \xi| \geqslant \delta_0 > 0$, then the equation $\bar{\Delta}(s) = \xi$ has no solution. This implies that $P\{\inf_s |\widehat{\Delta}(s) - \xi| \geqslant \delta_0/2\} \to 1$ and $I\left\{\bar{\Delta}(s) - \xi \in [-\widehat{\epsilon}_\Delta, \widehat{\epsilon}_\Delta]\right\} \xrightarrow{\mathcal{P}} 0$ in probability. Thus, $P(\widehat{W}_{\Delta 1} = 0) \to 1$ and $n^{\frac{1}{2}}\widehat{W}_{\Delta 1} \xrightarrow{\mathcal{P}} 0$.

For the setting where the solution to $\bar{\Delta}(s) = \xi$ exists, we have,

$$|\widehat{W}_{\Delta 1}| \leqslant \sum_{r=1}^{R_0}\int_{s_{k_r}}^{s_{k_r+1}} I\left\{s \in [\bar{\Delta}_{k_r}^{-1}(\xi - \widehat{\epsilon}_\Delta), \bar{\Delta}_{k_r}^{-1}(\xi + \widehat{\epsilon}_\Delta)]\right\}\left|\widehat{H}(\widehat{\boldsymbol\beta}; ds)\right|$$

$$\leqslant \sum_{r=1}^{R_0}\widehat{\epsilon}_H\left\{\left|\bar{\Delta}_{k_r}^{-1}(\xi - \widehat{\epsilon}_\Delta) - \bar{\Delta}_{k_r}^{-1}(\xi + \widehat{\epsilon}_\Delta)\right|\right\}$$

where $\bar{\Delta}_{k_r}^{-1}(\cdot)$ denotes the inverse function of $\bar{\Delta}$ within $[s_{k_r}, s_{k_r+1}]$. Given that $|\bar{\Delta}'(\mathfrak{s}_{k_r})| \geqslant L_{\Delta'}$ and $\widehat{\epsilon}_\Delta = o_p(1)$, $\bar{\Delta}_{k_r}^{-1}(\xi - \widehat{\epsilon}_\Delta) - \bar{\Delta}_{k_r}^{-1}(\xi + \widehat{\epsilon}_\Delta) = o_p(1)$, we have $\widehat{W}_{\Delta 1} = o_p(n^{-\frac{1}{2}})$.

For $\widehat{W}_{\Delta 2}$, since $\widehat{F}(\boldsymbol{\beta}; s)$ is a monotone function in s,

$$
\begin{aligned}
|\widehat{W}_{\Delta 2}| &\leqslant \int \left| I\{\widehat{\Delta}_\xi(s) \geqslant 0\} - I\{\Delta_\xi(s) < 0\} \right| |\Delta_\xi(s)| \widehat{F}(\widehat{\boldsymbol{\beta}}; ds) \\
&\leqslant \int I\{\Delta_\xi(s) \in [-\widehat{\epsilon}_\Delta, \widehat{\epsilon}_\Delta]\} \widehat{\epsilon}_\Delta \widehat{F}(\widehat{\boldsymbol{\beta}}; ds) = \widehat{\epsilon}_\Delta \left\{ \widehat{F}_\Delta(\widehat{\boldsymbol{\beta}}; \widehat{\epsilon}_\Delta) - \widehat{F}_\Delta(\widehat{\boldsymbol{\beta}}; -\widehat{\epsilon}_\Delta) \right\}
\end{aligned}
$$

where $\widehat{F}_\Delta(\boldsymbol{\beta}; d) = n^{-1} \sum_{i=1}^n I[\Delta\{D(\boldsymbol{\beta}, \mathbf{X}_i)\} \leqslant d]$.

It follows from a functional central limit theorem that $n^{\frac{1}{2}}\{\widehat{F}_\Delta(\boldsymbol{\beta}; d) - F_\Delta(\boldsymbol{\beta}; d)\}$ converges to a zero-mean Gaussian process and hence is equicontinuous in $\boldsymbol{\beta}$ and d. We have $\widehat{F}_\Delta(\widehat{\boldsymbol{\beta}}; \widehat{\epsilon}_\Delta) - \widehat{F}_\Delta(\widehat{\boldsymbol{\beta}}; -\widehat{\epsilon}_\Delta) = F_\Delta(\bar{\boldsymbol{\beta}}; \widehat{\epsilon}_\Delta) - F_\Delta(\bar{\boldsymbol{\beta}}; \widehat{\epsilon}_\Delta) + o_p(n^{-\frac{1}{2}})$. This implies $|\widehat{W}_{\Delta 2}| \leqslant 2\widehat{\epsilon}_\Delta^2 \sup_s |f_\Delta(\bar{\boldsymbol{\beta}}; s)| + o_p(n^{-\frac{1}{2}})$. Provided that $\widehat{\epsilon}_\Delta = o_p(n^{-1/4})$, we have $\widehat{W}_{\Delta 2} = o_p(n^{-\frac{1}{2}})$. Hence, $\widehat{W}_\Delta = o_p(n^{-\frac{1}{2}})$. The above results also indicate that $P\{\widehat{\mathcal{I}}_{\text{calib}}(\mathbf{X}) \neq \mathcal{I}_{\text{calib}}(\mathbf{X})\} \to 0$.

Next, we establish the the asymptotic convergence of $\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}}$ to $\mathbb{V}_{\mathcal{I}_{\text{calib}}}$. By the law of large numbers, $\widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + \bar{Y}_0 - \mathbb{V}_{\mathcal{I}_{\text{calib}}} \xrightarrow{\mathcal{P}} 0$. It remains to show that $\widehat{\mathbb{V}}_a(\bar{\Delta}, \widehat{\boldsymbol{\beta}}) - \widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) = o_p(1)$. It follows from similar arguments as given for the convergence of $\widehat{H}(\boldsymbol{\beta}; s)$ that $n^{\frac{1}{2}}\{\widehat{\mathbb{V}}_a(\bar{\Delta}, \boldsymbol{\beta}) - \mathbb{V}_a(\bar{\Delta}, \boldsymbol{\beta})\} = n^{-\frac{1}{2}} \sum_{i=1}^n \Psi_{ai}(\boldsymbol{\beta}) + o_p(1)$, which converges weakly to a zero-mean Gaussian process, where

$$
\Psi_{ai}(\boldsymbol{\beta}) = I[\bar{\Delta}_\xi\{D(\boldsymbol{\beta}; \mathbf{X}_i)\} \geqslant 0]\mathcal{Y}_{\xi_i} - m_{\Delta 1}(\boldsymbol{\beta}) \tfrac{I(G_i=1)}{\pi_1} - + m_{\Delta 0}(\boldsymbol{\beta}) \tfrac{I(G_i=0)}{\pi_0}
$$

and $m_{\Delta j}(\boldsymbol{\beta}) = E(I[\bar{\Delta}_\xi\{D(\boldsymbol{\beta}; \mathbf{X}_i)\} \geqslant 0]\mathcal{Y}_{\xi_i} \mid G_i = j)$. Hence,

$$
\widehat{\mathbb{V}}_a(\bar{\Delta}, \widehat{\boldsymbol{\beta}}) - \widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) = \mathbb{V}_a(\bar{\Delta}, \widehat{\boldsymbol{\beta}}) - \mathbb{V}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + o_p(n^{-\frac{1}{2}}) = \mathcal{V}'_{a2}(\bar{\Delta}, \boldsymbol{\beta})^\intercal (\widehat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}}) + o_p(n^{-\frac{1}{2}}).
$$

Therefore, $\widehat{\mathbb{V}}_a(\bar{\Delta}, \widehat{\boldsymbol{\beta}}) - \widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) = o_p(1)$ and thus $\widehat{\mathcal{W}}_{\mathbb{V}} = o_p(1)$, indicating the consistency of $\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}}$ for $\mathbb{V}_{\mathcal{I}_{\text{calib}}}$. For the weak convergence of $\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}}$, we note from the above expansions that

$$
\begin{aligned}
n^{\frac{1}{2}}(\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}} - \mathbb{V}_{\mathcal{I}_{\text{calib}}}) &= n^{\frac{1}{2}} \mathcal{V}'_{a2}(\bar{\Delta}, \boldsymbol{\beta})^\intercal (\widehat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}}) + n^{\frac{1}{2}}\{\widehat{\mathbb{V}}_a(\bar{\Delta}, \bar{\boldsymbol{\beta}}) + \bar{Y}_0 - \mathbb{V}_{\mathcal{I}_{\text{calib}}}\} + o_p(1) \\
&= n^{-\frac{1}{2}} \sum_{i=1}^n \Psi_{\mathcal{V}i} + o_p(1),
\end{aligned}
$$

where

$$\Psi_{\mathcal{V}i} = \mathcal{V}'_{a2}(\bar{\Delta}, \boldsymbol{\beta})^{\intercal}\Psi_{\boldsymbol{\beta}i} + \Psi_{ai}(\bar{\boldsymbol{\beta}}) + (Y_i - \mu_0)\frac{I(G_i = 0)}{\pi_0}. \quad \text{(A.1)}$$

It then follows from a central limit theorem that $n^{\frac{1}{2}}(\widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}} - \mathbb{V}_{\mathcal{I}_{\text{calib}}})$ converges in distribution to a normal with mean 0 and variance $\sigma^2_{\mathbb{V}} = E(\Psi^2_{\mathcal{V}i})$.

Web Appendix B: Asymptotic Properties of $\widehat{¢}$, $\widehat{c}$, and $\widehat{\mathbb{V}}_{\widehat{I}_{\widehat{c}}}$ when $\bar{\Delta}$ is increasing

In this section, we assume that $\bar{\Delta}(c)$ is monotone and there exists c_0 such that $\bar{\Delta}(c_0) = \xi$ and $\bar{\Delta}'(c_0) > 0$. The monotonicity implies that $V\{F^{-1}(u)\} = \mu_0 + \int_0^u \bar{\Delta}\{F^{-1}(v)\}dv$ is convex in u. Moreover, $\bar{\Delta}'(c_0)F'(c_0) > 0$ implies that $\mathbb{V}_{I_c}$ has a unique maximizer at c_0, where $I_c = I\{D(\bar{\boldsymbol{\beta}}; \mathbf{X}) \geqslant c\}$. From Newey and McFadden (1994), to show the consistency of $\widehat{¢}$, it suffices to show that $\widehat{\mathbb{V}}_{\widehat{I}_c} \xrightarrow{\mathcal{P}} \mathbb{V}_{I_c}$ uniformly in c. To this end, note that $\widehat{\mathbb{V}}(c, \boldsymbol{\beta}) = \widehat{\mathbb{V}}_a(c, \boldsymbol{\beta}) + \bar{Y}_0$, and $\widehat{\mathbb{V}}_a(c, \boldsymbol{\beta}) = n^{-1}\sum_{i=1}^{n} I\{D(\boldsymbol{\beta}; \mathbf{X}_i) \geqslant c\}\widehat{\mathcal{Y}}_{\xi_i}$. We have $\widehat{\mathbb{V}}_{\widehat{I}_c} = \widehat{\mathbb{V}}(c; \widehat{\boldsymbol{\beta}})$. It follows from the uniform law of large numbers (Pollard, 1990) that $\widehat{\mathbb{V}}(c; \boldsymbol{\beta})$ converges to $\mathbb{V}(c; \boldsymbol{\beta})$ uniformly in $\{\boldsymbol{\beta}, c\}$, where $\mathbb{V}(c; \boldsymbol{\beta}) = E[I\{D(\boldsymbol{\beta}; \mathbf{X}) \geqslant c\}\mathcal{Y}_{\xi}] + \mu_0$. Furthermore, by the functional central limit theorem (Pollard, 1990), $n^{\frac{1}{2}}\{\widehat{\mathbb{V}}(c, \boldsymbol{\beta}) - \mathbb{V}(c; \boldsymbol{\beta})\}$ converges weakly to a Gaussian process in $\boldsymbol{\beta}$ and c.

These results, together with the weak convergence of $n^{\frac{1}{2}}(\widehat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}})$, imply that

$$\sup_c \left| n^{\frac{1}{2}}\{\widehat{\mathbb{V}}_{\widehat{I}_c} - \mathbb{V}_{I_c}\} \right| = O_p(1) \quad \text{(B.1)}$$

The consistency of $\widehat{c}$ immediately follows from the monotonicity of $\bar{\Delta}(c)$, $\bar{\Delta}'(c_0) > 0$.

We next show that $\widehat{¢} - c^o = O_p(n^{-1/3})$ following the cubic-root asymptotic theory from Kim and Pollard (1990). For the ease of presentation, we assume that $\bar{\boldsymbol{\beta}}$ is known and hence

$$\widehat{¢} = \operatorname*{argmax}_c \widehat{\mathbb{V}}_{\widehat{I}_c} = \operatorname*{argmax}_c \{\widehat{\mathbb{V}}_{\widehat{I}_c} - \bar{Y}_0\} = n^{-1}\sum_{i=1}^{n} \mathbb{Y}_i I(D(\bar{\boldsymbol{\beta}}, \mathbf{X}_i) \geqslant c) + O_p(n^{-1})$$

where $\mathbb{Y}_i = \mathcal{Y}_{\xi_i} - \mu_1\{I(G_i = 1)/\pi_1 - 1\} + \mu_0\{I(G_i = 0)/\pi_0 - 1\}$. To verify the regularity conditions of the main theorem in §1.1 of Kim and Pollard (1990), we let $\eta_i = (\mathbb{Y}_i, D(\bar{\boldsymbol{\beta}}, \mathbf{X}_i))$, $g(\eta, c) = \mathbb{Y}_i I(D(\bar{\boldsymbol{\beta}}, \mathbf{X}_i) \geqslant c)$ and hence $n^{-1}\sum_{i=1}^{n} \mathbb{Y}_i I(D(\bar{\boldsymbol{\beta}}, \mathbf{X}_i) \geqslant c) = n^{-1}\sum_{i=1}^{n} g(\eta_i, c)$.

The main condition involves an envelope function for $g(\xi, c)$

$$G_R(\eta) = \sup\{g(\eta, c) : |c - c^o| \leqslant R\}$$

$$= \sup_{|c-c^o|\leqslant R} \left[\mathbb{Y}\{I(D(\bar{\boldsymbol\beta}, \mathbf{X}) \geqslant c) - I(D(\bar{\boldsymbol\beta}, \mathbf{X}) \geqslant c^o)\}\right] = |\mathbb{Y}| I(|D(\bar{\boldsymbol\beta}, \mathbf{X}) - c^o| < R)$$

It is easy to see that $E\{G_R(\eta_i)\} = E\{|\mathbb{Y}|I(|D(\bar{\boldsymbol\beta}, \mathbf{X}) - c^o| < R)\} = O(R)$ as $R \to 0$ and

$$E[G_R(\eta_i)I\{G_R(\eta_i) > \mathcal{K}\}] \leqslant E\{I(|D(\bar{\boldsymbol\beta}, \mathbf{X}) - c^o| < R)\mathbb{Y}^2 I(|\mathbb{Y}| > \mathcal{K})\}$$

$$\leqslant R\left[\max\{f_D(\cdot)\}E\{\mathbb{Y}^2 I(|\mathbb{Y}| > \mathcal{K}) \mid c^o - R < D(\bar{\boldsymbol\beta}, \mathbf{X}) < c^o + R\}\right].$$

Since $Y^{(k)}$ has a finite second moment given $D(\bar{\boldsymbol\beta}, \mathbf{X})$, $E\{\mathbb{Y}^2 I(|\mathbb{Y}| > \mathcal{K}) \mid c^o - R < D(\bar{\boldsymbol\beta}, \mathbf{X}) < c^o + R\} \to 0$ as $\mathcal{K} \to \infty$ and hence condition (vi) of Theorem §1.1 is verified. The rest of the regularity conditions follow mainly from the smoothness of $\Delta(\cdot)$ and f_D.

To derive the distribution of $\widehat{c}$, note that $\widehat{\Delta}_\xi(\widehat{c}; \widehat{\boldsymbol\beta}) = 0$, where $\widehat{\Delta}_\xi(c; \boldsymbol\beta) = \sum_{i=1}^n K_h\{D(\boldsymbol\beta; \mathbf{X}_i) - c\}\widehat{\mathcal{Y}}_{\xi i}/[\sum_{i=1}^n K_h\{D(\boldsymbol\beta; \mathbf{X}_i\}]$. From arguments as given in Cai et al. (2011), $\sup_c |\widehat{\Delta}_\xi(c; \widehat{\boldsymbol\beta}) - \widehat{\Delta}_\xi(c; \bar{\boldsymbol\beta})| = O_p(n^{-\frac{1}{2}})$ and $\sup_c |\widehat{\Delta}_\xi(c; \bar{\boldsymbol\beta}) - \bar{\Delta}_\xi(c; \bar{\boldsymbol\beta})| = O_p\{(nh)^{-1/2}\log(n)\}$. Then, by a taylor series expansion, $o = \widehat{\Delta}_\xi(\widehat{c}; \widehat{\boldsymbol\beta}) = \widehat{\Delta}_\xi(\widehat{c}; \bar{\boldsymbol\beta}) + O_p(n^{-\frac{1}{2}}) = \widehat{\Delta}_\xi(c_0; \bar{\boldsymbol\beta}) + (\widehat{c} - c_0)\widehat{\Delta}'_\xi(c_0; \bar{\boldsymbol\beta}) + O_p\{n^{-\frac{1}{2}} + (nh)^{-1}\log(n)^2\}$. It follows that

$$\widehat{c} - c_0 = \widehat{\Delta}_\xi(c_0; \bar{\boldsymbol\beta}) + o_p\{(nh)^{-1/2} + h^2\} = n^{-1}\sum_{i=1}^n K_h\{D(\boldsymbol\beta; \mathbf{X}_i) - c\}\frac{\mathcal{Y}_{\xi_i} - \bar{\Delta}_\xi(c_0; \bar{\boldsymbol\beta})}{F'(c)}$$

which is asymptotically normal as a typical kernel smoothed estimator.

For the asymptotic distribution of $\widehat{\mathbb{V}}_{\mathcal{I}_{\widehat{c}}} = \widehat{\mathbb{V}}(\widehat{c}, \widehat{\boldsymbol\beta})$, using similar arguments as given above, we can show that $\widehat{\mathbb{V}}(\widehat{c}, \widehat{\boldsymbol\beta}) - \widehat{\mathbb{V}}(c_0, \widehat{\boldsymbol\beta}) = o_p(n^{-\frac{1}{2}})$ and $\widehat{\mathbb{V}}(c_0; \widehat{\boldsymbol\beta}) - \widehat{\mathbb{V}}(c_0; \bar{\boldsymbol\beta}) = \mathbb{V}(c_0; \widehat{\boldsymbol\beta}) - \mathbb{V}(c_0; \bar{\boldsymbol\beta}) + o_p(n^{-\frac{1}{2}})$. Therefore,

$$\widehat{\mathbb{V}}_{\mathcal{I}_{\widehat{c}}} - \mathbb{V}_{\mathcal{I}_{\text{calib}}} = \mathbb{V}(c_0; \widehat{\boldsymbol\beta}) - \mathbb{V}(c_0; \bar{\boldsymbol\beta}) + \widehat{\mathbb{V}}(c_0; \bar{\boldsymbol\beta}) - \mathbb{V}_{\mathcal{I}_{\text{calib}}} + o_p(n^{-\frac{1}{2}})$$

$$= \mathbf{V}'_2(c_0; \bar{\boldsymbol\beta})^{\mathsf{T}}(\widehat{\boldsymbol\beta} - \bar{\boldsymbol\beta}) + \widehat{\mathbb{V}}(c_0; \bar{\boldsymbol\beta}) - \mathbb{V}_{\mathcal{I}_{\text{calib}}} + o_p(n^{-\frac{1}{2}}) = n^{-1}\sum_{i=1}^n \boldsymbol\Psi_{Vi}(c_0) + o_p(n^{-\frac{1}{2}})$$

where $\boldsymbol\Psi_{Vi}(c_0) = \mathbf{V}'_2(c_0; \bar{\boldsymbol\beta})^{\mathsf{T}}\boldsymbol\Psi_{\boldsymbol\beta i} + I(D_i \geqslant c_0)\mathcal{Y}_{\xi_i} - m_{\Delta 1}(\bar{\boldsymbol\beta})I(G_i = 1)/\pi_1 + m_{\Delta_0}(\bar{\boldsymbol\beta})I(G_i =$

$0)/\pi_0 + (Y_i - \mu_0)I(G_i = 0)/\pi_0$. Under the monotonicity of $\bar{\Delta}(\cdot)$, we see that $\mathbf{\Psi}_{Vi}(c_0) = \mathbf{\Psi}_{\mathcal{V}i}$ and hence $\widehat{\mathbb{V}}_{\widehat{I_c}} - \widehat{\mathbb{V}}_{\widehat{\mathcal{I}}_{\text{calib}}} = o_p(n^{-\frac{1}{2}})$.

References

Cai, T., Tian, L., Wong, P., and Wei, L. (2011). Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics* **12,** 270–282.

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *The Annals of Statistics* pages 191–219.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics* **4,** 2111–2245.

Pollard, D. (1990). Empirical processes: theory and applications. In *NSF-CBMS regional conference series in probability and statistics*. JSTOR.